

Chapter 3

Topology of Linear Spaces and Convex Sets

In this chapter, several basic results on topological linear spaces and convex sets are presented. We will characterize finite-dimensionality, metrizable, and normability of topological linear spaces. Among the important results are the Hahn–Banach Extension Theorem, the Separation Theorem, the Closed Graph Theorem, and the Open Mapping Theorem. We will also prove the Michael Selection Theorem, which will be applied in the proof of the Bartle–Graves Theorem.

3.1 Flats and Affine Functions

In this section, we present the basic properties of flats and affine functions. Let E be a linear space (over \mathbb{R}). We call $F \subset E$ a **flat**¹ if the straight line through every distinct two points of F is contained in F , i.e.,

$$(1 - t)x + ty \in F \text{ for each } x, y \in F \text{ and } t \in \mathbb{R}.$$

Evidently, the intersection and the product of flats are also flats. We have the following characterization of flats:

Proposition 3.1.1. *Let E be a linear space. For each non-empty subset $F \subset E$, the following conditions are equivalent:*

- (a) F is a flat;
- (b) For each $n \in \mathbb{N}$, if $v_1, \dots, v_n \in F$ and $\sum_{i=1}^n t_i = 1$, then $\sum_{i=1}^n t_i v_i \in F$;
- (c) $F - x$ is a linear subspace of E for any $x \in F$;
- (d) $F - x_0$ is a linear subspace of E for some $x_0 \in F$.

¹A flat is also called an **affine set**, a **linear manifold**, or a **linear variety**.

Proof. By induction on $n \in \mathbb{N}$, we can obtain (a) \Rightarrow (b). Condition (c) follows from the case $n = 3$ of (b) because, for each $x, y, z \in F$ and $a, b \in \mathbb{R}$,

$$a(y - x) + b(z - x) + x = (1 - a - b)x + ay + bz.$$

To see (c) \Rightarrow (a), let $x, y \in F$ and $t \in \mathbb{R}$. Since $F - x$ is a linear subspace of E by (c), we have $t(y - x) \in F - x$, which means $(1 - t)x + ty \in F$. The implication (c) \Rightarrow (d) is obvious.

(d) \Rightarrow (c): It suffices to show that if $F - x_0$ is a linear subspace of E , then $F - x = F - x_0$ for any $x \in F$. For every $z \in F$, we have

$$z - x = (z - x_0) - (x - x_0) \in F - x_0.$$

Here, take $z' \in F$ so that $(z - x_0) + (x - x_0) = z' - x_0$. Then, we have

$$z - x_0 = (z' - x_0) - (x - x_0) = z' - x \in F - x.$$

Consequently, we have $F - x = F - x_0$. □

In the proof of the implication (d) \Rightarrow (c), we actually proved the following:

Corollary 3.1.2. *Let F be a flat in a linear space E . Then, $F - x = F - y$ for any $x, y \in F$.* □

A maximal proper flat $H \subsetneq E$ is called a **hyperplane** in E . The following proposition shows the relationship between hyperplanes and linear functionals.

Proposition 3.1.3. *Let E be a linear space.*

- (1) *For each hyperplane $H \subset E$, there is a linear functional $f : E \rightarrow \mathbb{R}$ such that $H = f^{-1}(s)$ for some $s \in \mathbb{R}$;*
- (2) *For each non-trivial linear functional $f : E \rightarrow \mathbb{R}$ and $s \in \mathbb{R}$, $f^{-1}(s)$ is a hyperplane in E ;*
- (3) *For linear functionals $f_1, f_2 : E \rightarrow \mathbb{R}$, if $f_1^{-1}(s_1) = f_2^{-1}(s_2)$ for some $s_1, s_2 \in \mathbb{R}$, then $f_2 = rf_1$ for some $r \in \mathbb{R}$.*

Proof. (1): For a given $x_0 \in H$, $H_0 = H - x_0$ is a maximal proper linear subspace of E (Proposition 3.1.1). Let $x_1 \in E \setminus H_0$. For each $x \in E$, there exists a unique $t \in \mathbb{R}$ such that $x - tx_1 \in H_0$. Indeed, $E = H_0 + \mathbb{R}x_1$ because of the maximality of H_0 . Hence, we can write $x = z + tx_1$ for some $z \in H_0$ and $t \in \mathbb{R}$. Then, $x - tx_1 \in H_0$. Moreover, if $x - t'x_1 \in H_0$ and $t' \in \mathbb{R}$, then $(t - t')x_1 \in H_0$. Since $x_1 \notin H_0$, it follows that $t = t'$. Therefore, we have a function $f : E \rightarrow \mathbb{R}$ such that $x - f(x)x_1 \in H_0$. For each $x, y \in E$ and $a, b \in \mathbb{R}$,

$$(ax + by) - (af(x) + bf(y))x_1 = a(x - f(x)x_1) + b(y - f(y)x_1) \in H_0,$$

which means $f(ax + by) = af(x) + bf(y)$, i.e., f is linear. Observe that $f^{-1}(0) = H_0 = H - x_0$, hence it follows that $H = f^{-1}(f(x_0))$.

(2): From the non-triviality of f , it follows that $f(E) = \mathbb{R}$, and hence $\emptyset \subsetneq f^{-1}(s) \subsetneq E$. A simple calculation shows that $f^{-1}(s)$ is a flat. To prove the maximality, let $F \subset E$ be a flat with $f^{-1}(s) \subsetneq F$. Take $x_0 \in f^{-1}(s)$ and $x_1 \in F \setminus f^{-1}(s)$. Since $f(x_1) \neq f(x_0)$ and F is a flat, it follows that $f(F) = \mathbb{R}$. For each $x \in E$, we can choose $y \in F \setminus f^{-1}(s)$ so that $f(y) \neq f(x)$. Note that $s = tf(x) + (1-t)f(y)$ for some $t \in \mathbb{R} \setminus \{0\}$. Let $z = tx + (1-t)y \in f^{-1}(s) \subset F$. Then, $x = (1-t^{-1})y + t^{-1}z \in F$. Accordingly, we have $F = E$.

(3): When $f_1^{-1}(s_1) = f_2^{-1}(s_2) = \emptyset$, both f_1 and f_2 are trivial (i.e., $f_1(E) = f_2(E) = \{0\}$), and hence $f_1 = f_2$. If $f_1^{-1}(s_1) = f_2^{-1}(s_2) \neq \emptyset$, take $x_0 \in f_1^{-1}(s_1) = f_2^{-1}(s_2)$. Then, it follows that

$$f_1^{-1}(0) = f_1^{-1}(s_1) - x_0 = f_2^{-1}(s_2) - x_0 = f_2^{-1}(0).$$

Let $H_0 = f_1^{-1}(0) = f_2^{-1}(0)$ and $x_1 \in E \setminus H_0$. Analogous to (1), each $x \in E$ can be uniquely written as $x = y + tx_1$, where $y \in H_0$ and $t \in \mathbb{R}$. Then, $f_1(x) = tf_1(x_1)$ and $f_2(x) = tf_2(x_1)$, hence $f_2(x) = f_1(x)f_1(x_1)^{-1}f_2(x_1)$. Let $r = f_1(x_1)^{-1}f_2(x_1)$. It follows that $f_2 = rf_1$. \square

It is said that finitely many distinct points $v_1, \dots, v_n \in E$ are **affinely** (or **geometrically**) **independent** provided that, for $t_1, \dots, t_n \in \mathbb{R}$,

$$\sum_{i=1}^n t_i v_i = \mathbf{0}, \quad \sum_{i=1}^n t_i = 0 \Rightarrow t_1 = \dots = t_n = 0,$$

i.e., $v_1 - v_n, \dots, v_{n-1} - v_n$ are linearly independent. In this case, the subset $\{v_1, \dots, v_n\} \subset E$ is also said to be **affinely** (or **geometrically**) **independent**. An (infinite) subset $A \subset E$ is said to be **affinely** (or **geometrically**) **independent** if every finite subset of A is affinely independent. This condition is equivalent to the condition that $(A - v) \setminus \{\mathbf{0}\}$ is linearly independent for some/any $v \in A$.²

The smallest flat containing $A \subset E$ is called the **flat hull**³ of A and is denoted by $\text{fl } A$. Then, $\mathbb{R}^n = \text{fl}\{\mathbf{0}, \mathbf{e}_1, \dots, \mathbf{e}_n\}$, where $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is the canonical orthonormal basis for \mathbb{R}^n (i.e., $\mathbf{e}_i(i) = 1$ and $\mathbf{e}_i(j) = 0$ for $j \neq i$). Observe that

$$\begin{aligned} \text{fl}\{v_1, \dots, v_n\} &= \left\{ \sum_{i=1}^n t_i v_i \mid \sum_{i=1}^n t_i = 1 \right\} \text{ and} \\ \text{fl } A &= \bigcup \left\{ \text{fl}\{x_1, \dots, x_n\} \mid n \in \mathbb{N}, x_1, \dots, x_n \in A \right\}. \end{aligned}$$

²The phrase “for some/any” means that we can choose one of “some” or “any” in the sentence. By this choice, we have two different conditions. The condition using “some” is weaker than the condition using “any” in general. However, these two conditions can be equivalent in a certain situation.

³The flat hull is also called the **affine hull**.

By Zorn's Lemma, every non-empty subset $A \subset E$ contains a maximal affinely independent subset $A_0 \subset A$. Then, $\text{fl } A_0 = \text{fl } A$ and each $x \in \text{fl } A$ can be uniquely written as $x = \sum_{i=1}^n t_i v_i$, where $v_1, \dots, v_n \in A_0$ and $t_1, \dots, t_n \in \mathbb{R} \setminus \{0\}$ such that $\sum_{i=1}^n t_i = 1$. In fact, for some/any $v \in A_0$, $(A_0 - v) \setminus \{0\}$ ($= (A_0 \setminus \{v\}) - v$) is a Hamel basis for the linear subspace $\text{fl } A - v$ ($= \text{fl } A_0 - v$) of E .

The **dimension** of a flat $F \subset E$ is denoted by $\dim F$, and is defined by the dimension of the linear space $F - x$ for some/any $x \in F$, i.e., $\dim F = \dim(F - x)$. When $\dim F = n$ (resp. $\dim F < \infty$ or $\dim F = \infty$), it is said that F is **n -dimensional** (resp. **finite-dimensional** (abbrev. **f.d.**) or **infinite-dimensional** (abbrev. **i.d.**)). Therefore, every n -dimensional flat $F \subset E$ contains $n + 1$ points v_1, \dots, v_{n+1} such that $F = \text{fl}\{v_1, \dots, v_{n+1}\}$. In this case, v_1, \dots, v_{n+1} are affinely independent. Conversely, if $F = \text{fl}\{v_1, \dots, v_{n+1}\}$ for some $n + 1$ affinely independent points $v_1, \dots, v_{n+1} \in F$, then $\dim F = n$.

Let F and F' be flats in linear spaces E and E' , respectively. A function $f : F \rightarrow F'$ is said to be **affine** if it satisfies the following condition:

$$f((1-t)x + ty) = (1-t)f(x) + tf(y) \text{ for each } x, y \in F \text{ and } t \in \mathbb{R},$$

which is equivalent to the following:

$$f\left(\sum_{i=1}^n t_i v_i\right) = \sum_{i=1}^n t_i f(v_i)$$

$$\text{for each } n \in \mathbb{N}, v_i \in F, t_i \in \mathbb{R} \text{ with } \sum_{i=1}^n t_i = 1.$$

Recall that $F \subset E$ is a flat if and only if $F - x_0$ is a linear subspace of E for some/any $x_0 \in F$ (Proposition 3.1.1).

Proposition 3.1.4. *Let $f : F \rightarrow F'$ be a function between flats F and F' in linear spaces E and E' , respectively. In order that f is affine, it is necessary and sufficient that the following $f^{x_0} : F - x_0 \rightarrow F' - f(x_0)$ is linear for some/any $x_0 \in F$:*

$$f^{x_0}(x) = f(x + x_0) - f(x_0) \text{ for each } x \in F - x_0.$$

Proof. (Necessity) For each $x, y \in F - x_0$ and $a, b \in \mathbb{R}$,

$$\begin{aligned} f^{x_0}(ax + by) &= f(ax + by + x_0) - f(x_0) \\ &= f(a(x + x_0) + b(y + x_0) + (1 - a - b)x_0) - f(x_0) \\ &= af(x + x_0) + bf(y + x_0) + (1 - a - b)f(x_0) - f(x_0) \\ &= a(f(x + x_0) - f(x_0)) + b(f(y + x_0) - f(x_0)) \\ &= af^{x_0}(x) + bf^{x_0}(y). \end{aligned}$$

(Sufficiency) For each $x, y \in F$ and $t \in \mathbb{R}$,

$$\begin{aligned}
 f((1-t)x + ty) &= f^{x_0}((1-t)x + ty - x_0) + f(x_0) \\
 &= f^{x_0}((1-t)(x - x_0) + t(y - x_0)) + f(x_0) \\
 &= (1-t)f^{x_0}(x - x_0) + tf^{x_0}(y - x_0) + f(x_0) \\
 &= (1-t)(f^{x_0}(x - x_0) + f(x_0)) + t(f^{x_0}(y - x_0) + f(x_0)) \\
 &= (1-t)f(x) + tf(y). \quad \square
 \end{aligned}$$

Proposition 3.1.5. *Let A be a non-empty affinely independent subset of a linear space E . Then, every function $g : A \rightarrow E'$ to another linear space E' uniquely extends to an affine function $\tilde{g} : \text{fl } A \rightarrow E'$ such that $\tilde{g}(\text{fl } A) = \text{fl } g(A)$. Accordingly, every affine function f defined on $F = \text{fl } A$ is uniquely determined by $f|_A$ and the image $f(F)$ is a flat.*

Proof. Let $F' = \text{fl } g(A)$ and take $v_0 \in A$. Since $(A \setminus \{v_0\}) - v_0$ is a Hamel basis of the linear subspace $\text{fl } A - v_0$ of E , we have the unique linear function $h : \text{fl } A - v_0 \rightarrow F' - g(v_0)$ such that

$$h(v - v_0) = g(v) - g(v_0) \text{ for each } v \in A \setminus \{v_0\}.$$

Then, g uniquely extends to the affine function $\tilde{g} : \text{fl } A \rightarrow F'$ defined by

$$\tilde{g}(x) = h(x - v_0) + g(v_0) \text{ for each } x \in \text{fl } A.$$

It is easy to see that $\tilde{g}(\text{fl } A) = \text{fl } g(A)$. □

Additional Properties of Flats and Affine Functions 3.1.6.

In the following, let E and E' be linear spaces and $f : F \rightarrow E'$ be a function of a flat F in E .

- (1) If f is affine and F' is a flat in E' , then $f(F)$ and $f^{-1}(F')$ are flats in E' and E , respectively.
- (2) A function f is affine if and only if the graph $\text{Gr}(f) = \{(x, f(x)) \mid x \in F\}$ of f is a flat in $E \times E'$.

3.2 Convex Sets

In this section, we introduce the basic concepts of convex sets. A subset $C \subset E$ is said to be **convex** if the line segment with the end points in C is contained in C , i.e.,

$$(1-t)x + ty \in C \text{ for each } x, y \in C \text{ and } t \in \mathbf{I}.$$

By induction on n , it can be proved that every convex set $C \subset E$ satisfies the following condition:

$$\sum_{i=1}^n z(i)v_i \in C \text{ for each } n \in \mathbb{N}, v_i \in C \text{ and } z \in \Delta^{n-1},$$

where $\Delta^{n-1} = \{z \in \mathbf{I}^n \mid \sum_{i=1}^n z(i) = 1\}$ is the standard $(n-1)$ -simplex. The following is easy:

- If $A, B \subset E$ are convex, then $aA + bB$ is also convex for each $a, b \in \mathbb{R}$.

The **dimension** of a convex set $C \subset E$ is defined by the dimension of the flat hull $\text{fl } C$, i.e., $\dim C = \dim \text{fl } C$. Concerning the flat hull of a convex set, we have the following proposition:

Proposition 3.2.1. *For each convex set $C \subset E$,*

$$\text{fl } C = \{(1-t)x + ty \mid x, y \in C, t \in \mathbb{R}\}.$$

Proof. Each $z \in \text{fl } C$ can be written $z = \sum_{i=1}^n t_i x_i$, where $x_i \in C$ and $\sum_{i=1}^n t_i = 1$. We may assume that $t_1 \leq \dots \leq t_n \in \mathbb{R} \setminus \{0\}$. If $t_1 \geq 0$ then $z \in C$. Otherwise, $t_k < 0$ and $t_{k+1} > 0$ for some $k = 1, \dots, n-1$. Then, we have $t = \sum_{i=1}^{n-k} t_{k+i} > 0$, where $1-t = \sum_{i=1}^k t_i < 0$. Let

$$x = \sum_{i=1}^k (1-t)^{-1} t_i x_i, \quad y = \sum_{i=1}^{n-k} t^{-1} t_{k+i} x_{k+i} \in C.$$

Then, $z = (1-t)x + ty$. Accordingly, we have

$$\text{fl } C \subset \{(1-t)x + ty \mid x, y \in C, t \in \mathbb{R}\}.$$

The converse inclusion is obvious. □

The smallest convex set containing $A \subset E$ is called the **convex hull** of A and is denoted by $\langle A \rangle$. We simply write $\langle v_1, \dots, v_n \rangle = \langle \{v_1, \dots, v_n\} \rangle$. Then, $\Delta^{n-1} = \langle \mathbf{e}_1, \dots, \mathbf{e}_n \rangle$. Observe that

$$\begin{aligned} \langle v_1, \dots, v_n \rangle &= \left\{ \sum_{i=1}^n z(i)v_i \mid z \in \Delta^{n-1} \right\} \text{ and} \\ \langle A \rangle &= \bigcup \{ \langle x_1, \dots, x_n \rangle \mid n \in \mathbb{N}, x_1, \dots, x_n \in A \}. \end{aligned}$$

For each two non-empty subsets $A, B \subset E$,

$$\begin{aligned} \langle A \cup B \rangle &= \{(1-t)x + ty \mid x \in \langle A \rangle, y \in \langle B \rangle, t \in \mathbf{I}\} \text{ and} \\ \langle aA + bB \rangle &= a\langle A \rangle + b\langle B \rangle \text{ for } a, b \in \mathbb{R}. \end{aligned}$$

The second equality can be proved as follows: Because $a\langle A \rangle + b\langle B \rangle$ is convex and $aA + bB \subset a\langle A \rangle + b\langle B \rangle$, we have $\langle aA + bB \rangle \subset a\langle A \rangle + b\langle B \rangle$. To show that $a\langle A \rangle + b\langle B \rangle \subset \langle aA + bB \rangle$, let $x \in \langle A \rangle$ and $y \in \langle B \rangle$. Then, $x = \sum_{i=1}^n t_i x_i$ and $y = \sum_{j=1}^m s_j y_j$ for some $x_i \in A, y_j \in B$, and $t_i, s_j > 0$ with $\sum_{i=1}^n t_i = \sum_{j=1}^m s_j = 1$. Since $ax_i + by_j \in aA + bB$ and $\sum_{i=1}^n \sum_{j=1}^m t_i s_j = 1$, it follows that

$$\begin{aligned} ax + by &= \sum_{i=1}^n t_i(ax_i + by) = \sum_{i=1}^n t_i \left(\sum_{j=1}^m s_j(ax_i + by_j) \right) \\ &= \sum_{i=1}^n \sum_{j=1}^m t_i s_j(ax_i + by_j) \in \langle aA + bB \rangle. \end{aligned}$$

Let C and C' be non-empty convex sets in the linear spaces E and E' , respectively. A function $f : C \rightarrow C'$ is said to be **affine** (or **linear in the affine sense**) provided

$$f((1-t)x + ty) = (1-t)f(x) + tf(y) \text{ for each } x, y \in C \text{ and } t \in \mathbf{I}.$$

As in the definition of a flat, \mathbf{I} can be replaced by \mathbb{R} , i.e.,

$$\begin{aligned} x, y \in C, t \in \mathbb{R}, (1-t)x + ty \in C \\ \Rightarrow f((1-t)x + ty) = (1-t)f(x) + tf(y). \end{aligned}$$

Indeed, let $z = (1-t)x + ty \in C$ in the above expression. When $t < 0$, consider

$$x = \frac{1}{1-t}z + \frac{-t}{1-t}y, \quad \frac{1}{1-t} \in \mathbf{I}, \quad \frac{-t}{1-t} = 1 - \frac{1}{1-t}.$$

When $t > 1$, consider

$$y = \frac{1}{t}z + \frac{t-1}{t}x, \quad \frac{1}{t} \in \mathbf{I}, \quad \frac{t-1}{t} = 1 - \frac{1}{t}.$$

As is easily seen, $f : C \rightarrow C'$ is affine if and only if

$$f\left(\sum_{i=1}^n z(i)v_i\right) = \sum_{i=1}^n z(i)f(v_i) \text{ for each } n \in \mathbb{N}, v_i \in C \text{ and } z \in \Delta^{n-1},$$

which is equivalent to the following:

$$v_i \in C, t_i \in \mathbb{R}, \sum_{i=1}^n t_i v_i \in C, \sum_{i=1}^n t_i = 1 \Rightarrow f\left(\sum_{i=1}^n t_i v_i\right) = \sum_{i=1}^n t_i f(v_i).$$

For every affine function $f : C \rightarrow E'$ of a convex set $C \subset E$ into another linear space E' , the image $f(C)$ is also convex.

Proposition 3.2.2. *Let C and D be non-empty convex sets in the linear spaces E and E' , respectively. Every affine function $f : C \rightarrow D$ uniquely extends to an affine function $\tilde{f} : \text{fl } C \rightarrow \text{fl } D$. Moreover, if f is injective (or surjective) then so is \tilde{f} .*

Proof. Let C_0 be a maximal affinely independent subset of C . Then, $\text{fl } C = \text{fl } C_0$. Due to Proposition 3.1.5, $f|_{C_0}$ uniquely extends to an affine function $\tilde{f} : \text{fl } C \rightarrow \text{fl } D$. From the above remark, we can see that $\tilde{f}|_C = f$.

If f is injective, we show that \tilde{f} is also injective. By the definition of \tilde{f} in the proof of Proposition 3.1.5, it suffices to show that $f(C_0)$ is affinely independent. Assume that $f(C_0)$ is not affinely independent, i.e., there are distinct points $v_1, \dots, v_n \in C_0$ and $t_1, \dots, t_n \in \mathbb{R} \setminus \{0\}$ such that $\sum_{i=1}^n t_i f(v_i) = \mathbf{0}$ and $\sum_{i=1}^n t_i = 0$. Without loss of generality, it can be assumed that $t_1, \dots, t_k > 0$ and $t_{k+1}, \dots, t_n < 0$. Note that $1 < k < n$ and $\sum_{i=1}^k t_i = -\sum_{j=k+1}^n t_j > 0$. Let

$$x = \sum_{i=1}^k \frac{t_i}{s} v_i \quad \text{and} \quad y = \sum_{j=k+1}^n -\frac{t_j}{s} v_j, \quad \text{where } s = \sum_{i=1}^k t_i > 0.$$

Then, $x, y \in C$ and $f(x) = f(y)$ because

$$f(x) - f(y) = \frac{1}{s} \sum_{i=1}^n t_i f(v_i) = \mathbf{0}.$$

Since f is injective, we have $x = y$. Hence, it follows that $\sum_{i=1}^k t_i v_i = -\sum_{j=k+1}^n t_j v_j$, i.e., $\sum_{i=1}^n t_i v_i = \mathbf{0}$. Because C_0 is affinely independent, $t_1 = \dots = t_n = 0$, which is a contradiction.

Finally, we show that if f is surjective then so is \tilde{f} . By Proposition 3.2.1, each $z \in \text{fl } D$ can be written as follows:

$$z = (1-t)y + ty', \quad y, y' \in D, \quad t \in \mathbb{R}.$$

Since f is surjective, we have $x, x' \in C$ such that $f(x) = y$ and $f(x') = y'$. Then, $(1-t)x + tx' \in \text{fl } C$ and

$$\tilde{f}((1-t)x + tx') = (1-t)y + ty' = z.$$

Therefore, \tilde{f} is also surjective. □

Let C be a convex set in a linear space E . The following set is called the **radial interior** of C :

$$\text{rint } C = \{x \in C \mid \forall y \in C, \exists \delta > 0 \text{ such that } (1+\delta)x - \delta y \in C\}^4$$

⁴In Köthe's book, $\text{rint } C$ is denoted by C^i and called the **algebraic kernel** of C .

In the case $C = \langle v_1, \dots, v_n \rangle$, observe that

$$\text{rint}\langle v_1, \dots, v_n \rangle = \left\{ \sum_{i=1}^n z(i)v_i \mid z \in \Delta^{n-1} \cap (0, \infty)^n \right\}.$$

Indeed, let $x_0 = \sum_{i=1}^n n^{-1}v_i \in \langle v_1, \dots, v_n \rangle$. For each $x \in \text{rint}\langle v_1, \dots, v_n \rangle$, we have $y \in \langle v_1, \dots, v_n \rangle$ such that $x \in \langle x_0, y \rangle$, i.e., $x = (1-t)x_0 + ty$ for some $t \in (0, 1)$. Then, $y = \sum_{i=1}^n z(i)v_i$ for some $z \in \Delta^{n-1}$. It follows that $x = \sum_{i=1}^n ((1-t)n^{-1} + tz(i))v_i$, where $\sum_{i=1}^n ((1-t)n^{-1} + tz(i)) = 1$ and $(1-t)n^{-1} + tz(i) > 0$ for all $i = 1, \dots, n$. Thus, x is a point of the rightside set. Conversely, it is straightforward to prove that each point of the rightside set belongs to $\langle v_1, \dots, v_n \rangle$.

In particular, $\text{rint}\langle v_1, v_2 \rangle = \{(1-t)v_1 + tv_2 \mid 0 < t < 1\}$, and hence $\text{rint}\langle v_1, v_2 \rangle = \langle v_1, v_2 \rangle \setminus \{v_1, v_2\}$ if $v_1 \neq v_2$. The radial interior of C can also be defined as

$$\text{rint } C = \{x \in C \mid \forall y \in C, \exists z \in C \text{ such that } x \in \text{rint}\langle y, z \rangle\}.$$

For each $x \in C$, the following subset $C_x \subset C$ is called the **face** of C at x :

$$\begin{aligned} C_x &= \{y \in C \mid \exists \delta > 0 \text{ such that } (1 + \delta)x - \delta y \in C\} \\ &= \{y \in C \mid \exists z \in C \text{ such that } x \in \text{rint}\langle y, z \rangle\}.^5 \end{aligned}$$

By an easy observation, we have

$$\text{rint } C = \{x \in C \mid C_x = C\}, \quad \text{i.e., } x \in \text{rint } C \Leftrightarrow C_x = C.$$

When $C_x = \{x\}$, we call x an **extreme point** of C . It is said that $x \in E$ is **linearly accessible** from C if there is some $y \in C$ such that

$$\text{rint}\langle x, y \rangle \subset C \quad (\text{i.e., } \langle x, y \rangle \setminus \{x\} \subset C).$$

The **radial closure** $\text{rcl } C$ of C is the set of all linearly accessible points from C .⁶ It should be noted that $\text{rcl } C \subset \text{fl } C$ by Proposition 3.2.1, hence $\text{fl } \text{rcl } C = \text{fl } C$. Consequently, we have the following inclusions:

$$\text{rint } C \subset C \subset \text{rcl } C \subset \text{fl } C.$$

The set $\partial C = \text{rcl } C \setminus \text{rint } C$ is called the **radial boundary** of C .

Remark 1. Note that $A \subset B$ implies $\text{rcl } A \subset \text{rcl } B$, but it does not imply $\text{rint } A \subset \text{rint } B$. For example, consider $A = \mathbf{I}^n \times \{\mathbf{0}\} \subset B = \mathbf{I}^{n+1}$. Then, $A \cap \text{rint } B = \emptyset$.

⁵The face C_x is a little differently defined than the **supporting facet** of C through x in Köthe's book.

⁶In Köthe's book, $\text{rcl } C$ is denoted by C^a and called the **algebraic hull** of C .

For the Hilbert cube $\mathbf{Q} = [-1, 1]^{\mathbb{N}}$, we have

$$\text{rint } \mathbf{Q} = \{x \in \mathbf{Q} \mid \sup_{i \in \mathbb{N}} |x(i)| < 1\} \subsetneq (-1, 1)^{\mathbb{N}}.$$

Observe that $\text{rint}[-1, 1]_f^{\mathbb{N}} = (-1, 1)_f^{\mathbb{N}}$ but $\text{rint } \mathbf{I}_f^{\mathbb{N}} = \emptyset$, where

$$[-1, 1]_f^{\mathbb{N}} = \mathbb{R}_f^{\mathbb{N}} \cap [-1, 1]^{\mathbb{N}}, \quad (-1, 1)_f^{\mathbb{N}} = \mathbb{R}_f^{\mathbb{N}} \cap (-1, 1)^{\mathbb{N}}, \quad \text{and } \mathbf{I}_f^{\mathbb{N}} = \mathbb{R}_f^{\mathbb{N}} \cap \mathbf{I}^{\mathbb{N}}.^7$$

As is easily observed, $\mathbf{I}_f^{\mathbb{N}} = \text{rcl}(\mathbf{I}_f^{\mathbb{N}} \setminus \{\mathbf{0}\})$. It will be shown in Remark 3 that $\mathbf{I}_f^{\mathbb{N}} \setminus \{\mathbf{0}\} = \text{rcl } C$ for some convex set $C \subset \mathbb{R}_f^{\mathbb{N}}$.

Remark 2. The unit closed ball \mathbf{B}_{c_0} of the Banach space c_0 has no extreme points. In fact, every $x \in \mathbf{B}_{c_0}$ is the midpoint of two distinct points $y, z \in \mathbf{B}_{c_0}$, i.e., $x = \frac{1}{2}y + \frac{1}{2}z$. For example, choose $n \in \mathbb{N}$ so that $|x(n)| < \frac{1}{2}$ and let $y, z \in \mathbf{B}_{c_0}$ such that $y(i) = z(i) = x(i)$ for $i \neq n$, $y(n) = x(n) + \frac{1}{2}$, and $z(n) = x(n) - \frac{1}{2}$.

Proposition 3.2.3. *Let $C \subset E$ be a convex set. If $x \in \text{rint } C$, $y \in \text{rcl } C$, and $0 \leq t < 1$, then $(1-t)x + ty \in \text{rint } C$, i.e., $\langle x, y \rangle \setminus \{y\} \subset \text{rint } C$.*

Proof. For each $z \in C$, we have to find $v \in C$ and $0 < s < 1$ such that

$$(1-t)x + ty = (1-s)z + sv \in \text{rint}\langle z, v \rangle.$$

Take $w \in C$ so that $\text{rint}\langle w, y \rangle \subset C$, and choose $0 < r < 1$ so that

$$z' = (1+r)x - rz, \quad w' = (1+r)x - rw \in C.$$

The desired v is to be written as

$$v = t_1y + t_2w + t_3w' + t_4z' = (t_1 + t_2)u + (t_3 + t_4)u' \in C,$$

where $t_1 + t_2 + t_3 + t_4 = 1$, $t_1, t_2, t_3, t_4 > 0$,

$$u = \frac{t_1}{t_1 + t_2}y + \frac{t_2}{t_1 + t_2}w, \quad u' = \frac{t_3}{t_3 + t_4}w' + \frac{t_4}{t_3 + t_4}z' \in C.$$

Then, we have

$$\begin{aligned} (1-s)z + sv &= (1-s)z + s(t_1y + t_2w + t_3w' + t_4z') \\ &= st_1y + s(t_2 - t_3r)w + s(t_3 + t_4)(1+r)x + (1-s - st_4r)z. \end{aligned}$$

To obtain $(1-s)z + sv = (1-t)x + ty$, it is enough to find $t_1, t_2, t_3, t_4 > 0$ and $0 < s < 1$ satisfying the simultaneous equations: $st_1 = t$, $t_2 = t_3r$, $s(t_3 + t_4)(1+r) =$

⁷It is known that $[-1, 1]_f^{\mathbb{N}} \approx \mathbf{I}_f^{\mathbb{N}}$.

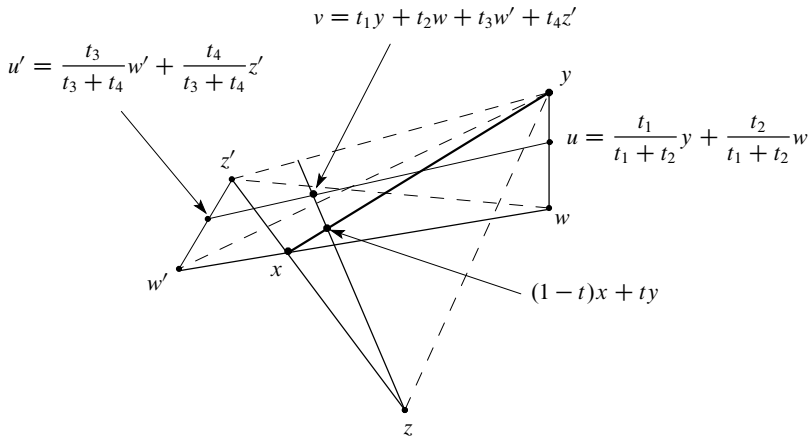


Fig. 3.1 $(1-t)x + ty \in \text{rint } C$

$1-t$, and $1-s = st_4r$, i.e.,

$$(*) \quad t_1 = \frac{t}{s}, \quad t_4 = \frac{1-s}{rs}, \quad t_3 = \frac{1}{r} - \frac{1+rt}{(1+r)rs}, \quad t_2 = 1 - \frac{1+rt}{(1+r)s}.$$

Since $t_1, t_4 < 1$ and $0 < t_2 (< t_3)$, it is necessary to satisfy

$$\max \left\{ t, \frac{1}{1+r}, \frac{1+rt}{1+r} \right\} < s < 1.$$

We can take such an s because the left side of the above inequality is less than 1. Then, we can define $t_1, t_2, t_3, t_4 > 0$ as in (*), which satisfies $t_1 + t_2 + t_3 + t_4 = 1$. Thus, we have the desired $v = t_1 y + t_2 w + t_3 w' + t_4 z' \in C$ — Fig. 3.1. □

Although we verified in Remark 1 that $A \subset B$ does not imply $\text{rint } A \subset \text{rint } B$ in general, we do have the following corollary:

Corollary 3.2.4. *Let A and B be non-empty convex sets in E . If $A \subset B$ and $A \cap \text{rint } B \neq \emptyset$, then $\text{rint } A \subset \text{rint } B$.*

Proof. Let $x \in A \cap \text{rint } B$. For each $y \in \text{rint } A$, we have $z \in A$ such that $y \in \text{rint}\langle x, z \rangle$. Since $\text{rint}\langle x, z \rangle \subset \text{rint } B$ by Proposition 3.2.3, it follows that $y \in \text{rint } B$. □

Proposition 3.2.5. *For each convex set $C \subset E$, the following statements hold:*

- (1) Both $\text{rint } C$ and $\text{rcl } C$ are convex;
- (2) $\text{rint } \text{rint } C = \text{rint } C \subset \text{rint } \text{rcl } C$;
- (3) $\text{rint } C \neq \emptyset \Rightarrow \text{rint } \text{rcl } C = \text{rint } C$, $\text{rcl } \text{rint } C = \text{rcl } \text{rcl } C = \text{rcl } C$,
in which case $\partial \text{rint } C = \partial \text{rcl } C = \partial C$;

- (4) $\text{rint } C \neq \emptyset \Rightarrow \text{fl } C = \text{fl } \text{rint } C$;
 (5) $\text{rint } C \neq \emptyset, \text{rcl } C = \text{fl } C \Rightarrow \text{rint } C = C = \text{fl } C$;
 (6) $\partial C \neq \emptyset \Leftrightarrow \emptyset \neq C \subsetneq \text{fl } C$;
 (7) C_x is convex and $C_x = C \cap \text{fl } C_x$ for $x \in C$;
 (8) $x \in \text{rint } C_x$ for $x \in C$, hence $(C_x)_x = C_x$;
 (9) $(C_x)_y = C_y$ for $x \in C$ and $y \in C_x$;
 (10) $C_x = C_y$ for $x \in C$ and $y \in \text{rint } C_x$.

Proof. (1): To prove the convexity of $\text{rint } C$, we can apply Proposition 3.2.3. It is now quite straightforward to show the convexity of $\text{rcl } C$.

(2): To show $\text{rint } C \subset \text{rint } \text{rint } C$, we can apply Proposition 3.2.3. Because $\text{rint}(\text{rint } C) \subset \text{rint } C$ by Corollary 3.2.4, we have $\text{rint } \text{rint } C = \text{rint } C$.

For each $x \in \text{rint } C$ and $y \in \text{rcl } C$, $\frac{1}{2}x + \frac{1}{2}y \in \text{rint } C$ by Proposition 3.2.3. Then, we have $\delta > 0$ such that $(1 + \delta)x - \delta(\frac{1}{2}x + \frac{1}{2}y) \in C$, i.e., $(1 + \frac{1}{2}\delta)x - \frac{1}{2}\delta y \in C$. Hence, $x \in \text{rint } \text{rcl } C$.

(3): Let $x_0 \in \text{rint } C$. For each $x \in \text{rint } \text{rcl } C$, we have $y \in \text{rcl } C$ such that $x \in \text{rint}\langle x_0, y \rangle$, which implies that $x \in \text{rint } C$ by Proposition 3.2.3. Combining this with (2) yields $\text{rint } \text{rcl } C = \text{rint } C$.

We now have $x_0 \in \text{rint } C = \text{rint } \text{rcl } C$. If $x \in \text{rcl } \text{rcl } C$, then $\text{rint}\langle x_0, x \rangle \subset \text{rint } \text{rcl } C = \text{rint } C$ by Proposition 3.2.3, which means that $x \in \text{rcl } \text{rint } C$. Since $\text{rcl } \text{rint } C \subset \text{rcl } C \subset \text{rcl } \text{rcl } C$, we have $\text{rcl } \text{rint } C = \text{rcl } C = \text{rcl } \text{rcl } C$.

(4): Let $x_0 \in \text{rint } C$. For each $x \in C$, $\frac{1}{2}x + \frac{1}{2}x_0 \in \text{fl } \text{rint } C$ by Proposition 3.2.3. Then, it follows from Proposition 3.2.1 that $x = 2(\frac{1}{2}x + \frac{1}{2}x_0) - x_0 \in \text{fl } \text{rint } C$. Accordingly, we have $C \subset \text{fl } \text{rint } C$, which implies $\text{fl } C \subset \text{fl } \text{rint } C$. Since $\text{fl } \text{rint } C \subset \text{fl } C$, we have $\text{fl } C = \text{fl } \text{rint } C$.

(5): Let $x_0 \in \text{rint } C$. For each $x \in \text{fl } C$, $2x - x_0 \in \text{fl } C = \text{rcl } C$. Then, $x = \frac{1}{2}x_0 + \frac{1}{2}(2x - x_0) \in \text{rint } C \subset C$ by Proposition 3.2.3.

(6): Assume $\emptyset \neq C \subsetneq \text{fl } C$. Then, we have $x \in \text{fl } C \setminus C$, which can be written as $x = (1 + t)y - tz$ for some $y \neq z \in C$ and $t > 0$ by Proposition 3.2.1. Let

$$s = \inf \{t > 0 \mid (1 + t)y - tz \notin C\} \geq 0.$$

Then, $(1 + s)y - sz \in \text{rcl } C \setminus \text{rint } C = \partial C$.

When $C = \text{fl } C$, i.e., C is a flat, we have $\text{rcl } C = \text{rint } C = C$ by definition, which means $\partial C = \emptyset$. Therefore, $\partial C \neq \emptyset$ implies $\emptyset \neq C \subsetneq \text{fl } C$.

(7): First, we show that C_x is convex. For each $y, z \in C_x$, we can choose $\delta > 0$ so that $(1 + \delta)x - \delta y \in C$ and $(1 + \delta)x - \delta z \in C$. Then, for each $t \in \mathbf{I}$,

$$\begin{aligned} (1 + \delta)x - \delta((1 - t)y + tz) \\ = (1 - t)((1 + \delta)x - \delta y) + t((1 + \delta)x - \delta z) \in C, \end{aligned}$$

which means $(1 - t)y + tz \in C_x$.

Because $C_x \subset C \cap \text{fl } C_x$, it remains to show $C \cap \text{fl } C_x \subset C_x$. By Proposition 3.2.1, each $y \in C \cap \text{fl } C_x$ can be written as $y = (1 - t)y' + ty''$ for some $y', y'' \in C_x$

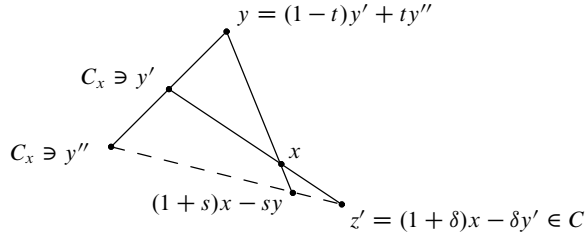


Fig. 3.2 $C \cap \text{ri } C_x \subset C_x$

and $t \in \mathbb{R}$. Because of the convexity of C_x , we have $y \in C_x$ if $t \in \mathbf{I}$. Then, we may assume that $t < 0$ (if $t > 1$, exchange y' with y''). We have $\delta > 0$ such that $z' = (1 + \delta)x - \delta y' \in C$. Observe that

$$\begin{aligned} (1 + s)x - sy &= (1 + s)\left(\frac{\delta}{1 + \delta}y' + \frac{1}{1 + \delta}z'\right) - s((1 - t)y' + ty'') \\ &= \left(\frac{(1 + s)\delta}{1 + \delta} - s(1 - t)\right)y' + \frac{1 + s}{1 + \delta}z' - sty''. \end{aligned}$$

Let $s = \delta/(1 - t - t\delta) > 0$. Then, since $1 + s = (1 - t)(1 + \delta)/(1 - t - t\delta)$, it follows that

$$(1 + s)x - sy = \frac{1 - t}{1 - t - t\delta}z' + \frac{-t\delta}{1 - t - t\delta}y'' \in C,$$

which implies that $y \in C_x$ (Fig. 3.2).

(8): From the definition of $\text{rint } C_x$, it easily follows that $x \in \text{rint } C_x$.

(9): Because $C_x \subset C$, we have $(C_x)_y \subset C_y$ by definition. We will show that $C_y \subset C_x$, which implies $C_y = (C_y)_y \subset (C_x)_y$ by (8) and the definition. For each $z \in C_y$, choose $\delta_1 > 0$ so that $u = (1 + \delta_1)y - \delta_1 z \in C$. On the other hand, since $y \in C_x$, we have $\delta_2 > 0$ such that $v = (1 + \delta_2)x - \delta_2 y \in C$. Then,

$$\frac{(1 + \delta_1)(1 + \delta_2)}{1 + \delta_1 + \delta_2}x - \frac{\delta_1\delta_2}{1 + \delta_1 + \delta_2}z = \frac{1 + \delta_1}{1 + \delta_1 + \delta_2}v + \frac{\delta_2}{1 + \delta_1 + \delta_2}u \in C,$$

which means that $z \in C_x$.

(10): Since $y \in \text{rint } C_x$, we have $(C_x)_y = C_x$. On the other hand, $(C_x)_y = C_y$ by (9). □

Remark 3. It should be noted that, in general, $\text{rcl rcl } C \neq \text{rcl } C$. For example, let C be the convex set in $\mathbb{R}_f^{\mathbb{N}}$ defined as follows:

$$\begin{aligned} C = \{x \in \mathbf{I}_f^{\mathbb{N}} \mid \exists k \in \mathbb{N} \text{ such that } \sum_{i \in \mathbb{N}} x(i) \geq k^{-1}, \\ x(i) \neq 0 \text{ at least } k \text{ many } i \in \mathbb{N}\}. \end{aligned}$$

It is easy to see that $\mathbf{0} \notin \text{rcl } C$, i.e., $\text{rcl } C \subset \mathbf{I}_f^{\mathbb{N}} \setminus \{\mathbf{0}\}$. For each $x \in \mathbf{I}_f^{\mathbb{N}} \setminus \{\mathbf{0}\}$, choose $k \in \mathbb{N}$ so that $k^{-1} \leq \sum_{i \in \mathbb{N}} x(i)$, and let $y \in C$ such that $y(i) = k^{-2}$ for $i \leq k$ and $y(i) = 0$ for $i > k$. If $0 < t \leq 1$, then $(1-t)x + ty \in C$ because $(1-t)x(i) + ty(i) \neq 0$ for at least k many $i \in \mathbb{N}$ and

$$\sum_{i \in \mathbb{N}} ((1-t)x(i) + ty(i)) = (1-t) \sum_{i \in \mathbb{N}} x(i) + t \sum_{i \in \mathbb{N}} y(i) \geq k^{-1}.$$

Therefore, $\text{rcl } C = \mathbf{I}_f^{\mathbb{N}} \setminus \{\mathbf{0}\}$. As observed in Remark 1, $\text{rcl}(\mathbf{I}_f^{\mathbb{N}} \setminus \{\mathbf{0}\}) = \mathbf{I}_f^{\mathbb{N}}$. Hence, we have $\text{rcl } \text{rcl } C \neq \text{rcl } C$. It should also be noted that $\text{rint } C = \emptyset$.

In the finite-dimensional case, we have the following proposition:

Proposition 3.2.6. *Every non-empty finite-dimensional convex set C has a non-empty radial interior; i.e., $\text{rint } C \neq \emptyset$, and therefore*

$$\text{rcl } \text{rint } C = \text{rcl } \text{rcl } C = \text{rcl } C \quad \text{and} \quad \partial \text{rint } C = \partial \text{rcl } C = \partial C.$$

Proof. We have a maximal affinely independent finite subset $\{v_1, \dots, v_n\} \subset C$. Then, $v_0 = \sum_{i=1}^n n^{-1} v_i \in \text{rint } C$. Indeed, since $C \subset \text{fl}\{v_1, \dots, v_n\}$, each $x \in C$ can be written as $x = \sum_{i=1}^n t_i v_i$, where $\sum_{i=1}^n t_i = 1$. Observe that

$$\begin{aligned} (1 + \delta)v_0 - \delta x &= (1 + \delta) \sum_{i=1}^n n^{-1} v_i - \delta \sum_{i=1}^n t_i v_i \\ &= \sum_{i=1}^n (n^{-1} + \delta(n^{-1} - t_i)) v_i. \end{aligned}$$

When $v_0 \neq x$, we have $s = \min\{n^{-1} - t_i \mid i=1, \dots, n\} < 0$. Let $\delta = 1/(-sn) > 0$. Then, $n^{-1} + \delta(n^{-1} - t_i) \geq 0$ for every $i = 1, \dots, n$, which implies that $(1 + \delta)v_0 - \delta x \in C$. \square

Additional Results for Convex Sets 3.2.7.

(1) For every two convex sets C and D ,

$$(C \cap D)_x = C_x \cap D_x \quad \text{for each } x \in C \cap D.$$

(2) For every two convex sets C and D with $\text{rint } C \cap \text{rint } D \neq \emptyset$,

$$\text{rint}(C \cap D) = \text{rint } C \cap \text{rint } D.$$

In general, $\text{rint } C \cap \text{rint } D \subset \text{rint}(C \cap D)$.

Sketch of Proof. To show that $\text{rint}(C \cap D) \subset \text{rint } C \cap \text{rint } D$, let $x_0 \in \text{rint } C \cap \text{rint } D$. For each $x \in \text{rint}(C \cap D)$, take $y \in C \cap D$ so that $x \in \text{rint}\langle x_0, y \rangle$. Since $\text{rint}\langle x_0, y \rangle \subset \text{rint } C$ by Proposition 3.2.3, it follows that $x \in \text{rint } C$. Hence, $\text{rint}(C \cap D) \subset \text{rint } C$. Similarly, we have $\text{rint}(C \cap D) \subset \text{rint } D$.

- (3) Let C and D be convex sets in the linear spaces E and E' , respectively. Then, $C \times D$ is also convex,

$$\text{rint}(C \times D) = \text{rint } C \times \text{rint } D \quad \text{and} \quad \text{rcl}(C \times D) = \text{rcl } C \times \text{rcl } D.$$

Moreover, $(C \times D)_{(x,y)} = C_x \times D_y$ for each $(x, y) \in C \times D$.

- (4) Let $f : C \rightarrow E'$ be an affine function of a convex set C in a linear space E into another linear space E' , and D be a convex set in E' . Then, $f(C)$ and $f^{-1}(D)$ are convex and

$$f^{-1}(D)_x = C_x \cap f^{-1}(D_{f(x)}) \quad \text{for each } x \in f^{-1}(D) (\subset C).$$

In particular, $C_x \subset f^{-1}(f(C)_{f(x)})$ (i.e., $f(C_x) \subset f(C)_{f(x)}$) for each $x \in C$. When f is injective, $f(C_x) = f(C)_{f(x)}$ for each $x \in C$.

Sketch of Proof. It is easy to see that $f(f^{-1}(D)_x) \subset D_{f(x)}$, hence $f^{-1}(D)_x \subset f^{-1}(D_{f(x)})$. Also, $f^{-1}(D)_x \subset C_x$ because $f^{-1}(D) \subset C$. Accordingly, $f^{-1}(D)_x \subset C_x \cap f^{-1}(D_{f(x)})$. To prove the converse inclusion, for each $y \in f^{-1}(D_{f(x)}) \cap C_x$, choose $\delta > 0$ so that $(1 + \delta)f(x) - \delta f(y) \in D$ and $(1 + \delta)x - \delta y \in C$. Then, $(1 + \delta)x - \delta y \in f^{-1}(D)$.

- (5) For every (bounded) subset A of a normed linear space $E = (E, \|\cdot\|)$, the following hold:

- (i) $\|x - y\| \leq \sup_{z \in A} \|x - z\|$ for each $x \in E$ and $y \in \langle A \rangle$;
- (ii) $\text{diam}\langle A \rangle = \text{diam } A$.

Sketch of Proof. (i): Write $y = \sum_{i=1}^n z(i)x_i$ for some $x_1, \dots, x_n \in A$ and $z \in \Delta^{n-1}$.

(ii): For each $x, y \in \langle A \rangle$,

$$\|x - y\| \leq \sup_{z \in A} \|x - z\| \leq \sup_{z \in A} \sup_{z' \in A} \|z - z'\| = \text{diam } A.$$

Remark 4. In (2) above, $\text{rint}(C \cap D) \neq \text{rint } C \cap \text{rint } D$ in general. Consider the case that $C \cap D \neq \emptyset$ but $\text{rint } C \cap \text{rint } D = \emptyset$.

In (4) above, $f(C_x) \neq f(C)_{f(x)}$ in general. For instance, let $C = \{(s, t) \in \mathbb{R}^2 \mid |s| \leq t \leq 1\} \subset \mathbb{R}^2$. Then, $\text{pr}_1(C) = [-1, 1]$, $\text{pr}_1(C_0) = \{0\}$, and $\text{pr}_1(C)_0 = \text{pr}_1(C)$.

3.3 The Hahn–Banach Extension Theorem

We now prove the Hahn–Banach Extension Theorem and present a relationship between the sublinear functionals and the convex sets.

Let E be a linear space. A functional $p : E \rightarrow \mathbb{R}$ is **sublinear** if it satisfies the following conditions:

- (SL₁) $p(x + y) \leq p(x) + p(y)$ for each $x, y \in E$, and

(SL₂) $p(tx) = tp(x)$ for each $x \in E$ and $t > 0$.

Note that if $p : E \rightarrow \mathbb{R}$ is sublinear then $p(\mathbf{0}) = 0$ and $-p(-x) \leq p(x)$. For each $x, y \in E$ and $t \in \mathbf{I}$,

$$p((1-t)x + ty) \leq (1-t)p(x) + tp(y).$$

When $p : E \rightarrow \mathbb{R}$ is a non-negative sublinear functional, $p^{-1}([0, r]) = rp^{-1}([0, 1])$ and $p^{-1}([0, r]) = rp^{-1}(\mathbf{I})$ are convex for each $r > 0$.

In the following Hahn–Banach Extension Theorem, no topological concepts appear (even in the proof). Nevertheless, this theorem is very important in the study of topological linear spaces.

Theorem 3.3.1 (HAHN–BANACH EXTENSION THEOREM). *Let $p : E \rightarrow \mathbb{R}$ be a sublinear functional of a linear space E and F be a linear subspace of E . If $f : F \rightarrow \mathbb{R}$ is a linear functional such that $f(x) \leq p(x)$ for every $x \in F$, then f extends to a linear functional $\tilde{f} : E \rightarrow \mathbb{R}$ such that $\tilde{f}(x) \leq p(x)$ for every $x \in E$.*

Proof. Let \mathcal{F} be the collection of all linear functionals $f' : F' \rightarrow \mathbb{R}$ of a linear subspace $F' \subset E$ such that $F \subset F'$, $f'|_F = f$, and $f'(x) \leq p(x)$ for every $x \in F'$. For $f', f'' \in \mathcal{F}$, we define $f' \leq f''$ if f'' is an extension of f' . Then, $\mathcal{F} = (\mathcal{F}, \leq)$ is an inductive ordered set, i.e., every totally ordered subset of \mathcal{F} is upper bounded. By Zorn's Lemma, \mathcal{F} has a maximal element $f_0 : F_0 \rightarrow \mathbb{R}$. It suffices to show that $F_0 = E$.

Assume that $F_0 \neq E$. Taking $x_1 \in E \setminus F_0$, we have a linear subspace $F_1 = F_0 + \mathbb{R}x_1 \supsetneq F_0$. We show that f_0 has a linear extension $f_1 : F_1 \rightarrow \mathbb{R}$ in \mathcal{F} , which contradicts the maximality of f_0 . By assigning x_1 to $\alpha \in \mathbb{R}$, f_1 can be defined, i.e., $f_1(x + tx_1) = f_0(x) + t\alpha$ for $x \in F_0$ and $t \in \mathbb{R}$. In order that $f_1 \in \mathcal{F}$, we have to choose α so that for every $x \in F_0$ and $t > 0$,

$$f_0(x) + t\alpha \leq p(x + tx_1) \quad \text{and} \quad f_0(x) - t\alpha \leq p(x - tx_1).$$

Dividing by t , we obtain the following equivalent condition:

$$f_0(y) - p(y - x_1) \leq \alpha \leq p(y + x_1) - f_0(y) \quad \text{for every } y \in F_0.$$

Hence, such an $\alpha \in \mathbb{R}$ exists if

$$\sup\{f_0(y) - p(y - x_1) \mid y \in F_0\} \leq \inf\{p(y + x_1) - f_0(y) \mid y \in F_0\}.$$

This inequality can be proved as follows: for each $y, y' \in F_0$,

$$f_0(y) + f_0(y') = f_0(y + y') \leq p(y + y') \leq p(y - x_1) + p(y' + x_1),$$

hence $f_0(y) - p(y - x_1) \leq p(y' + x_1) - f_0(y')$, which implies the desired inequality. \square

Let F be a flat in a linear space E and $A \subset F$. The following set is called the **core** of A in F :

$$\text{core}_F A = \{x \in A \mid \forall y \in F, \exists \delta > 0 \text{ such that} \\ |t| \leq \delta \Rightarrow (1-t)x + ty \in A\},$$

where $|t| \leq \delta$ can be replaced by $-\delta \leq t \leq 0$ (or $0 \leq t \leq \delta$). Each point of $\text{core}_F A$ is called a **core point** of A in F . In the case that A is convex,

$$x \in \text{core}_F A \Leftrightarrow \forall y \in F, \exists \delta > 0 \text{ such that } (1+\delta)x - \delta y \in A \\ \Leftrightarrow \forall y \in F, \exists \delta > 0 \text{ such that } (1-\delta)x + \delta y \in A.$$

When $F = E$, we can omit the phrase “in E ” and simply write $\text{core } A$ by removing the subscript E . By definition, $A \subset B \subset F$ implies $\text{core}_F A \subset \text{core}_F B$. We also have the following fact:

Fact. For each $A \subset F$, $\text{core}_F A \neq \emptyset$ if and only if $\text{fl } A = F$.

Indeed, the “if” part is trivial. To show the “only if” part, let $x \in \text{core}_F A$. For each $y \in F$, we have $\delta > 0$ such that $z = (1+\delta)x - \delta y \in A$. Then, $y = \delta^{-1}(1+\delta)x - \delta^{-1}z \in \text{fl } A$. Note that $\text{fl } A \subset F$ because $A \subset F$. Consequently, $\text{fl } A = F$.

Proposition 3.3.2. For every convex set $A \subset E$, $\text{core}_{\text{fl } A} A = \text{rint } A$, which is also convex. Hence, $\text{core } A \neq \emptyset$ implies $\text{core } A = \text{rint } A$ and $\text{core core } A = \text{core } A$.

Proof. Because $\text{core}_{\text{fl } A} A \subset \text{rint } A$ by definition, it suffices to show that $\text{rint } A \subset \text{core}_{\text{fl } A} A$. For each $x \in \text{rint } A$ and $y \in \text{fl } A$, we need to find some $s > 0$ such that $(1+s)x - sy \in A$. This can be done using the same proof of the inclusion $C \cap \text{fl } C_x \subset C_x$ in Proposition 3.2.5(7). \square

Remark 5. When A is a finite-dimensional convex set, $\text{core}_F A \neq \emptyset$ if and only if $F = \text{fl } A$ according to Propositions 3.3.2 and 3.2.6. However, this does not hold for an infinite-dimensional convex set. For example, consider the convex set $\mathbf{I}_f^{\mathbb{N}}$ in $\mathbb{R}^{\mathbb{N}}$. Then, $\mathbb{R}_f^{\mathbb{N}} = \text{fl } \mathbf{I}_f^{\mathbb{N}}$ and $\text{core}_{\mathbb{R}_f^{\mathbb{N}}} \mathbf{I}_f^{\mathbb{N}} = \text{rint } \mathbf{I}_f^{\mathbb{N}} = \emptyset$.

With regard to convex sets defined by a non-negative sublinear functional, we have the following proposition:

Proposition 3.3.3. Let $p : E \rightarrow \mathbb{R}$ be a non-negative sublinear functional of a linear subspace E . Then,

$$p^{-1}([0, 1)) = \text{core } p^{-1}([0, 1)) = \text{core } p^{-1}(\mathbf{I}).$$

Proof. The inclusion $\text{core } p^{-1}([0, 1)) \subset \text{core } p^{-1}(\mathbf{I})$ is obvious.

Let $x \in p^{-1}([0, 1))$. For each $y \in E$, we can choose $\delta > 0$ so that $\delta p(x - y) < 1 - p(x)$. Then,

$$0 \leq p((1 + \delta)x - \delta y) = p(x + \delta(x - y)) \leq p(x) + \delta p(x - y) < 1,$$

i.e., $x \in \text{core } p^{-1}([0, 1])$. Hence, $p^{-1}([0, 1]) \subset \text{core } p^{-1}([0, 1])$.

If $p(x) \geq 1$, then $x \notin \text{core } p^{-1}(\mathbf{I})$ because

$$p((1 + t)x - t\mathbf{0}) = (1 + t)p(x) > 1 \text{ for any } t > 0.$$

This means that $\text{core } p^{-1}(\mathbf{I}) \subset p^{-1}([0, 1])$. □

For each $A \subset E$ with $\mathbf{0} \in \text{core } A$, the **Minkowski functional** $p_A : E \rightarrow \mathbb{R}_+$ can be defined as follows:

$$p_A(x) = \inf \{s > 0 \mid x \in sA\} = \inf \{s > 0 \mid s^{-1}x \in A\}.$$

Then, for each $x \in E$ and $t > 0$,

$$\begin{aligned} p_A(tx) &= \inf \{s > 0 \mid s^{-1}tx \in A\} = \inf \{ts > 0 \mid (ts)^{-1}tx \in A\} \\ &= t \inf \{s > 0 \mid s^{-1}x \in A\} = tp_A(x), \end{aligned}$$

i.e., p_A satisfies (SL_2) . In the above, $p_A(tx) = p_{t^{-1}A}(x)$. Then, it follows that $p_{t^{-1}A} = tp_A$ for each $t > 0$. Replacing t by t^{-1} , we have

$$p_{tA} = t^{-1}p_A \text{ for each } t > 0.$$

If $A \subset E$ is convex, the Minkowski functional p_A has the following desirable properties:

Proposition 3.3.4. *Let $A \subset E$ be a convex set with $\mathbf{0} \in \text{core } A$. Then, the Minkowski functional p_A is sublinear and*

$$\text{rint } A = \text{core } A = p_A^{-1}([0, 1]) \subset A \subset p_A^{-1}(\mathbf{I}) = \text{rcl } A,$$

so $\partial A = p_A^{-1}(1)$. Moreover,

$$p_A(x) = 0 \Leftrightarrow \mathbb{R}_+x \subset A.$$

In order that p_A is a norm on E , it is necessary and sufficient that $\mathbb{R}_+x \not\subset A$ if $x \neq \mathbf{0}$ and $tA \subset A$ if $|t| < 1$.

Proof. First, we prove that p_A is sublinear. As already observed, p_A satisfies (SL_2) . To show that p_A satisfies (SL_1) , let $x, y \in E$. Since A is convex, we have

$$s^{-1}x, t^{-1}y \in A \Rightarrow (s + t)^{-1}(x + y) = \frac{s}{s + t}s^{-1}x + \frac{t}{s + t}t^{-1}y \in A,$$

which implies that $p_A(x + y) \leq p_A(x) + p_A(y)$.

The first equality $\text{rint } A = \text{core } A$ has been stated in Proposition 3.3.2. It easily follows from the definitions that $\text{core } A \subset p_A^{-1}([0, 1]) \subset A \subset p_A^{-1}(\mathbf{1})$ and $p_A^{-1}(1) \subset \text{rcl } A$, so $p_A^{-1}(\mathbf{1}) \subset \text{rcl } A$. By Propositions 3.3.2 and 3.3.3, we have

$$\text{core } A = \text{core core } A \subset \text{core } p_A^{-1}([0, 1]) = p_A^{-1}([0, 1]) \subset \text{core } A,$$

which means the second equality $\text{core } A = p_A^{-1}([0, 1])$. To obtain the third equality $p_A^{-1}(\mathbf{1}) = \text{rcl } A$, it remains to show that $\text{rcl } A \subset p_A^{-1}(\mathbf{1})$. Let $x \in \text{rcl } A$. Since $\mathbf{0} \in \text{rint } A$, it follows from Proposition 3.2.3 that $s^{-1}x \in \text{rint } C \subset C$ for each $s > 1$, which implies that $p_A(x) \leq 1$, i.e., $x \in p_A^{-1}(\mathbf{1})$.

By definition, $p_A(x) = 0$ if and only if $tx \in A$ for an arbitrarily large $t > 0$, which means that $\mathbb{R}_+x \subset A$ because A is convex.

Because p_A is sublinear, p_A is a norm if and only if $p_A(x) \neq 0$ and $p_A(x) = p_A(-x)$ for every $x \in E \setminus \{\mathbf{0}\}$. Because $p_A(x) \neq 0$ if and only if $\mathbb{R}_+x \not\subset A$, it remains to show that $p_A(x) = p_A(-x)$ for every $x \in E \setminus \{\mathbf{0}\}$ if and only if $tA \subset A$ whenever $|t| < 1$.

Assume that $p_A(x) = p_A(-x)$ for each $x \in E$. If $x \in A$ and $|t| < 1$ then $p_A(tx) = p_A(|t|x) = |t|p_A(x) < 1$, which implies that $tx \in A$. Hence, $tA \subset A$ whenever $|t| < 1$.

Conversely, assume that $tA \subset A$ whenever $|t| < 1$. For each $s > p_A(x)$, $r^{-1}x \in A$ for some $0 < r < s$, and we have $s^{-1}(-x) = (-s^{-1}r)r^{-1}x \in A$, hence $p_A(-x) \leq p_A(x)$. Replacing x with $-x$, we have $p_A(x) \leq p_A(-x)$. Therefore, $p_A(x) = p_A(-x)$. \square

When the Minkowski functional p_A is a norm on E , we call it the **Minkowski norm**. In this case, $\text{rcl } A$, $\text{rint } A$, and ∂A are the unit closed ball, the unit open ball, and the unit sphere, respectively, of the normed linear space $E = (E, p_A)$. Then, $\text{rcl } A$ and $\text{rint } A$ are symmetric about $\mathbf{0}$, i.e., $\text{rcl } A = -\text{rcl } A$ and $\text{rint } A = -\text{rint } A$. We should note that a convex set $A \subset E$ is symmetric about $\mathbf{0}$ if and only if $tA \subset A$ whenever $|t| \leq 1$ (in the next section, A is said to be circled).

A subset $W \subset E$ is called a **wedge** if $x + y \in W$ for each $x, y \in W$ and $tx \in W$ for each $x \in W, t \geq 0$, or equivalently, W is convex and $tW \subset W$ for every $t \geq 0$. Note that if $A \subset E$ is convex then \mathbb{R}_+A is a wedge. For a wedge $W \subset E$, the following statements are true:

- (1) $\mathbf{0} \in \text{core } W \Leftrightarrow W = E$;
- (2) $W \neq E, x \in \text{core } W \Rightarrow -x \notin W$.

A **cone** $C \subset E$ is a wedge with $C \cap (-C) = \{\mathbf{0}\}$. Each translation of a cone is also called a **cone**.

Using the Hahn–Banach Extension Theorem, we can prove the following separation theorem:

Theorem 3.3.5 (SEPARATION THEOREM). *Let A and B be convex sets in E such that $\text{core } A \neq \emptyset$ and $(\text{core } A) \cap B = \emptyset$. Then, there exists a linear functional $f : E \rightarrow \mathbb{R}$ such that $f(x) < f(y)$ for every $x \in \text{core } A$ and $y \in B$, and $\sup f(A) \leq \inf f(B)$.*

Proof. Recall that $\text{core } A = \text{rint } A$ (Proposition 3.3.2). For a linear functional $f : E \rightarrow \mathbb{R}$, if $f(x) < f(y)$ for every $x \in \text{core } A$ and $y \in B$, then $\sup f(A) \leq \inf f(B)$. Indeed, let $x \in A$, $y \in B$, $v \in \text{core } A$, and $0 \leq t < 1$. Since $(1-t)v + tx \in \text{core } A$ by Proposition 3.2.3, we have

$$(1-t)f(v) + tf(x) = f((1-t)v + tx) < f(y),$$

where the left side tends to $f(x)$ as $t \rightarrow 1$, and hence $f(x) \leq f(y)$.

Note that $W = \mathbb{R}_+(A - B)$ is a wedge. Moreover, $(\text{core } A) - B \subset \text{core } W$. Indeed, let $x \in \text{core } A$ and $y \in B$. For each $z \in E$, choose $\delta > 0$ so that $(1 + \delta)x - \delta(y + z) \in A$. Then,

$$(1 + \delta)(x - y) - \delta z = (1 + \delta)x - \delta(y + z) - y \in A - B \subset W.$$

Therefore, it suffices to construct a linear functional $f : E \rightarrow \mathbb{R}$ such that $f(\text{core } W) \subset (-\infty, 0)$.

Now, we shall show that $W \cap (B - \text{core } A) = \emptyset$. Assume that there exist $x_0 \in A$, $x_1 \in \text{core } A$, $y_0, y_1 \in B$, and $t_0 \geq 0$ such that $t_0(x_0 - y_0) = y_1 - x_1$. Note that $\text{rint}\langle x_0, x_1 \rangle \subset \text{rint } A = \text{core } A$ by Proposition 3.2.3. Hence,

$$\frac{t_0}{t_0 + 1}x_0 + \frac{1}{t_0 + 1}x_1 = \frac{t_0}{t_0 + 1}y_0 + \frac{1}{t_0 + 1}y_1 \in (\text{core } A) \cap B,$$

which contradicts the fact that $(\text{core } A) \cap B = \emptyset$.

Take $v_0 \in (\text{core } A) - B \subset \text{core } W$. Then, note that $-v_0 \notin W$. For each $x \in E$, we have $\delta > 0$ such that $(1 + \delta)v_0 - \delta(-x) \in W$, which implies $x + \delta^{-1}(1 + \delta)v_0 \in W$. Then, we can define $p : E \rightarrow \mathbb{R}$ by

$$p(x) = \inf \{t \geq 0 \mid x + tv_0 \in W\}.$$

Because W is a wedge, we see that p is sublinear. Since $-v_0 \notin W$, it follows that $p(s(-v_0)) = s$ and $p(sv_0) = 0$ for every $s \geq 0$. Applying the Hahn–Banach Extension Theorem 3.3.1, we can obtain a linear functional $f : E \rightarrow \mathbb{R}$ such that $f(s(-v_0)) = s$ for each $s \in \mathbb{R}$ and $f(x) \leq p(x)$ for every $x \in E$ (see Fig. 3.3). For each $z \in \text{core } W$, we have $\delta > 0$ such that $(1 + \delta)z - \delta(z + v_0) \in W$, i.e., $z - \delta v_0 \in W$. Accordingly, $(z - \delta v_0) + tv_0 \in W$ for every $t \geq 0$, which means $p(z - \delta v_0) = 0$. Thus, we have

$$f(z) < f(z) + \delta = f(z - \delta v_0) \leq p(z - \delta v_0) = 0. \quad \square$$

Remark 6. Using the Hahn–Banach Extension Theorem, we have proved the Separation Theorem. Conversely, the Hahn–Banach Extension Theorem can be derived from the Separation Theorem. Indeed, under the assumption of the Hahn–Banach Extension Theorem 3.3.1, we define

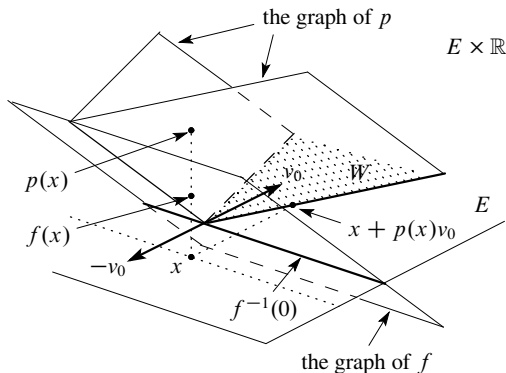


Fig. 3.3 The graphs of p and f

$$A = \{(x, t) \in E \times \mathbb{R} \mid t > p(x)\} \text{ and } B = \{(x, f(x)) \in E \times \mathbb{R} \mid x \in F\},$$

where $B = \text{Gr}(f)$ is the graph of f . Then, A and B are disjoint convex sets in $E \times \mathbb{R}$. It is straightforward to show that $\text{core } A = A \neq \emptyset$. By the Separation Theorem 3.3.5, we have a linear functional $\varphi : E \times \mathbb{R} \rightarrow \mathbb{R}$ such that $A \subset \varphi^{-1}((-\infty, r])$ and $B \subset \varphi^{-1}([r, \infty))$ for some $r \in \mathbb{R}$. Then, $r \leq 0$ because $0 = \varphi(0, 0) \in \varphi(B)$. If $\varphi(z) < 0$ for some $z \in B$, then $\varphi(tz) = t\varphi(z) < r$ for sufficiently large $t > 0$. This is a contradiction because $tz \in B$. If $\varphi(z) > 0$ for some $z \in B$, then $-z \in B$ and $\varphi(-z) = -\varphi(z) < 0$, which is a contradiction. Therefore, $B \subset \varphi^{-1}(0)$. Note that $\varphi(0, 1) < 0$ because $(0, 1) \in A$. Since $\varphi(x, t) = \varphi(x, 0) + t\varphi(0, 1)$ for each $x \in E$, we have $\varphi(\{x\} \times \mathbb{R}) = \mathbb{R}$. Observe that $(\{x\} \times \mathbb{R}) \cap \varphi^{-1}(0)$ is a singleton. Then, f extends to the linear functional $\tilde{f} : E \rightarrow \mathbb{R}$ whose graph is $\varphi^{-1}(0)$, i.e., $(x, \tilde{f}(x)) \in \varphi^{-1}(0)$ for each $x \in E$. Since $\varphi^{-1}(0) \subset (E \times \mathbb{R}) \setminus A$, it follows that $\tilde{f}(x) \leq p(x)$ for every $x \in E$.

The Separation Theorem 3.3.5 can also be obtained as a corollary of the following two theorems, where we do not use the Hahn–Banach Extension Theorem 3.3.1.

Theorem 3.3.6. *For each pair of disjoint non-empty convex sets $A, B \subset E$, there exists a pair of disjoint convex sets $\tilde{A}, \tilde{B} \subset E$ such that $A \subset \tilde{A}$, $B \subset \tilde{B}$, and $\tilde{A} \cup \tilde{B} = E$.*

Proof. Let \mathcal{P} be the collection of pairs (C, D) of disjoint convex sets such that $A \subset C$ and $B \subset D$. For $(C, D), (C', D') \in \mathcal{P}$, we define $(C, D) \leq (C', D')$ if $C \subset C'$ and $D \subset D'$. Then, it is easy to see that $\mathcal{P} = (\mathcal{P}, \leq)$ is an inductive ordered set. Due to Zorn’s Lemma, \mathcal{P} has a maximal element (\tilde{A}, \tilde{B}) .

To show that $\tilde{A} \cup \tilde{B} = E$, assume the contrary, i.e., there exists a point $v_0 \in E \setminus (\tilde{A} \cup \tilde{B})$. By the maximality of (\tilde{A}, \tilde{B}) , we can obtain two points

$$x \in \tilde{A} \cap \langle \tilde{B} \cup \{v_0\} \rangle \text{ and } y \in \tilde{B} \cap \langle \tilde{A} \cup \{v_0\} \rangle.$$

Then, $x \in \langle v_0, y_1 \rangle$ for some $y_1 \in \widetilde{B}$ and $y \in \langle v_0, x_1 \rangle$ for some $x_1 \in \widetilde{A}$. Note that $x \in \text{rint}\langle v_0, y_1 \rangle$ and $y \in \text{rint}\langle v_0, x_1 \rangle$. Consider the triangle $\langle v_0, \widetilde{x}_1, y_1 \rangle$. It is easy to see that $\langle x_1, x \rangle$ and $\langle y_1, y \rangle$ meet at a point v_1 . Since $\langle x_1, x \rangle \subset \widetilde{A}$ and $\langle y_1, y \rangle \subset \widetilde{B}$, it follows that $v_1 \in \widetilde{A} \cap \widetilde{B}$, which is a contradiction. \square

Theorem 3.3.7. *For each pair of disjoint non-empty convex sets $C, D \subset E$ with $C \cup D = E$, $\text{rcl } C \cap \text{rcl } D$ is a hyperplane if $\text{rcl } C \cap \text{rcl } D \neq E$.*

Proof. First, we show that $\text{rcl } C \cap \text{rcl } D = \partial C = \partial D$. To prove that $\partial C \subset \partial D$, let $x \in \partial C$. It suffices to find $y \in C$ such that

$$(1-t)x + ty \in C \text{ for } 0 < t \leq 1 \text{ and} \\ (1+t)x - ty \in E \setminus C = D \text{ for } t > 0.$$

To this end, take $y', y'' \in C$ such that $(1-t)x + ty' \in C$ for $0 < t \leq 1$ and $(1+t)x - ty'' \notin C$ for $t > 0$. Then, $y = \frac{1}{2}y' + \frac{1}{2}y'' \in C$ is the desired point. Indeed, for each $0 < t \leq 1$,

$$(1-t)x + ty = (1-t)x + \frac{1}{2}ty' + \frac{1}{2}ty'' \\ = (1 - \frac{1}{2}t) \left(\frac{1-t}{1 - \frac{1}{2}t}x + \frac{\frac{1}{2}t}{1 - \frac{1}{2}t}y' \right) + \frac{1}{2}ty'' \in C.$$

Moreover, note that

$$(1-s)((1+t)x - ty) + sy' \\ = (1-s)(1+t)x - \frac{1}{2}(1-s)ty' - \frac{1}{2}(1-s)ty'' + sy'.$$

For each $t > 0$, let $s = t/(2+t) \in (0, 1)$. Then, $(1-s)t = 2s$. Therefore, we have

$$(1-s)((1+t)x - ty) + sy' = (1+s)x - sy'' \notin C,$$

which means that $(1+t)x - ty \notin C$ (Fig. 3.4). Similarly, we have $\partial D \subset \partial C$. Hence, $\partial C = \partial D$. Since $\text{rint } C \cap \text{rint } D = \emptyset$, it follows that $\text{rcl } C \cap \text{rcl } D = \partial C = \partial D$.

Next, we show that ∂C is a flat. It suffices to show that if $x, y \in \partial C$ and $t > 0$, then $x' = (1+t)x - ty \in \partial C$. If $x' \notin \partial C$, then $x' \in \text{rint } C$ or $x' \in \text{rint } D$. In this case, $x \in \text{rint}\langle x', y \rangle \subset \text{rint } C$ or $x \in \text{rint}\langle x', y \rangle \subset \text{rint } D$ by Proposition 3.2.3. This is a contradiction. Therefore, $x' \in \partial C$.

It remains to show that if $\partial C \neq E$ then ∂C is a hyperplane. We have $v \in E \setminus \partial C$. It suffices to prove that $E = \text{fl}(\partial C \cup \{v\})$. Without loss of generality, we may assume that $v \in \text{rint } C$. On the other hand, $\partial C \neq \emptyset$ because $C \neq E$. Let $z \in \partial C$. Then, $w = z - (v - z) = 2z - v \in \text{rint } D$. Otherwise, $w \in \text{rcl } C$, from which, using Proposition 3.2.3, it would follow that $z = \frac{1}{2}v + \frac{1}{2}w \in \text{rint}\langle v, w \rangle \subset \text{rint } C$, which is a contradiction.

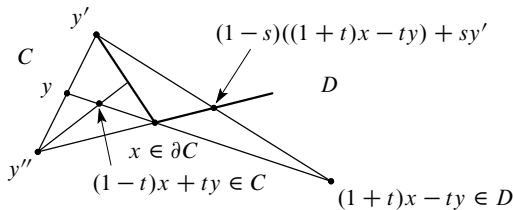


Fig. 3.4 $\partial C \subset \partial D$

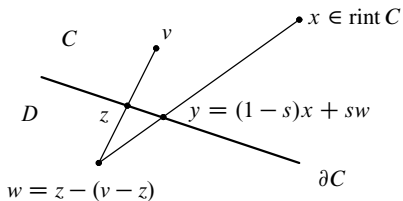


Fig. 3.5 The case $x \in \text{rint } C$

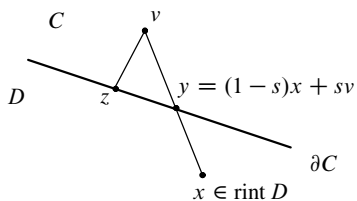


Fig. 3.6 The case $x \in \text{rint } D$

For each $x \in E \setminus \partial C$, $x \in \text{rint } C$ or $x \in \text{rint } D$. When $x \in \text{rint } C$, let

$$s = \sup \{t \in \mathbf{I} \mid (1-t)x + tw \in C\}.$$

Refer to Fig. 3.5. Then, $y = (1-s)x + sw \in \partial C$, which implies that

$$x = \frac{1}{1-s}y - \frac{2s}{1-s}z + \frac{s}{1-s}v \in \text{fl}(\partial C \cup \{v\}).$$

In the case that $x \in \text{rint } D$, let

$$s = \sup \{t \in \mathbf{I} \mid (1-t)x + tv \in D\}.$$

Now, refer to Fig. 3.6. Then, $y = (1-s)x + sv \in \partial D = \partial C$, which implies that

$$x = \frac{1}{1-s}y + \frac{-s}{1-s}v \in \text{fl}(\partial C \cup \{v\}).$$

Consequently, it follows that $E = \text{fl}(\partial C \cup \{v\})$. □

Remark 7. In the above, the condition $\text{rcl } C \cap \text{rcl } D \neq E$ is necessary. For example, define the convex set C in the linear space $\mathbb{R}_f^{\mathbb{N}}$ as follows:

$$C = \{x \in \mathbb{R}_f^{\mathbb{N}} \mid n = \max\{i \mid x(i) \neq 0\} \Rightarrow x(n) > 0\}.$$

Let $D = \mathbb{R}_f^{\mathbb{N}} \setminus C = (-C) \setminus \{0\}$. Then, D is also convex. As is easily observed, $\text{rcl } C = \text{rcl } D = \mathbb{R}_f^{\mathbb{N}}$, hence $\text{rcl } C \cap \text{rcl } D = \mathbb{R}_f^{\mathbb{N}}$.

The Separation Theorem 3.3.5 can also be obtained as a corollary of Theorems 3.3.6 and 3.3.7. In fact, let $A, B \subset E$ be convex sets with $\text{core } A \neq \emptyset$ and $(\text{core } A) \cap B = \emptyset$. Then, $\text{core } A = \text{rint } A$ is convex. We apply Theorem 3.3.6 to obtain disjoint non-empty convex sets C and D such that $\text{core } A \subset C$, $B \subset D$, and $C \cup D = E$. Observe that $\text{core } A \cap \text{rcl } D = \emptyset$, hence $\text{rcl } D \neq E$. It follows from Theorem 3.3.7 that $\text{rcl } C \cap \text{rcl } D$ is a hyperplane. Then, we have a linear functional $f : E \rightarrow \mathbb{R}$ such that $\text{rcl } C \cap \text{rcl } D = f^{-1}(s)$ for some $s \in \mathbb{R}$ (Proposition 3.1.3(1)). Since $\text{core } A \subset E \setminus f^{-1}(s)$, we have $\text{core } A \subset f^{-1}((s, \infty))$ or $\text{core } A \subset f^{-1}((-\infty, s))$. If $\text{core } A \subset f^{-1}((s, \infty))$, by replacing f and s by $-f$ and $-s$, it can be assumed that $\text{core } A \subset f^{-1}((-\infty, s))$.

We now show that $\text{rcl } C \subset f^{-1}((-\infty, s])$. Let $x \in \text{core } A (\subset \text{rint } C)$. Then, $x \in \text{rint } C$ and $f(x) < s$. If $f(y) > s$ for some $y \in \text{rcl } C$, we have $z \in \text{rint}\langle x, y \rangle \cap f^{-1}(s)$. Because $z \in \text{rcl } D$, $\text{rint}\langle w, z \rangle \subset D$ for some $w \in D$. On the other hand, $z \in \text{rint}\langle x, y \rangle \subset \text{rint } C$ (Proposition 3.2.3). Because $\text{rint } C = \text{core } C$, $\langle v, z \rangle \subset C = E \setminus D$ for some $v \in \text{rint}\langle w, z \rangle$, which is a contradiction.

Since $C \subset f^{-1}((-\infty, s])$, it follows that $D \supset f^{-1}((s, \infty))$. Observe that $\text{rint } D \supset f^{-1}((s, \infty))$. So, we have $x \in \text{rint } D$ and $f(x) > s$. Likewise for $\text{rcl } D$, we can show that $\text{rcl } D \subset f^{-1}([s, \infty))$. Accordingly, we have

$$\text{rcl } C = f^{-1}((-\infty, s]) \text{ and } \text{rcl } D = f^{-1}([s, \infty)).$$

Since $\text{core } A \subset f^{-1}((-\infty, s))$ and $B \subset f^{-1}([s, \infty))$, we have the desired result.

3.4 Topological Linear Spaces

A **topological linear space** E is a linear space with a topology such that the algebraic operations of addition $(x, y) \mapsto x + y$ and scalar multiplication $(t, x) \mapsto tx$ are continuous.⁸ Every linear space E has such a topology. In fact,

⁸Here, we only consider linear spaces over \mathbb{R} . Recall that topological spaces are assumed to be **Hausdorff**. For topological linear spaces (more generally for topological groups), it suffices to assume axiom T_0 , which implies **regularity** (Proposition 3.4.2 and its footnote). The continuity of scalar multiplication implies the continuity of the operation $x \mapsto -x$ because $(-1)x = -x$.

E has a Hamel basis B . As a linear subspace of the product space \mathbb{R}^B , \mathbb{R}_f^B is a topological linear space that is linearly isomorphic to E by the linear isomorphism $\varphi : \mathbb{R}_f^B \rightarrow E$ defined by $\varphi(x) = \sum_{v \in B} x(v)v$. Then, φ induces a topology that makes E a topological linear space. In the next section, it will be seen that if E is finite-dimensional, then such a topology is unique. However, an infinite-dimensional linear space has various topologies for which the algebraic operations are continuous.

In the following proposition, we present the basic properties of a neighborhood basis at $\mathbf{0}$ in a topological linear space.

Proposition 3.4.1. *Let E be a topological linear space and \mathcal{U} be a neighborhood basis at $\mathbf{0}$ in E . Then, \mathcal{U} has the following properties:*

- (1) For each $U, V \in \mathcal{U}$, there is some $W \in \mathcal{U}$ such that $W \subset U \cap V$;
- (2) For each $U \in \mathcal{U}$, there is some $V \in \mathcal{U}$ such that $V + V \subset U$;
- (3) For each $U \in \mathcal{U}$, there is some $V \in \mathcal{U}$ such that $[-1, 1]V \subset U$;
- (4) For each $x \in E$ and $U \in \mathcal{U}$, there is some $a > 0$ such that $x \in aU$;
- (5) $\bigcap \mathcal{U} = \{\mathbf{0}\}$.

Conversely, let E be a linear space with \mathcal{U} a collection of subsets satisfying these conditions. Then, E has a topology such that addition and scalar multiplication are continuous and \mathcal{U} is a neighborhood basis at $\mathbf{0}$.

Sketch of Proof. Property (1) is trivial; (2) comes from the continuity of addition at $(0, 0) \in E \times E$; (3) is obtained by the continuity of scalar multiplication at each $(t, 0) \in [-1, 1] \times E$ and the compactness of $[-1, 1]$; (4) follows from the continuity of scalar multiplication at $(0, x) \in \mathbb{R} \times E$; the Hausdorffness of E implies (5).

Given \mathcal{U} with these properties, an open set in E is defined as a subset $W \subset E$ satisfying the condition that, for each $x \in W$, there is some $U \in \mathcal{U}$ such that $x + U \subset W$. (Verify the axioms of open sets, i.e., the intersection of finite open sets is open; every union of open sets is open.)

For each $x \in E$ and $U \in \mathcal{U}$, $x + U$ is a neighborhood of x in this topology.⁹ Indeed, let

$$W = \{y \in E \mid \exists V \in \mathcal{U} \text{ such that } y + V \subset x + U\}.$$

Then, $x \in W \subset x + U$ because of (5). For each $y \in W$, we have $V \in \mathcal{U}$ such that $y + V \subset x + U$. Take $V' \in \mathcal{U}$ so that $V' + V' \subset V$ as in (2). Then, $y + V' \subset W$ because $(y + y') + V' \subset y + V \subset x + U$ for every $y' \in V'$. Therefore, W is open in E , so $x + U$ is a neighborhood of x in E . By the definition of the topology, $\{x + U \mid U \in \mathcal{U}\}$ is a neighborhood basis at x . In particular, \mathcal{U} is a neighborhood basis at $\mathbf{0}$.

Since $\{x + U \mid U \in \mathcal{U}\}$ is a neighborhood basis at x , the continuity of addition follows from (2). Using (3), we can show that the operation $x \mapsto -x$ is continuous.

For scalar multiplication, let $x \in E$, $\alpha \in \mathbb{R}$, and $U \in \mathcal{U}$. Because of the continuity of $x \mapsto -x$, it can be assumed that $\alpha \geq 0$. Then, we can write $\alpha = n + t$, where $n \in \omega$ and $0 \leq t < 1$. Using (2) inductively, we can find $V_1 \supset \cdots \supset V_n \supset V_{n+1}$ in \mathcal{U} such that

⁹If E is a topological linear space, $x + U$ is a neighborhood of $x \in E$ for any neighborhood U of $\mathbf{0}$.

$$V_1 + \cdots + V_n + (V_{n+1} + V_{n+1}) \subset U.$$

By (3), we have $W \in \mathcal{U}$ such that $[-1, 1]W \subset V_{n+1}$. Then, $x \in rW$ for some $r > 0$ by (4). Choose $\delta > 0$ so that $\delta < \min\{1/r, 1-t\}$. Let $y \in x + W$ and $|\alpha - \beta| < \delta$. Then, we can write $\beta = n + s$, where $t - \delta < s < t + \delta$. It follows that

$$\begin{aligned} \beta y - \alpha x &= (n + s)y - (n + t)x = n(y - x) + s(y - x) + (s - t)x \\ &\in nW + [-1, 1]W + \delta[-1, 1](rW) \\ &\subset nV_{n+1} + V_{n+1} + V_{n+1} \\ &\subset \underbrace{V_{n+1} + \cdots + V_{n+1}}_{n+2 \text{ many}} \subset U,^{10} \end{aligned}$$

hence $\beta y \in \alpha x + U$.

To see the Hausdorffness, let $x \neq y \in E$. By (5), we have $U \in \mathcal{U}$ such that $x - y \notin U$. By (2) and (3), we can find $V \in \mathcal{U}$ such that $V - V \subset U$. Then, $x + V$ and $y + V$ are neighborhoods of x and y , respectively. Observe that $(x + V) \cap (y + V) = \emptyset$.

It is said that $A \subset E$ is **circled** if $tA \subset A$ for every $t \in [-1, 1]$. It should be noted that the closure of a circled set A is also circled.

Indeed, let $x \in \text{cl } A$ and $t \in [-1, 1]$. If $t = 0$, then $tx = 0 \in A \subset \text{cl } A$. When $t \neq 0$, for each neighborhood U of tx in E , since $t^{-1}U$ is a neighborhood of $t^{-1}x$, $t^{-1}U \cap A \neq \emptyset$, which implies that $U \cap tA \neq \emptyset$. Because $tA \subset A$, $U \cap A \neq \emptyset$. Thus, it follows that $tx \in \text{cl } A$.

In (3) above, $W = [-1, 1]V$ is a neighborhood of $\mathbf{0} \in E$ that is circled, i.e., $tW \subset W$ for every $t \in [-1, 1]$. Consequently, (3) is equivalent to the following condition:

(3)' $\mathbf{0} \in E$ has a neighborhood basis consisting of circled (open) sets.

A **topological group** G is a group with a topology such that the algebraic operations of multiplication $(x, y) \mapsto xy$ and taking inverses $x \mapsto x^{-1}$ are both continuous.¹¹ Then, G is **homogeneous**, that is, for each distinct $x_0, x_1 \in G$, there is a homeomorphism $h : G \rightarrow G$ such that $h(x_0) = x_1$. Such an h can be defined by $h(x) = x_0x^{-1}x_1$, where not only $h(x_0) = x_1$ but also $h(x_1) = x_0$. Every topological linear space is a topological group with respect to addition, so it is homogeneous.

Proposition 3.4.2. *Every topological group G has a closed neighborhood basis at each $g \in G$, i.e., it is regular.¹² For a topological linear space E , $\mathbf{0} \in E$ has a circled closed neighborhood basis.*

¹⁰It should be noted that, in general, $2V \subset V + V$ but $V + V \not\subset 2V$.

¹¹These two operations are continuous if and only if the operation $(x, y) \rightarrow x^{-1}y$ is continuous.

¹²A topological group G is assumed to be **Hausdorff**, but it suffices to assume axiom T_0 . In fact, axiom T_0 implies T_1 for a topological group G because of the homogeneity of G .

Sketch of Proof. Each neighborhood U of the unit $\mathbf{1} \in G$ contains a neighborhood V of $\mathbf{1}$ such that $V^{-1}V \subset U$. For each $x \in \text{cl } V$, we have $y \in Vx \cap V$. Consequently, $x \in V^{-1}y \subset V^{-1}V \subset U$, so we have $\text{cl } V \subset U$.

For the additional statement, recall that if V is circled then $\text{cl } V$ is also circled.

Proposition 3.4.3. *Let G be a topological group and H be a subgroup of G .*

- (1) *If H is open in G then H is closed in G .*
- (2) *The closure $\text{cl } H$ of H is a subgroup of G .*

Sketch of Proof. (1): For each $x \in G \setminus H$, Hx is an open neighborhood of x in G and $Hx \subset G \setminus H$.

(2): For each $x, y \in \text{cl } H$, show that $x^{-1}y \in \text{cl } H$, i.e., each neighborhood W of $x^{-1}y$ meets H . To this end, choose neighborhoods U and V of x and y , respectively, so that $U^{-1}V \subset W$.

Due to Proposition 3.4.3(1), a connected topological group G has no open subgroups except for G itself. Observe that every topological linear space E is path-connected. Consequently, E has no open linear subspaces except for E itself, i.e., every proper linear subspace of E is not open in E .

The continuity of linear functionals is characterized as follows:

Proposition 3.4.4. *Let E be a topological linear space. For a linear functional $f : E \rightarrow \mathbb{R}$ with $f(E) \neq \{0\}$, the following are equivalent:*

- (a) *f is continuous;*
- (b) *$f^{-1}(0)$ is closed in E ;*
- (c) *$f^{-1}(0)$ is not dense in E ;*
- (d) *$f(V)$ is bounded for some neighborhood V of $\mathbf{0} \in E$.*

Proof. The implication (a) \Rightarrow (b) is obvious, and (b) \Rightarrow (c) follows from $f(E) \neq \{0\}$ (i.e., $f^{-1}(0) \neq E$).

(c) \Rightarrow (d): We have $x \in E$ and a circled neighborhood V of $\mathbf{0} \in E$ such that $(x + V) \cap f^{-1}(0) = \emptyset$. Then, $f(V)$ is bounded. Indeed, if $f(V)$ is unbounded, then there is some $z \in V$ such that $|f(z)| > |f(x)|$. In this case, $f(tz) = tf(z) = -f(x)$ for some $t \in [-1, 1]$, which implies that $-f(x) \in f(V)$. It follows that $0 \in f(x) + f(V) = f(x + V)$, which contradicts the fact that $(x + V) \cap f^{-1}(0) = \emptyset$.

(d) \Rightarrow (a): For each $\varepsilon > 0$, we have $n \in \mathbb{N}$ such that $f(V) \subset (-n\varepsilon, n\varepsilon)$. Then, $n^{-1}V$ is a neighborhood of $\mathbf{0}$ in E and $f(n^{-1}V) \subset (-\varepsilon, \varepsilon)$. Therefore, f is continuous at $\mathbf{0} \in E$. Since f is linear, it follows that f is continuous at every point of E . □

Proposition 3.4.5. *Let E be a topological linear space and $A, B \subset E$.*

- (1) *If B is open in E then $A + B$ is open in E .*
- (2) *If A is compact and B is closed in E then $A + B$ is also closed in E .*

Sketch of Proof. (1): Note that $A + B = \bigcup_{x \in A} (x + B)$.

(2): To show that $E \setminus (A + B)$ is open in E , let $z \in E \setminus (A + B)$. For each $x \in A$, because $z - x \in E \setminus B$, we have open neighborhoods U_x, V_x of x, z in E such that $V_x - U_x \subset E \setminus B$. Since A is compact, $A \subset \bigcup_{i=1}^n U_{x_i}$ for some $x_1, \dots, x_n \in A$. Then, $V = \bigcap_{i=1}^n V_{x_i}$ is an open neighborhood of z in E . We can show that $V \cap (A + B) = \emptyset$, i.e., $V \subset E \setminus (A + B)$.

Remark 8. In (2) above, we cannot assert that $A + B$ is closed in E even if both A and B are closed and convex in E . For example, $A = \mathbb{R} \times \{0\}$ and $B = \{(x, y) \in \mathbb{R}^2 \mid x > 0, y \geq x^{-1}\}$ are closed convex sets in \mathbb{R}^2 , but $A + B = \mathbb{R} \times (0, \infty)$ is not closed in \mathbb{R}^2 .

Proposition 3.4.6. *Let F be a closed linear subspace of a topological linear space E . Then, the quotient linear space E/F with the quotient topology is also a topological linear space, and the quotient map $q : E \rightarrow E/F$ (i.e., $p(x) = x + F \in E/F$) is open, hence if \mathcal{U} is a neighborhood basis at $\mathbf{0}$ in E , then $q(\mathcal{U}) = \{q(U) \mid U \in \mathcal{U}\}$ is a neighborhood basis $\mathbf{0}$ in E/F .*

Sketch of Proof. Apply Proposition 3.4.5(1) to show that the quotient map $q : E \rightarrow E/F$ is open. Then, in the diagrams below, $q \times q$ and $q \times \text{id}_{\mathbb{R}}$ are open, so they are quotient maps:

$$\begin{array}{ccc}
 E \times E & \longrightarrow & E \\
 q \times q \downarrow & & \downarrow q \\
 E/F \times E/F & \longrightarrow & E/F,
 \end{array}
 \qquad
 \begin{array}{ccc}
 E \times \mathbb{R} & \longrightarrow & E \\
 q \times \text{id}_{\mathbb{R}} \downarrow & & \downarrow q \\
 E/F \times \mathbb{R} & \longrightarrow & E/F.
 \end{array}$$

Accordingly, the continuity of addition and scalar multiplication are clear. Note that E/F is Hausdorff if and only if F is closed in E .

For convex sets in a topological linear space, we have the following:

Proposition 3.4.7. *For each convex set C in a topological linear space E , the following hold:*

- (1) $\text{cl } C$ is convex and $\text{rcl } C \subset \text{cl } C$, hence $\text{rcl } C = C$ if C is closed in E ;
- (2) $\text{int}_F C = \emptyset$ for any flat F with $\text{fl } C \subsetneq F$;
- (3) $\text{int}_{\text{fl } C} C \neq \emptyset$ implies $\text{int}_{\text{fl } C} C = \text{core}_{\text{fl } C} C = \text{rint } C$.

Proof. By the definition and the continuity of algebraic operations, we can easily obtain (1). For (2), observe $\text{int}_F C \subset \text{core}_F C$. If $\text{int}_F C \neq \emptyset$ then $\text{fl } C = F$ by the Fact stated in the previous section.

(3): Due to Proposition 3.3.2, $\text{core}_{\text{fl } C} C = \text{rint } C$. Note that $\text{int}_{\text{fl } C} C \subset \text{core}_{\text{fl } C} C$. Without loss of generality, we may assume that $\mathbf{0} \in \text{int}_{\text{fl } C} C$. Then, for each $x \in \text{rint } C$, we can find $0 < s < 1$ such that $x \in sC$. Since $(1 - s)C$ is a neighborhood of $\mathbf{0} = x - x$ in $\text{fl } C$, we have a neighborhood U of x in $\text{fl } C$ such that $U - x \subset (1 - s)C$. Then, it follows that $U \subset (1 - s)C + sC = C$. Therefore, $x \in \text{int}_{\text{fl } C} C$. □

Remark 9. In the above, we cannot assert any one of $\text{cl } C = \text{rcl } C$, $\text{int}_{\text{fl } C} C = \text{core}_{\text{fl } C} C$, or $\text{int}_{\text{fl } C} C \neq \emptyset$. For example, $[-1, 1]_f^{\mathbb{N}}$ is a convex set in $\mathbb{R}^{\mathbb{N}}$ such that

$\text{rc}[-1, 1]_f^{\mathbb{N}} = [-1, 1]_f^{\mathbb{N}}$ but $\text{cl}[-1, 1]_f^{\mathbb{N}} = [-1, 1]^{\mathbb{N}}$. Note that $\text{fl}[-1, 1]_f^{\mathbb{N}} = \mathbb{R}_f^{\mathbb{N}}$. Regard $[-1, 1]_f^{\mathbb{N}}$ as a convex set in $\mathbb{R}_f^{\mathbb{N}}$. Then,

$$\text{int}_{\mathbb{R}_f^{\mathbb{N}}}[-1, 1]_f^{\mathbb{N}} = \emptyset \text{ but } \text{core}_{\mathbb{R}_f^{\mathbb{N}}}[-1, 1]_f^{\mathbb{N}} = \text{rint}[-1, 1]_f^{\mathbb{N}} = (-1, 1)_{f, \mathbb{N}}.$$

By Proposition 3.4.7(1), if A is a subset of a topological linear space E , then $\text{cl}\langle A \rangle$ is the smallest closed convex set containing A , which is called the **closed convex hull** of A .

Remark 10. In general, $\langle A \rangle$ is not closed in E even if A is compact. For example, let $A = \{a_n \mid n \in \omega\} \subset \ell_1$, where $a_0(i) = 2^{-i}$ for every $i \in \mathbb{N}$ and, for each $n \in \mathbb{N}$, $a_n(i) = 2^{-i}$ if $i \leq n$ and $a_n(i) = 0$ if $i > n$. Then, A is compact and $\langle A \rangle = \bigcup_{n \in \mathbb{N}} \langle a_0, a_1, \dots, a_n \rangle$. For each $n \in \mathbb{N}$, let

$$x_n = 2^{-n}a_0 + 2^{-1}a_1 + \dots + 2^{-n}a_n \in \langle a_0, a_1, \dots, a_n \rangle.$$

Then, $x_n(i) = 2^{-2i+1}$ if $i \leq n$ and $x_n(i) = 2^{-n-i}$ if $i > n$. Hence, $(x_n)_{n \in \mathbb{N}}$ converges to $x_0 \in \ell_1$, where $x_0(i) = 2^{-2i+1}$ for each $i \in \mathbb{N}$. However, $x_0 \notin \langle A \rangle$. Otherwise, $x_0 \in \langle a_0, a_1, \dots, a_n \rangle$ for some $n \in \mathbb{N}$, where we can write

$$x_0 = \sum_{i=0}^n z(i+1)a_i, \quad z \in \Delta^n.$$

Then, we have the following:

$$\begin{aligned} z(1)a_0(n+1) &= x_0(n+1) = 2^{-2n-1} = 2^{-n}a_0(n+1) \text{ and} \\ z(1)a_0(n+2) &= x_0(n+2) = 2^{-2n-3} = 2^{-n-1}a_0(n+2), \end{aligned}$$

hence $z(1) = 2^{-n}$ and $z(1) = 2^{-n-1}$. This is a contradiction. Therefore, $\langle A \rangle$ is not closed in ℓ_1 .

The following is the topological version of the Separation Theorem 3.3.5:

Theorem 3.4.8 (SEPARATION THEOREM). *Let A and B be convex sets in a topological linear space E such that $\text{int } A \neq \emptyset$ and $(\text{int } A) \cap B = \emptyset$. Then, there is a continuous linear functional $f : E \rightarrow \mathbb{R}$ such that $f(x) < f(y)$ for each $x \in \text{int } A$ and $y \in B$, and $\sup f(A) \leq \inf f(B)$.*

Proof. First, $\text{int } A \neq \emptyset$ implies $\text{core } A = \text{int } A \neq \emptyset$ by Proposition 3.4.7(3). Then, by the Separation Theorem 3.3.5, we have a linear functional $f : E \rightarrow \mathbb{R}$ such that $f(x) < f(y)$ for every $x \in \text{int } A$ and $y \in B$, and $\sup f(A) \leq \inf f(B)$. Note that $B - \text{int } A$ is open in E and $f(z) > 0$ for every $z \in B - \text{int } A$. Thus, $f^{-1}(0)$ is not dense in E . Therefore, f is continuous by Proposition 3.4.4. \square

A topological linear space E is **locally convex** if $\mathbf{0} \in E$ has a neighborhood basis consisting of (open) convex sets; equivalently, open convex sets make up an open basis for E . It follows from Proposition 3.4.6 that for each locally convex topological space E and each closed linear subspace $F \subset E$, the quotient linear space E/F is also locally convex. For locally convex topological linear spaces, we have the following separation theorem:

Theorem 3.4.9 (STRONG SEPARATION THEOREM). *Let A and B be disjoint closed convex sets in a locally convex topological linear space E . If at least one of A and B is compact, then there is a continuous linear functional $f : E \rightarrow \mathbb{R}$ such that $\sup f(A) < \inf f(B)$.*

Proof. By Proposition 3.4.5(2), $B - A$ is closed in E . Since $A \cap B = \emptyset$, it follows that $\mathbf{0} \notin B - A$. Choose an open convex neighborhood U of $\mathbf{0}$ so that $U \cap (B - A) = \emptyset$. By the Separation Theorem 3.4.8, we have a nontrivial continuous linear functional $f : E \rightarrow \mathbb{R}$ such that $\sup f(U) \leq \inf f(B - A)$. Then, $\sup f(A) + \sup f(U) \leq \inf f(B)$, where $\sup f(U) > 0$ by the non-triviality of f . Thus, we have the result. \square

As a particular case, we have the following:

Corollary 3.4.10. *Let E be a locally convex topological linear space. For each pair of distinct points $x, y \in E$, there exists a continuous linear functional $f : E \rightarrow \mathbb{R}$ such that $f(x) \neq f(y)$.* \square

Concerning the continuity of sublinear functionals, we have the following:

Proposition 3.4.11. *Let $p : E \rightarrow \mathbb{R}$ be a non-negative sublinear functional of a topological linear space E . Then, p is continuous if and only if $p^{-1}([0, 1])$ is a neighborhood of $\mathbf{0} \in E$.*

Proof. The “only if” part follows from $p^{-1}([0, 1]) = p^{-1}((-1, 1))$. To see the “if” part, let $\varepsilon > 0$. Since $p^{-1}([0, \varepsilon]) = \varepsilon p^{-1}([0, 1])$ is a neighborhood of $\mathbf{0} \in E$, each $x \in E$ has the following neighborhood:

$$U = (x + p^{-1}([0, \varepsilon])) \cap (x - p^{-1}([0, \varepsilon])).$$

For each $y \in U$, since $p(y - x) < \varepsilon$ and $p(x - y) < \varepsilon$, it follows that

$$\begin{aligned} p(y) &\leq p(y - x) + p(x) < p(x) + \varepsilon \quad \text{and} \\ p(y) &\geq p(x) - p(x - y) > p(x) - \varepsilon, \end{aligned}$$

which means that p is continuous at x . \square

For each convex set $C \subset E$ with $\mathbf{0} \in \text{int } C$, we have $\text{int } C = \text{core } C = p_C^{-1}([0, 1])$ by Propositions 3.3.4 and 3.4.7(3). Then, the following is obtained from Proposition 3.4.11.

Corollary 3.4.12. *Let E be a topological linear space. For each convex set $C \subset E$ with $\mathbf{0} \in \text{int } C$, the Minkowski functional $p_C : E \rightarrow \mathbb{R}$ is continuous. Moreover, $p_C^{-1}([0, 1)) = \text{int } C = \text{rint } C$ and $p_C^{-1}(\mathbf{1}) = \text{cl } C = \text{rcl } C$, hence $p_C^{-1}(1) = \text{bd } C = \partial C$. \square*

The boundedness is a metric concept, but it can be extended to subsets of a topological linear space E . A subset $A \subset E$ is **topologically bounded**¹³ provided that, for each neighborhood U of $\mathbf{0} \in E$, there exists some $r > 0$ such that $A \subset rU$. If $A \subset E$ is topologically bounded and $B \subset A$, then B is also topologically bounded. Recall that every neighborhood U of $\mathbf{0} \in E$ contains a circled neighborhood V of $\mathbf{0} \in E$ (cf. Proposition 3.4.1(3)). Since $sV \subset tV$ for $0 < s < t$, it is easy to see that every compact subset of E is topologically bounded. When E is a normed linear space, $A \subset E$ is topologically bounded if and only if A is bounded in the metric sense. Applying Minkowski functionals, we can show the following:

Theorem 3.4.13. *Let E be a topological linear space. Each pair of topologically bounded closed convex sets $C, D \subset E$ with $\text{int } C \neq \emptyset$ and $\text{int } D \neq \emptyset$ are homeomorphic to each other by a homeomorphism of E onto itself, hence $(C, \text{bd } C) \approx (D, \text{bd } D)$ and $\text{int } C \approx \text{int } D$.*

Proof. Without loss of generality, we may assume that $\mathbf{0} \in \text{int } C \cap \text{int } D$. Let p_C and p_D be the Minkowski functionals for C and D , respectively. By the topological boundedness of C and D , it is easy to see that $p_C(x), p_D(x) > 0$ for every $x \in E \setminus \{\mathbf{0}\}$. Then, we can define maps $\varphi, \psi : E \rightarrow E$ as follows: $\varphi(\mathbf{0}) = \psi(\mathbf{0}) = \mathbf{0}$,

$$\varphi(x) = \frac{p_C(x)}{p_D(x)}x \quad \text{and} \quad \psi(x) = \frac{p_D(x)}{p_C(x)}x \quad \text{for each } x \in E \setminus \{\mathbf{0}\}.$$

It follows from the continuity of p_C and p_D (Corollary 3.4.12) that φ and ψ are continuous at each $x \in E \setminus \{\mathbf{0}\}$.

To verify the continuity of φ at $\mathbf{0} \in E$, let U be a neighborhood of $\mathbf{0} \in E$. Since D is topologically bounded and C is a neighborhood of $\mathbf{0}$, there is an $r > 0$ such that $D \subset rC$. Then, $p_C(x) \leq rp_D(x)$ for every $x \in E$. Choose a circled neighborhood V of $\mathbf{0} \in E$ so that $rV \subset U$. Then, $\varphi(V) \subset U$. Indeed, for each $x \in V \setminus \{\mathbf{0}\}$,

$$\varphi(x) = \frac{p_C(x)}{p_D(x)}x \in \frac{p_C(x)}{p_D(x)}V \subset rV \subset U.$$

Similarly, ψ is continuous at $\mathbf{0} \in E$.

For each $x \in E \setminus \{\mathbf{0}\}$, since $\varphi(x) \neq \mathbf{0}$,

¹³Usually, we say simply *bounded* but here add *topologically* in order to distinguish the metric sense. It should be noted that every metrizable space has an admissible bounded metric.

$$\psi\varphi(x) = \frac{p_D(\varphi(x))}{p_C(\varphi(x))}\varphi(x) = \frac{\frac{p_C(x)}{p_D(x)}p_D(x)}{\frac{p_C(x)}{p_D(x)}p_C(x)} \cdot \frac{p_C(x)}{p_D(x)}x = x.$$

Hence, $\psi\varphi = \text{id}$. Similarly, $\varphi\psi = \text{id}$. Therefore, φ is a homeomorphism with $\varphi^{-1} = \psi$. Moreover, observe that $\varphi(C) \subset D$ and $\psi(D) \subset C$, hence $\varphi(C) = D$. Thus, we have the result. \square

The norm of a normed linear space E is the Minkowski functional for the unit closed ball \mathbf{B}_E of E . Since $\text{bd } \mathbf{B}_E$ is the unit sphere \mathbf{S}_E of E , we have the following:

Corollary 3.4.14. *Let $E = (E, \|\cdot\|)$ be a normed linear space. For every bounded closed convex set $C \subset E$ with $\text{int } C \neq \emptyset$, the pair $(C, \text{bd } C)$ is homeomorphic to the pair $(\mathbf{B}_E, \mathbf{S}_E)$ of the unit closed ball and the unit sphere of E . \square*

It is easy to see that every normed linear space $E = (E, \|\cdot\|)$ is homeomorphic to the unit open ball $\mathbf{B}(\mathbf{0}, 1) = \mathbf{B}_E \setminus \mathbf{S}_E$ of E .

In fact, the following are homeomorphisms (each of them is the inverse of the other):

$$E \ni x \mapsto \frac{1}{1 + \|x\|}x \in \mathbf{B}(\mathbf{0}, 1); \quad \mathbf{B}(\mathbf{0}, 1) \ni y \mapsto \frac{1}{1 - \|y\|}y \in E.$$

By applying the Minkowski functional, this can be extended as follows:

Theorem 3.4.15. *Every open convex set V in a topological linear space E is homeomorphic to E itself.*

Proof. Without loss of generality, it can be assumed that $\mathbf{0} \in \text{int } V = V$. Then, we have $V = \text{int } V = p_V^{-1}([0, 1))$ by Corollary 3.4.12. Using the Minkowski functional p_V , we can define maps $\varphi : V \rightarrow E$ and $\psi : E \rightarrow V$ as follows:

$$\varphi(x) = \frac{1}{1 - p_V(x)}x \quad \text{for } x \in V; \quad \psi(y) = \frac{1}{1 + p_V(y)}y \quad \text{for } y \in E.$$

Observe that $\psi\varphi = \text{id}_V$ and $\varphi\psi = \text{id}_E$. This means that φ is a homeomorphism with $\psi = \varphi^{-1}$. \square

3.5 Finite-Dimensionality

Here, we prove that every finite-dimensional linear space has the unique topology that is compatible with the algebraic operations, and that a topological linear space is finite-dimensional if and only if it is locally compact.

First, we show the following proposition:

Proposition 3.5.1. *Every finite-dimensional flat F in an arbitrary linear space E has the unique (Hausdorff) topology such that the following operation is continuous:*

$$F \times F \times \mathbb{R} \ni (x, y, t) \mapsto (1-t)x + ty \in F.$$

With respect to this topology, every affine bijection $f : \mathbb{R}^n \rightarrow F$ is a homeomorphism, where $n = \dim F$. Then, F is affinely homeomorphic to \mathbb{R}^n . Moreover, if E is a topological linear space then F is closed in E .

Proof. As mentioned at the beginning of Sect. 3.4, E has a topology that makes E a topological linear space. With respect to the topology of F inherited from this topology, the above operation is continuous.

Note that there exists an affine bijection $f : \mathbb{R}^n \rightarrow F$, where $\dim F = n$. We shall show that any affine bijection $f : \mathbb{R}^n \rightarrow F$ is a homeomorphism with respect to any other topology of F such that the above operation is continuous, which implies that such a topology is unique and F is affinely homeomorphic to \mathbb{R}^n .

Since f is affine, we have

$$f(z) = \left(1 - \sum_{i=1}^n z(i)\right) f(\mathbf{0}) + \sum_{i=1}^n z(i) f(\mathbf{e}_i) \quad \text{for each } z \in \mathbb{R}^n.$$

Note that the following function is continuous:

$$\mathbb{R}^n \ni z \mapsto \left(1 - \sum_{i=1}^n z(i), z(1), \dots, z(n)\right) \in \text{fl } \Delta^n \subset \mathbb{R}^{n+1}.$$

Then, the continuity of f follows from the claim:

Claim. Given $v_1, \dots, v_k \in F$, $k \leq n$, the following function is continuous:

$$\varphi_k : \text{fl } \Delta^{k-1} \ni z \mapsto \sum_{i=1}^k z(i) v_i \in F.$$

Since $\text{fl } \Delta^0 = \Delta^0$ is a singleton, the continuity of φ_1 is obvious. Assuming the continuity of φ_k , we shall show the continuity of φ_{k+1} . Let $\psi : \text{fl } \Delta^{k-1} \times \mathbb{R} \rightarrow \text{fl } \Delta^k$ be the map defined by $\psi(z, t) = ((1-t)z, t)$. Observe that

$$\varphi_{k+1} \psi(z, t) = (1-t) \sum_{i=1}^k z(i) v_i + t v_{k+1} = (1-t) \varphi_k(z) + t v_{k+1}.$$

From the property of the topology of F and the continuity of φ_k , it follows that $\varphi_{k+1} \psi$ is continuous. For each $i = 1, \dots, k+1$, let $p_i = \text{pr}_i | \text{fl } \Delta^k : \text{fl } \Delta^k \rightarrow \mathbb{R}$ be the restriction of the projection onto the i -th factor. Note that

$$\psi|_{\text{fl } \Delta^{k-1} \times (\mathbb{R} \setminus \{1\})} : \text{fl } \Delta^{k-1} \times (\mathbb{R} \setminus \{1\}) \rightarrow \text{fl } \Delta^k \setminus p_{k+1}^{-1}(1)$$

is a homeomorphism. Hence, $\varphi_{k+1}|_{\text{fl } \Delta^k \setminus p_{k+1}^{-1}(1)}$ is continuous. Replacing the $(k+1)$ -th coordinates with the i -th coordinates, we can see the continuity of $\varphi_{k+1}|_{\text{fl } \Delta^k \setminus p_i^{-1}(1)}$. Since $\text{fl } \Delta^k = \bigcup_{i=1}^{k+1} (\text{fl } \Delta^k \setminus p_i^{-1}(1))$, it follows that φ_{k+1} is continuous. Thus, the claim can be obtained by induction.

It remains to show the openness of f . On the contrary, assume that f is not open. Then, we have $x \in \mathbb{R}^n$ and $\varepsilon > 0$ such that $f(\text{B}(x, \varepsilon))$ is not a neighborhood of $f(x)$ in F . Since $\text{bd } \text{B}(x, \varepsilon)$ is a bounded closed set of \mathbb{R}^n , it is compact, hence $f(\text{bd } \text{B}(x, \varepsilon))$ is closed in F . Then, $F \setminus f(\text{bd } \text{B}(x, \varepsilon))$ is a neighborhood of $f(x)$ in F . Using the compactness of \mathbf{I} , we can find an open neighborhood U of $f(x)$ in F such that

$$(1-t)f(x) + tU \subset F \setminus f(\text{bd } \text{B}(x, \varepsilon)) \text{ for every } t \in \mathbf{I}.$$

Then, $U \cap f(\text{bd } \text{B}(x, \varepsilon)) = \emptyset$. Since $f(\text{B}(x, \varepsilon))$ is not a neighborhood of $f(x)$, it follows that $U \not\subset f(\text{B}(x, \varepsilon))$, and so we can take a point $y \in U \setminus f(\overline{\text{B}}(x, \varepsilon))$. Now, we define a linear path $g : \mathbf{I} \rightarrow \mathbb{R}^n$ by $g(t) = (1-t)x + tf^{-1}(y)$. Since f is affine and $y \in U$, it follows that

$$fg(t) = (1-t)f(x) + ty \in F \setminus f(\text{bd } \text{B}(x, \varepsilon)) \text{ for every } t \in \mathbf{I}.$$

Since f is a bijection, we have

$$g(\mathbf{I}) \subset \mathbb{R}^n \setminus \text{bd } \text{B}(x, \varepsilon) = \text{B}(x, \varepsilon) \cup (\mathbb{R}^n \setminus \overline{\text{B}}(x, \varepsilon)).$$

Then, $g(\mathbf{0}) = x \in \text{B}(x, \varepsilon)$ and $g(1) = f^{-1}(y) \in \mathbb{R}^n \setminus \overline{\text{B}}(x, \varepsilon)$, which contradicts the connectedness of \mathbf{I} . Thus, f is open.

In the case when E is a topological linear space, to prove that F is closed in E , take a point $x \in E \setminus F$ and consider the flat $F_x = \text{fl}(F \cup \{x\})$. It is easy to construct an affine bijection $f : \mathbb{R}^{n+1} \rightarrow F_x$ such that $f(\mathbb{R}^n \times \{\mathbf{0}\}) = F$. As we saw in the above, f is a homeomorphism, hence F is closed in F_x . Since $F_x \setminus F$ is open in F_x , we have an open set U in E such that $U \cap F_x = F_x \setminus F$. Then, U is a neighborhood of x in E and $U \subset E \setminus F$. Therefore, $E \setminus F$ is open in E , that is, F is closed in E . \square

If a linear space E has a topology such that the operation

$$E \times E \times \mathbb{R} \ni (x, y, t) \mapsto (1-t)x + ty \in E$$

is continuous, then scalar multiplication and addition are also continuous with this topology because they can be written as follows:

$$E \times \mathbb{R} \ni (x, t) \mapsto tx = (1-t)\mathbf{0} + tx \in E;$$

$$E \times E \ni (x, y) \mapsto x + y = 2\left(\frac{1}{2}x + \frac{1}{2}y\right) \in E.$$

Then, the following is obtained by Proposition 3.5.1:

Corollary 3.5.2. *Every finite-dimensional linear space E has the unique (Hausdorff) topology compatible with the algebraic operations (addition and scalar multiplication), and then it is linearly homeomorphic to \mathbb{R}^n , where $n = \dim E$. \square*

Moreover, we have the following:

Corollary 3.5.3. *Let E be a topological linear space and F a finite-dimensional flat in another topological linear space. Then, every affine function $f : F \rightarrow E$ is continuous, and if f is injective then f is a closed embedding.*

Proof. By Proposition 3.5.1, F can be replaced with \mathbb{R}^n , where $n = \dim F$. Then, we can write

$$f(x) = \left(1 - \sum_{i=1}^n x(i)\right)f(\mathbf{0}) + \sum_{i=1}^n x(i)f(\mathbf{e}_i) \text{ for each } x \in \mathbb{R}^n,$$

where $\mathbf{e}_1, \dots, \mathbf{e}_n$ is the canonical orthonormal basis for \mathbb{R}^n . Thus, the continuity of f is obvious. Since $f(\mathbb{R}^n)$ is a finite-dimensional flat in E , $f(\mathbb{R}^n)$ is closed in E by Proposition 3.5.1. If f is injective then $f : \mathbb{R}^n \rightarrow f(\mathbb{R}^n)$ is an affine bijection, which is a homeomorphism by Proposition 3.5.1. Hence, f is a closed embedding. \square

Combining Proposition 3.2.2 and Corollary 3.5.3, we have

Corollary 3.5.4. *Let E be a topological linear space and C a finite-dimensional convex set in another topological linear space. Then, every affine function $f : C \rightarrow E$ is continuous. Moreover, if f is injective then f is an embedding. \square*

For finite-dimensional convex sets in a linear space, we have the following:

Proposition 3.5.5. *Let C be a finite-dimensional convex set in an arbitrary linear space E . Then, $\text{rint } C = \text{int}_{\text{fl } C} C$ with respect to the unique topology for $\text{fl } C$ as in Proposition 3.5.1.*

Proof. We may assume that E is a topological linear space. By Proposition 3.4.7(3), it suffices to show that $\text{int}_{\text{fl } C} C \neq \emptyset$. We have affinely independent $v_0, v_1, \dots, v_n \in C$ with $\text{fl } C = \text{fl}\{v_0, v_1, \dots, v_n\}$, where $n = \dim C$. We have an affine bijection $f : \mathbb{R}^n \rightarrow \text{fl } C$ such that $f(\mathbf{0}) = v_0, f(\mathbf{e}_1) = v_1, \dots, f(\mathbf{e}_n) = v_n$. Then, f is a homeomorphism by Proposition 3.5.1, hence

$$\text{int}_{\text{fl } C} C \supset \text{int}_{\text{fl } C} \langle v_0, v_1, \dots, v_n \rangle = f(\text{int}_{\mathbb{R}^n} \langle \mathbf{0}, \mathbf{e}_1, \dots, \mathbf{e}_n \rangle) \neq \emptyset. \quad \square$$

Note that every compact set in a topological linear space is topologically bounded and closed. For an n -dimensional convex set C in a linear space, the flat hull $\text{fl } C$ is affinely isomorphic to \mathbb{R}^n . Combining Propositions 3.5.1 and 3.5.5 with Corollary 3.4.14, we have the following:

Corollary 3.5.6. *For every n -dimensional compact convex set C in an arbitrary topological linear space E , the pair $(C, \partial C)$ is homeomorphic to the pair $(\mathbf{B}^n, \mathbf{S}^{n-1})$ of the unit closed n -ball and the unit $(n - 1)$ -sphere. \square*

Remark 11. It should be noted that every bounded closed set in Euclidean space \mathbb{R}^n is compact. More generally, we can prove the following:

Proposition 3.5.7. *Let E be an arbitrary topological linear space and $A \subset E$ with $\dim \text{fl } A < \infty$. Then, A is compact if and only if A is topologically bounded and closed in E .*

Sketch of Proof. Using Proposition 3.5.1, this can be reduced to the case of \mathbb{R}^n .

The following convex version of Proposition 3.5.1 is not trivial.

Proposition 3.5.8. *Let C be an n -dimensional convex set in an arbitrary linear space E . If (1) C is the convex hull of a finite set¹⁴ or (2) $C = \text{rint } C$, then C has the unique (Hausdorff) topology such that the following operation is continuous:*

$$C \times C \times \mathbf{I} \ni (x, y, t) \mapsto (1 - t)x + ty \in C.$$

In case (1), $\text{rcl } C = C$ and $(C, \partial C) \approx (\mathbf{B}^n, \mathbf{S}^{n-1})$; in case (2), $C \approx \mathbb{R}^n$.

Proof. Like Proposition 3.5.1, it suffices to see the uniqueness and the additional statement. To this end, suppose that C has such a topology, but it is unknown whether this is induced from a topology of $\text{fl } C$ or not.

Case (1): Let $C = \langle v_1, \dots, v_k \rangle$ and define $f : \Delta^{k-1} \rightarrow C$ by $f(z) = \sum_{i=1}^k z(i)v_i$. In the same way as for the claim in the proof of Proposition 3.5.1, we can see that the continuity of the operation above induces the continuity of f . Since Δ^{k-1} is compact, f is a closed map, hence it is quotient. Thus, the topology of C is unique and C is compact with respect to this topology. Giving any topology on E so that E is a topological linear space, we have $\text{rcl } C = C$ by Proposition 3.4.7(i) and $(C, \partial C) \approx (\mathbf{B}^n, \mathbf{S}^{n-1})$ by Corollary 3.5.6.

Case (2): Let $f : \mathbb{R}^n \rightarrow \text{fl } C$ be an affine bijection, where $n = \dim \text{fl } C = \dim C$. Since $D = f^{-1}(C)$ is an n -dimensional convex set in \mathbb{R}^n , $D = \text{rint } D = \text{int } D$ is open in \mathbb{R}^n by Proposition 3.5.5, hence $D \approx \mathbb{R}^n$ by Proposition 3.4.15. Then, it suffices to show that $f|_D : D \rightarrow C$ is a homeomorphism. For each $x \in D$, choose $\delta > 0$ so that $x + \delta \mathbf{B}^n = \overline{\mathbf{B}}(x, \delta) \subset D$. Let $v_0 = x - \delta \hat{\Delta}^{n-1}$, where $\hat{\Delta}^{n-1}$ is the barycenter of the standard $(n - 1)$ -simplex $\Delta^{n-1} = \langle \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n \rangle \subset \mathbb{R}^n$. For each $i = 1, \dots, n$, let $v_i = x + \delta \mathbf{e}_i$. Then, v_0, v_1, \dots, v_n are affinely independent and

$$x \in \text{int}_{\mathbb{R}^n} \langle v_0, v_1, \dots, v_n \rangle \subset x + \delta \mathbf{B}^n \subset D,$$

hence $\langle v_0, v_1, \dots, v_n \rangle$ is a neighborhood of x in D . On the other hand, we have the affine homeomorphism $\varphi : \Delta^n \rightarrow \langle v_0, v_1, \dots, v_n \rangle$ defined by $\varphi(z) = \sum_{i=0}^n z(i) +$

¹⁴In this case, C is called a **cell** or a **(convex) linear cell** (cf. Sect. 4.1).

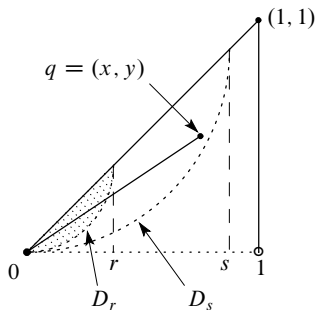


Fig. 3.7 The continuity of the operation at $(\mathbf{0}, q, 0) \in C \times C \times \mathbf{I}$

$1)v_i$. Since $f\varphi(z) = \sum_{i=0}^n z(i+1)f(v_i)$, the continuity of the operation above implies that of $f\varphi$, hence $f|\langle v_0, v_1, \dots, v_n \rangle$ is continuous at x . Then, it follows that $f|D$ is continuous at x .

Since D is open in \mathbb{R}^n , we can apply the same argument as in the proof of Proposition 3.5.1 to prove that $f|D : D \rightarrow C$ is open. Consequently, $f|D : D \rightarrow C$ is a homeomorphism. \square

Remark 12. For an arbitrary finite-dimensional convex C , Proposition 3.5.8 does not hold in general. For example, let

$$C = \{\mathbf{0}\} \cup \{(x, y) \in (0, 1]^2 \mid x \geq y\} \subset \mathbb{R}^2.$$

Then, C is a convex set that has a finer topology than usual such that the operation in Proposition 3.5.8 is continuous. Such a topology is generated by open sets in the usual topology and the following sets:

$$D_r = \{\mathbf{0}\} \cup (\mathbf{B}((0, r), r) \cap C), \quad r > 0.$$

Note that this topology induces the same relative topology on $C \setminus \{\mathbf{0}\}$ as usual. Since $D_{\varepsilon/\sqrt{2}} \subset \mathbf{B}(\mathbf{0}, \varepsilon)$ for each $\varepsilon > 0$, $\{D_r \mid r > 0\}$ is a neighborhood basis at $\mathbf{0} \in C$ with respect to this topology.

We shall show that the operation

$$C \times C \times \mathbf{I} \ni (p, q, t) \mapsto (1-t)p + tq \in C$$

is continuous at $(p, q, t) \in C \times C \times \mathbf{I}$. If $(1-t)p + tq \neq \mathbf{0}$, it follows from the continuity with respect to the usual topology. The continuity at $(0, 0, t)$ follows from the convexity of $D_r, r > 0$.

To see the continuity at $(\mathbf{0}, q, 0)$ ($q \neq \mathbf{0}$), let $q = (x, y)$, where $0 < y \leq x \leq 1$. Choose $s > 0$ so that $q \in D_s$ (i.e., $s > (x^2 + y^2)/2y$). For each $0 < r < \min\{1, s\}$, let $0 \leq t \leq r/2s$, $p' \in D_{r/2}$, and $q' \in D_s$ (Fig. 3.7). Observe that

$$\frac{1-t}{r/s-t}(r/s)p' \in \frac{1-t}{r/s-t}(r/s)D_{r/2} \subset \frac{1}{r/2s}(r/s)D_{r/2} = D_r$$

and $(r/s)q' \in (r/s)D_s = D_r$. Since D_r is convex, it follows that

$$(1-t)p' + tq' = \left(1 - \frac{t}{r/s}\right) \frac{1-t}{r/s-t} (r/s)p' + \frac{t}{r/s} (r/s)q' \in D_r.$$

Thus, the operation is continuous at $(0, q, 0)$. The continuity at $(p, \mathbf{0}, 1)$ ($p \neq \mathbf{0}$) is the same.

A subset A of a topological linear space E is **totally bounded** provided, for each neighborhood U of $\mathbf{0} \in E$, there exists some finite set $M \subset E$ such that $A \subset M + U$. In this definition, M can be taken as a subset of A .

Indeed, for each neighborhood U of $\mathbf{0} \in E$, we have a circled neighborhood V such that $V + V \subset U$. Then, $A \subset M + V$ for some finite set $M \subset E$, where it can be assumed that $(x + V) \cap A \neq \emptyset$ for every $x \in M$. For each $x \in M$, choose $a_x \in A$ so that $a_x \in x + V$. Then, $x \in a_x - V = a_x + V$. It follows that $A \subset \bigcup_{x \in M} (x + V) \subset \bigcup_{x \in M} (a_x + V + V) \subset \bigcup_{x \in M} (a_x + U)$.

If $A \subset E$ is totally bounded and $B \subset A$, then B is also totally bounded. It is easy to see that every compact subset of E is totally bounded and every totally bounded subset of E is topologically bounded. In other words, we have:

$$\boxed{\text{compact}} \Rightarrow \boxed{\text{totally bounded}} \Rightarrow \boxed{\text{topologically bounded}}$$

For topological linear spaces, the finite-dimensionality can be simply characterized as follows:

Theorem 3.5.9. *Let E be a topological linear space. The following are equivalent:*

- (a) E is finite-dimensional;
- (b) E is locally compact;
- (c) $\mathbf{0} \in E$ has a totally bounded neighborhood in E .

Proof. Since each n -dimensional topological linear space is linearly homeomorphic to \mathbb{R}^n (Corollary 3.5.2), we have (a) \Rightarrow (b). Since every compact subset of E is totally bounded, the implication (b) \Rightarrow (c) follows.

(c) \Rightarrow (a): Let U be a totally bounded neighborhood of $\mathbf{0} \in E$. By Proposition 3.4.1, we have a circled neighborhood V of $\mathbf{0}$ such that $V + V \subset U$. Then, V is also totally bounded. First, we show the following:

Claim. For each closed linear subspace $F \subsetneq E$, there is some $x \in U$ such that $(x + V) \cap F = \emptyset$.

Contrary to the claim, suppose that $(x + V) \cap F \neq \emptyset$ for every $x \in U$. Since $V = -V$, it follows that $U \subset F + V$, so we have $V + V \subset F + V$. If $(n-1)V \subset F + V$ then

$$nV \subset (n-1)V + V \subset F + V + V \subset F + F + V = F + V.$$

By induction, we have $nV \subset F + V$ for every $n \in \mathbb{N}$, which implies that $V \subset \bigcap_{n \in \mathbb{N}} (F + n^{-1}V)$.

Take $z \in E \setminus F$. Since F is closed in E , we have a circled neighborhood W of $\mathbf{0} \in E$ such that $W \subset V$ and $(z + W) \cap F = \emptyset$. The total boundedness of V implies the topological boundedness, hence $V \subset mW$ for some $m \in \mathbb{N}$. On the other hand, $k^{-1}z \in V$ for some $k \in \mathbb{N}$. Since $k^{-1}z \in V \subset F + (km)^{-1}V$, it follows that $z \in F + m^{-1}V \subset F + W$. This contradicts the fact that $(z + W) \cap F = \emptyset$.

Now, assume that E is infinite-dimensional. Let $v_1 \in U \setminus \{\mathbf{0}\}$ and $F_1 = \mathbb{R}v_1$. Then, F_1 is closed in E (Proposition 3.5.1) and $F_1 \neq E$. Applying the claim above, we have $v_2 \in U$ such that $(v_2 + V) \cap F_1 = \emptyset$. Note that $v_2 \notin v_1 + V$. Let $F_2 = \mathbb{R}v_1 + \mathbb{R}v_2$. Since F_2 is closed in E (Proposition 3.5.1) and $F_2 \neq E$, we can again apply the claim to find $v_3 \in U$ such that $(v_3 + V) \cap F_2 = \emptyset$. Then, note that $v_3 \notin v_i + V$ for $i = 1, 2$. By induction, we have $v_n \in U$, $n \in \mathbb{N}$, such that $v_n \notin v_i + V$ for $i < n$. Then, $\{v_n \mid n \in \mathbb{N}\}$ is not totally bounded. This is a contradiction. Consequently, E is finite-dimensional. \square

By Theorem 3.5.9, every infinite-dimensional topological linear space is not locally compact.

3.6 Metrizable and Normability

In this section, we prove metrization and normability theorems for topological linear spaces. The metrizable of a topological linear space has the following very simple characterization:

Theorem 3.6.1. *A topological linear space E is metrizable if and only if $\mathbf{0} \in E$ has a countable neighborhood basis.*

In a more general setting, we shall prove a stronger result. A metric d on a group G is said to be **left** (resp. **right**) **invariant** if $d(x, y) = d(zx, zy)$ (resp. $d(x, y) = d(xz, yz)$) for each $x, y, z \in G$; equivalently, $d(x, y) = d(x^{-1}y, \mathbf{1})$ (resp. $d(x, y) = d(xy^{-1}, \mathbf{1})$) for each $x, y \in G$. When both of two metrics d and d' on a group G are left (or right) invariant, they are uniformly equivalent to each other if and only if they induce the same topology. It is said that d is **invariant** if it is left and right invariant. Every invariant metric d on a group G induces the topology on G that makes G a topological group. In fact,

$$d(x, y) = d(x^{-1}xy^{-1}, x^{-1}yy^{-1}) = d(y^{-1}, x^{-1}) = d(x^{-1}, y^{-1}) \quad \text{and} \\ d(xy, x'y') \leq d(xy, x'y) + d(x'y, x'y') = d(x, x') + d(y, y').$$

It is easy to verify that a left (or right) invariant metric d on a group G is invariant if $d(x, y) = d(x^{-1}, y^{-1})$ for each $x, y \in G$. Theorem 3.6.1 comes from the following:

Theorem 3.6.2. *For a topological group G , the following are equivalent:*

- (a) G is metrizable;
- (b) The unit $\mathbf{1} \in G$ has a countable neighborhood basis;
- (c) G has an admissible bounded left invariant (right invariant) metric.

Proof. Since the implications (a) \Rightarrow (b) and (c) \Rightarrow (a) are obvious, it suffices to show the implication (b) \Rightarrow (c).

(b) \Rightarrow (c):¹⁵ We shall construct a left invariant metric $\rho \in \text{Metr}(G)$. Then, a right invariant metric $\rho' \in \text{Metr}(G)$ can be defined by $\rho'(x, y) = \rho(x^{-1}, y^{-1})$. By condition (b), we can find an open neighborhood basis $\{V_n \mid n \in \mathbb{N}\}$ at $\mathbf{1} \in G$ such that

$$V_n^{-1} = V_n \text{ and } V_{n+1}V_{n+1}V_{n+1} \subset V_n \text{ for each } n \in \mathbb{N}.^{16}$$

Let $V_0 = G$, and define

$$p(x) = \inf\{2^{-i} \mid x \in V_i\} \in \mathbf{I} \text{ for each } x \in G.$$

Since $V_n = V_n^{-1}$ for each $n \in \mathbb{N}$, it follows that $p(x) = p(x^{-1})$ for every $x \in G$. Note that $\bigcap_{n \in \omega} V_n = \{\mathbf{1}\}$.¹⁷ Hence, for every $x \in G$,

$$p(x) = 0 \Leftrightarrow x = \mathbf{1}.$$

By induction on n , we shall prove the following:

$$(*) \quad p(x_0^{-1}x_n) \leq 2 \sum_{i=1}^n p(x_{i-1}^{-1}x_i) \text{ for each } x_0, x_1, \dots, x_n \in G.^{18}$$

The case $n = 1$ is obvious. Assume $(*)$ for $m < n$. If $\sum_{i=1}^n p(x_{i-1}^{-1}x_i) = 0$ or $\sum_{i=1}^n p(x_{i-1}^{-1}x_i) \geq \frac{1}{2}$, it is trivial. When $2^{-k-1} \leq \sum_{i=1}^n p(x_{i-1}^{-1}x_i) < 2^{-k}$ for some $k \in \mathbb{N}$, choose $1 \leq m \leq n$ so that

$$\sum_{i=1}^{m-1} p(x_{i-1}^{-1}x_i) < 2^{-k-1} \quad \text{and} \quad \sum_{i=m+1}^n p(x_{i-1}^{-1}x_i) < 2^{-k-1}.$$

¹⁵The idea of the proof is the same as that of Theorem 2.4.1 (b) \Rightarrow (a).

¹⁶Note that $\{V_n x \mid n \in \mathbb{N}\}$ is an open neighborhood basis at $x \in G$. For each $x, y \in G$ and $n \in \mathbb{N}$, $V_{n+1}x \cap V_{n+1}y \neq \emptyset$ implies $V_{n+1}y \subset V_n x$. Indeed, $ux = vy$ for some $u, v \in V_{n+1}$, hence $V_{n+1}y = V_{n+1}v^{-1}ux \subset V_n x$. Thus, the metrizable of G can be obtained by the Frink Metrization Theorem 2.4.1. On the other hand, $\mathcal{V}_n = \{V_n x \mid x \in G\} \in \text{cov}(G)$ and $\text{st } \mathcal{V}_{n+1} < \mathcal{V}_n$. Indeed, $\text{st}(V_{n+1}x, \mathcal{V}_{n+1}) \subset V_n x$. Thus, the metrizable of G can also be obtained by Corollary 2.4.4.

¹⁷It is assumed that G is Hausdorff.

¹⁸For each $x, y \in G$, let $\delta(x, y) = p(x^{-1}y)$. Then, this inequality is simply the one given in the sketch of the direct proof for Corollary 2.4.4.

Note that $p(x_{m-1}^{-1}x_m) < 2^{-k}$. By the inductive assumption, $p(x_0^{-1}x_{m-1}) < 2^{-k}$ and $p(x_m^{-1}x_n) < 2^{-k}$. Then, $x_0^{-1}x_{m-1}, x_{m-1}^{-1}x_m, x_m^{-1}x_n \in V_{k+1}$. Since $V_{k+1}V_{k+1}V_{k+1} \subset V_k$, it follows that $x_0^{-1}x_n \in V_k$, hence

$$p(x_0^{-1}x_n) \leq 2^{-k} \leq 2 \sum_{i=1}^n p(x_{i-1}^{-1}x_i).$$

Now, we can define a metric ρ on G as follows:

$$\rho(x, y) = \inf \left\{ \sum_{i=1}^n p(x_{i-1}^{-1}x_i) \mid n \in \mathbb{N}, x_i \in G, x_0 = x, x_n = y \right\}.$$

By the definition, ρ is left invariant. Note that $\rho(x, y) \leq p(x^{-1}y) \leq 1$. Then, $x^{-1}y \in V_n$ implies $\rho(x, y) \leq p(x^{-1}y) \leq 2^{-n} < 2^{-n+1}$, which means $xV_n \subset B_\rho(x, 2^{-n+1})$. On the other hand, if $\rho(x, y) < 2^{-n}$ then $p(x^{-1}y) \leq 2\rho(x, y) < 2^{-n+1}$ by (*), which implies $x^{-1}y \in V_n$. Thus, $B_\rho(x, 2^{-n}) \subset xV_n$. Therefore, ρ is admissible. \square

In the above proof, a right invariant metric $\rho \in \text{Metr}(G)$ can be directly defined as follows:

$$\rho(x, y) = \inf \left\{ \sum_{i=1}^n p(x_{i-1}x_i^{-1}) \mid n \in \mathbb{N}, x_i \in G, x_0 = x, x_n = y \right\}.$$

Every metrizable topological linear space E has an admissible (bounded) metric ρ that is not only invariant but also satisfies the following:

$$(\#) \quad |t| \leq 1 \Rightarrow \rho(tx, \mathbf{0}) \leq \rho(x, \mathbf{0}).$$

To verify this, let us recall how to define the metric ρ in the above proof. Taking a neighborhood basis $\{V_n \mid n \in \mathbb{N}\}$ at $\mathbf{0} \in E$ so that $V_n = -V_n$ and $V_{n+1} + V_{n+1} + V_{n+1} \subset V_n$ for each $n \in \mathbb{N}$, we define the admissible invariant metric ρ as follows:

$$\rho(x, y) = \inf \left\{ \sum_{i=1}^n p(x_i - x_{i-1}) \mid n \in \mathbb{N}, x_i \in E, x_0 = x, x_n = y \right\},$$

where $p(x) = \inf\{2^{-i} \mid x \in V_i\}$. Since E is a topological linear space, the condition that $V_n = -V_n$ can be replaced by a stronger condition that V_n is circled, i.e., $tV_n \subset V_n$ for $t \in [-1, 1]$. Then, $p(tx) \leq p(x)$ for each $x \in E$ and $t \in [-1, 1]$, which implies that $\rho(tx, \mathbf{0}) \leq \rho(x, \mathbf{0})$ for each $x \in E$ and $t \in [-1, 1]$.

Let d be an invariant metric on a linear space E . Addition on a linear space E is clearly continuous with respect to d . On the other hand, scalar multiplication on E is continuous with respect to d if and only if d satisfies the following three conditions:

- (i) $d(x_n, \mathbf{0}) \rightarrow 0 \Rightarrow \forall t \in \mathbb{R}, d(tx_n, \mathbf{0}) \rightarrow 0$;
- (ii) $t_n \rightarrow 0 \Rightarrow \forall x \in E, d(t_n x, \mathbf{0}) \rightarrow 0$;
- (iii) $d(x_n, \mathbf{0}) \rightarrow 0, t_n \rightarrow 0 \Rightarrow d(t_n x_n, \mathbf{0}) \rightarrow 0$.

Indeed, the “only if” part is trivial. To show the “if” part, observe

$$t_n x_n - tx = (t_n - t)(x_n - x) + t(x_n - x) + (t_n - t)x.$$

Since d is invariant, it follows that

$$\begin{aligned} d(t_n x_n, tx) &= d((t_n - t)(x_n - x) + t(x_n - x) + (t_n - t)x, \mathbf{0}) \\ &\leq d((t_n - t)(x_n - x), \mathbf{0}) + d(t(x_n - x), \mathbf{0}) + d((t_n - t)x, \mathbf{0}), \end{aligned}$$

where $d(t_n x_n, tx) \rightarrow 0$ if $t_n \rightarrow t$ and $d(x_n, x) \rightarrow 0$. Thus, the above three conditions imply the continuity of scalar multiplication on E with respect to d .

It should be remarked that condition (‡) implies condition (iii).

An invariant metric d on E satisfying these conditions is called a **linear metric**. A linear space with a linear metric is called a **metric linear space**. Then, every metric linear space is a metrizable topological linear space. Conversely, we have the following fact:

Fact. *Every admissible invariant metric for a metrizable topological linear space is a linear metric.*

For subsets of a metric linear space, the total boundedness coincides with that in the metric sense. On the other hand, the topological boundedness does not coincide with the metric boundedness. In fact, every metrizable topological linear space E has an admissible bounded invariant metric. For instance, given an admissible invariant metric d for E , the following are admissible bounded invariant metrics:

$$\min \{1, d(x, y)\}, \quad \frac{d(x, y)}{1 + d(x, y)}.$$

For a linear metric ρ on E with the condition (‡), the functional $E \ni x \mapsto \rho(x, \mathbf{0}) \in \mathbb{R}$ is called an **F -norm**. In other words, a functional $\|\cdot\| : E \rightarrow \mathbb{R}$ on a linear space E is called an **F -norm** if it satisfies the following conditions:

- (F_1) $\|x\| \geq 0$ for every $x \in E$;
- (F_2) $\|x\| = 0 \Rightarrow x = \mathbf{0}$;
- (F_3) $|t| \leq 1 \Rightarrow \|tx\| \leq \|x\|$ for every $x \in E$;
- (F_4) $\|x + y\| \leq \|x\| + \|y\|$ for every $x, y \in E$;
- (F_5) $\|x_n\| \rightarrow 0 \Rightarrow \|tx_n\| \rightarrow 0$ for every $t \in \mathbb{R}$;
- (F_6) $t_n \rightarrow 0 \Rightarrow \|t_n x\| \rightarrow 0$ for every $x \in E$.

Conditions (F_3), (F_5), and (F_6) correspond to conditions (‡), (i), and (ii), respectively. The converse of (F_2) is true because $\|\mathbf{0}\| = 0$ by (F_6). Then, $\|x\| = 0$ if and only if $x = \mathbf{0}$. Condition (F_3) implies that $\|-x\| = \|x\|$ for every $x \in E$. Furthermore, conditions (F_3) and (F_4) imply condition (F_5). Indeed, using (F_4) inductively, we have $\|nx\| \leq n\|x\|$ for every $n \in \mathbb{N}$. Each $t \in [0, \infty)$ can be written as $t = [t] + s$ for some $s \in [0, 1)$, where $[t]$ is the greatest integer $\leq t$. Since $\|sx\| \leq \|x\|$ by (F_3), it follows that $\|tx\| \leq ([t] + 1)\|x\|$. Because $\|-x\| = \|x\|$, $\|tx\| \leq ([|t|] + 1)\|x\|$ for every $t \in \mathbb{R}$. This implies condition (F_5). Thus, condition (F_5) is unnecessary.

A linear space E given an F -norm $\|\cdot\|$ is called an **F -normed linear space**. Every norm is an F -norm, hence every normed linear space is an F -normed space. An F -norm $\|\cdot\|$ induces the linear metric $d(x, y) = \|x - y\|$. Then, every F -normed linear space is a metric linear space. An F -norm on a topological linear

space E is said to be **admissible** if it induces the topology for E . As we saw above, if E is metrizable, then E has an admissible invariant metric ρ satisfying $(\#)$, which induces the F -norm. Therefore, we have the following:

Theorem 3.6.3. *A topological linear space has an admissible F -norm if and only if it is metrizable. \square*

For each metrizable topological linear space, there exists an F -norm with the following stronger condition than (F_3) :

$$(F_3^*) \quad x \neq \mathbf{0}, |t| < 1 \Rightarrow \|tx\| < \|x\|,$$

which implies that $\|sx\| < \|tx\|$ for each $x \neq \mathbf{0}$ and $0 < s < t$. The following proposition guarantees the existence of an F -norm with the condition (F_3^*) :

Proposition 3.6.4. *Every (completely) metrizable topological linear space E has an admissible invariant (complete) metric d such that $d(tx, \mathbf{0}) < d(x, \mathbf{0})$ if $x \neq \mathbf{0}$ and $|t| < 1$, which induces an admissible F -norm satisfying (F_3^*) . If an admissible invariant metric ρ for E is given, d can be chosen so that $d \geq \rho$ (hence, if ρ is complete, then so is d). Moreover, if ρ is bounded, d can be chosen to be bounded.*

Proof. Given an admissible (bounded) invariant metric ρ for E , we define $d_1(x, y) = \sup_{0 < s \leq 1} \rho(sx, sy)$. Then, d_1 is an invariant metric on E with $d_1 \geq \rho$ (if ρ is bounded then so is d_1). For each $\varepsilon > 0$, since the scalar multiplication $E \times \mathbb{R} \ni (x, s) \rightarrow sx \in E$ is continuous at $(\mathbf{0}, s)$ and \mathbf{I} is compact, we can find $\delta > 0$ such that $\rho(x, \mathbf{0}) < \delta$ implies $\rho(sx, \mathbf{0}) < \varepsilon$ for every $s \in \mathbf{I}$, hence $\rho(x, y) < \delta$ implies $d_1(x, y) = \sup_{0 < s \leq 1} \rho(sx, sy) \leq \varepsilon$. Thus, d_1 is uniformly equivalent to ρ . In particular, d_1 is admissible. For $r > 0$, we define an admissible invariant metric d_r for E by $d_r(x, y) = d_1(rx, ry) (= \sup_{0 < s \leq r} \rho(sx, sy))$. Observe that $d_r(tx, \mathbf{0}) \leq d_r(x, \mathbf{0})$ for each $x \in E$ and $t \in \mathbf{I}$.

Now, let $\mathbb{Q} \cap (0, 1] = \{r_n \mid n \in \mathbb{N}\}$, where $r_1 = 1$. We define $d(x, y) = \sum_{n \in \mathbb{N}} 2^{-n+1} d_{r_n}(x, y)$. Then, d is an invariant metric on E and

$$\rho(x, y) \leq d_1(x, y) \leq d(x, y) \leq 2d_1(x, y),$$

hence d is admissible (if ρ is bounded then so is d). It also follows that $d(tx, \mathbf{0}) \leq d(x, \mathbf{0})$ for each $x \in E$ and $t \in \mathbf{I}$. It remains to show that $d(tx, \mathbf{0}) < d(x, \mathbf{0})$ for each $x \in E \setminus \{\mathbf{0}\}$ and $0 < t < 1$. Since $\mathbb{Q} \cap (0, 1)$ is dense in $(0, 1)$, it suffices to show that $d(tx, \mathbf{0}) < d(x, \mathbf{0})$ for each $x \in E \setminus \{\mathbf{0}\}$ and $t \in \mathbb{Q} \cap (0, 1)$. Assume that there exists some $x \in E \setminus \{\mathbf{0}\}$ and $t \in \mathbb{Q} \cap (0, 1)$ such that $d(tx, \mathbf{0}) = d(x, \mathbf{0})$. Note that $d_r(tx, \mathbf{0}) = d_r(x, \mathbf{0})$ for each $r \in \mathbb{Q} \cap (0, 1)$. Then, it follows that

$$\begin{aligned} d_t(x, \mathbf{0}) &= d_t(tx, \mathbf{0}) = d_{t^2}(x, \mathbf{0}) = d_{t^2}(tx, \mathbf{0}) \\ &= d_{t^3}(x, \mathbf{0}) = d_{t^3}(tx, \mathbf{0}) = \dots, \end{aligned}$$

so $d_t(x, \mathbf{0}) = d_{t^{n+1}}(x, \mathbf{0}) = d_t(t^n x, \mathbf{0})$ for every $n \in \mathbb{N}$. Since $\lim_{n \rightarrow \infty} t^n = 0$, it follows that $d_t(x, \mathbf{0}) = \lim_{n \rightarrow \infty} d_t(t^n x, \mathbf{0}) = 0$, hence $x = \mathbf{0}$, which is a contradiction. \square

The topological linear space $\mathbb{R}^{\mathbb{N}} = \mathfrak{s}$ (the space of sequences) has the following admissible F -norms:

$$\sup_{i \in \mathbb{N}} \min \{1/i, |x(i)|\}, \sum_{i \in \mathbb{N}} \min \{2^{-i}, |x(i)|\}, \sum_{i \in \mathbb{N}} \frac{2^{-i} |x(i)|}{1 + |x(i)|}, \dots$$

The first two do not satisfy condition (F_3^*) , but the third does.

We now consider the completion of metric linear spaces (cf. 2.3.10).

Proposition 3.6.5. *Let G be a topological group such that the topology is induced by an invariant metric d . The completion $\tilde{G} = (\tilde{G}, \tilde{d})$ of (G, d) is a group such that G is its subgroup and \tilde{d} is invariant. Similarly, the completion of a metric (F -normed or normed) linear space E is a metric (F -normed or normed) linear space containing E as a linear subspace.*

Proof. We define the algebraic operations on \tilde{G} as follows: for each $x, y \in \tilde{G}$, choose sequences $(x_i)_{i \in \mathbb{N}}$ and $(y_i)_{i \in \mathbb{N}}$ in G so as to converge to x and y , respectively. Since d is invariant, $(x_i y_i)_{i \in \mathbb{N}}$ and $(x_i^{-1})_{i \in \mathbb{N}}$ are Cauchy sequences in G . Then, define xy and x^{-1} as the limits of $(x_i y_i)_{i \in \mathbb{N}}$ and $(x_i^{-1})_{i \in \mathbb{N}}$, respectively. It is easily verified that these are well-defined. Since $\tilde{d}(x, y) = \lim_{i \rightarrow \infty} d(x_i, y_i)$, it is also easy to see that \tilde{d} is invariant, which implies the continuity of the algebraic operations $(x, y) \mapsto xy$ and $x \mapsto x^{-1}$.

For the completion \tilde{E} of a metric linear space E , we can define not only addition but also scalar multiplication in the same way. To see the continuity of scalar multiplication, let $x \in \tilde{E}$ and $t \in \mathbb{R}$. Choose a sequence $(x_i)_{i \in \mathbb{N}}$ in E so as to converge to x . For each $\varepsilon > 0$, we can choose $\delta_0 > 0$ (depending on t) so that

$$z \in E, d(z, \mathbf{0}) < \delta_0, |t - t'| < \delta_0 \Rightarrow d(t'z, \mathbf{0}) < \varepsilon/4.$$

Then, we have $n_0 \in \mathbb{N}$ such that $d(x_n, x_{n_0}) < \delta_0$ for every $n \geq n_0$. Choose $\delta_1 > 0$ so that $\delta_1 < \delta_0$ and

$$|s| < \delta_1 \Rightarrow d(sx_{n_0}, \mathbf{0}) < \varepsilon/4.$$

Now, let $x' \in \tilde{E}$ and $t' \in \mathbb{R}$ such that $\tilde{d}(x, x') < \delta_0$ and $|t - t'| < \delta_1$. Take a sequence $(x'_i)_{i \in \mathbb{N}}$ in E so as to converge to x' and choose $n_1 \in \mathbb{N}$ so that $n_1 \geq n_0$ and $d(x_n, x'_n) < \delta_0$ for every $n \geq n_1$. Then, for every $n \geq n_1$, it follows that

$$\begin{aligned} d(tx_n, t'x'_n) &\leq d(tx_n, tx_{n_0}) + d(tx_{n_0}, t'x_{n_0}) + d(t'x_{n_0}, t'x_n) + d(t'x_n, t'x'_n) \\ &= d(t(x_n - x_{n_0}), \mathbf{0}) + d((t - t')x_{n_0}, \mathbf{0}) \\ &\quad + d(t'(x_{n_0} - x_n), \mathbf{0}) + d(t'(x_n - x'_n), \mathbf{0}) \\ &< \varepsilon/4 + \varepsilon/4 + \varepsilon/4 + \varepsilon/4 = \varepsilon. \end{aligned}$$

When E is an F -normed (or normed) linear space, it is easy to see that the F -norm (or norm) for E naturally extends to \tilde{E} . \square

Concerning the completeness of admissible invariant metrics, we have the following:

Theorem 3.6.6. *Let G be a completely metrizable topological group. Every admissible invariant metric for G is complete. In particular, a metric linear space is complete if it is absolutely G_δ (i.e., completely metrizable).*

Proof. Let d be an admissible invariant metric for G and \tilde{G} be the completion of (G, d) . Note that \tilde{G} is a topological group by Proposition 3.6.5. It suffices to show that $\tilde{G} = G$. Since G is completely metrizable, G is a dense G_δ -set in \tilde{G} (Theorem 2.5.2), hence we can write $\tilde{G} \setminus G = \bigcup_{n \in \mathbb{N}} F_n$, where each F_n is a nowhere dense closed set in \tilde{G} . Assume $\tilde{G} \setminus G \neq \emptyset$ and take $x_0 \in \tilde{G} \setminus G$. Since $x_0 x \in \tilde{G} \setminus G$ for every $x \in G$, it follows that $G \subset \bigcup_{n \in \mathbb{N}} x_0^{-1} F_n$, where each $x_0^{-1} F_n$ is also a nowhere dense closed set in \tilde{G} . Then, we have

$$\tilde{G} = \bigcup_{n \in \mathbb{N}} F_n \cup \bigcup_{n \in \mathbb{N}} x_0^{-1} F_n,$$

which is the countable union of nowhere dense closed sets. This contradicts the complete metrizability of \tilde{G} (the Baire Category Theorem 2.5.1). \square

Corollary 3.6.7. *Let G be a metrizable topological group. Every completely metrizable Abelian subgroup H of G is closed in G . Hence, in a metrizable topological linear space, every completely metrizable linear subspace is closed.*

Proof. By Theorem 3.6.2, G has an admissible left invariant metric d . Because H is an Abelian subgroup of G , the restriction of d on H is an admissible invariant metric for H , which is complete by Theorem 3.6.6. Hence, it follows that H is closed in G . \square

It is said that an F -norm (or an F -normed space) is **complete** if the metric induced by the F -norm is complete. It should be noted that every metrizable topological linear space has an admissible F -norm (Proposition 3.6.4) and that every admissible F -norm for a completely metrizable topological linear space is complete (Theorem 3.6.6). A completely metrizable topological linear space (or a complete F -normed linear space) is called an **F -space**. A **Fréchet space** is a locally convex F -space, that is, a completely metrizable locally convex topological linear space. Every Banach space is a Fréchet space, but the converse does not hold. In fact, $\mathfrak{s} = \mathbb{R}^{\mathbb{N}}$ is a Fréchet space but it is not normable (Proposition 1.2.1).

Concerning the quotient of an F -normed (or normed) linear space, we have the following:

Proposition 3.6.8. *Let $E = (E, \|\cdot\|)$ be an F -normed (or normed) linear space and F a closed linear subspace of E . Then, the quotient space E/F has the admissible F -norm (or norm) $\|\xi\| = \inf_{x \in \xi} \|x\|$, where if $\|\cdot\|$ is complete then so is $\|\cdot\|$. Hence, if E is (completely) metrizable or (completely) normable then so is E/F .*

Proof. It is easy to see that $\|\cdot\|$ is an F -norm (or a norm). It should be noted that the closedness of F is necessary for condition (F_2) . Let $q : E \rightarrow E/F$ be the natural linear surjection, i.e., $q(x) = x + F$. Then, for each $\varepsilon > 0$,

$$\{q(x) \mid \|x\| < \varepsilon\} = \{\xi \in E/F \mid \|\xi\| < \varepsilon\},$$

which means that $q : E \rightarrow (E/F, \|\cdot\|)$ is open and continuous, so it is a quotient map. Then, $\|\cdot\|$ induces the quotient topology, i.e., $\|\cdot\|$ is admissible for the quotient topology. It also follows that if E is locally convex then so is E/F .

We should remark the following fact:

Fact. $\|\xi - \xi'\| = \inf \{\|x - x'\| \mid x' \in \xi'\}$ for each $x \in \xi$.

Indeed, the left side is not greater than the right side by definition. For each $x, y \in \xi$ and $y' \in \xi'$,

$$\|y - y'\| = \|x - (y' + x - y)\| \geq \inf \{\|x - x'\| \mid x' \in \xi'\}$$

because $y' + x - y \in \xi'$. Thus, the left side is not less than the right side.

We shall show that if $\|\cdot\|$ is complete then so is $\|\cdot\|$. To see the completeness of $\|\cdot\|$, it suffices to prove that each Cauchy sequence $(\xi_i)_{i \in \mathbb{N}}$ in E/F contains a convergent subsequence. Then, by replacing $(\xi_i)_{i \in \mathbb{N}}$ with its subsequence, we may assume that $\|\xi_i - \xi_{i+1}\| < 2^{-i}$ for each $i \in \mathbb{N}$. Using the fact above, we can inductively choose $x_i \in \xi_i$ so that $\|x_i - x_{i+1}\| < 2^{-i}$. Then, $(x_i)_{i \in \mathbb{N}}$ is a Cauchy sequence in E , which converges to some $x \in E$. It follows that $(\xi_i)_{i \in \mathbb{N}}$ converges to some $x + F$. \square

In the above, E/F is called the **quotient F -normed** (or **normed**) **linear space** with the F -norm (or norm) $\|\cdot\|$, which is called the **quotient F -norm** (or **norm**). Note that E/F is locally convex if so is E . If E is a Banach space, a Fréchet space, or an F -space, then so is E/F for any closed linear subspace F of E .

Recall that $A \subset E$ is topologically bounded if, for each neighborhood U of $\mathbf{0} \in E$, there exists some $r \in \mathbb{R}$ such that $A \subset rU$.

Theorem 3.6.9. *A topological linear space E is normable if and only if there is a topologically bounded convex neighborhood of $\mathbf{0} \in E$.*

Proof. The “only if” part is trivial. To see the “if” part, let V be a topologically bounded convex neighborhood of $\mathbf{0} \in E$. Then, $W = V \cap (-V)$ is a topologically bounded circled convex neighborhood of $\mathbf{0} \in E$. Hence, the Minkowski functional p_W is a norm on E by Proposition 3.3.4. By Corollary 3.4.12,

$$\{x \in E \mid p_W(x) < \varepsilon\} = \varepsilon p_W^{-1}([0, 1)) = \varepsilon \operatorname{int} W \quad \text{for each } \varepsilon > 0.$$

For each neighborhood U of $\mathbf{0} \in E$, we can choose $r > 0$ such that $W \subset rU$. Then,

$$\{x \in E \mid p_W(x) < r^{-1}\} = r^{-1} \operatorname{int} W \subset r^{-1} W \subset U,$$

hence p_W induces the topology for E . \square

For the local convexity, we have the following:

Theorem 3.6.10. *A (metrizable) topological linear space E is locally convex if and only if E is linearly homeomorphic to a linear subspace of the (countable) product $\prod_{\lambda \in \Lambda} E_\lambda$ of normed linear spaces E_λ .*

Proof. As is easily observed, the product of locally convex topological linear spaces is locally convex, and so is any linear subspace of a locally convex topological linear space. Moreover, the countable product of metrizable spaces is metrizable. Then, the “if” part follows.

We show the “only if” part. By the local convexity, E has a neighborhood basis $\{V_\lambda \mid \lambda \in \Lambda\}$ of $\mathbf{0} \in E$ consisting of circled closed convex sets (cf. Proposition 3.4.2), where $\text{card } \Lambda = \aleph_0$ if E is metrizable (Theorem 3.6.1). For each $\lambda \in \Lambda$, let F_λ be a maximal linear subspace of E contained in V_λ . (The existence of F_λ is guaranteed by Zorn’s Lemma.) Then, F_λ is closed in E . Let $q_\lambda : E \rightarrow E/F_\lambda$ be the natural linear surjection, where we do not give the quotient topology to E/F_λ but we want to define a norm on E/F_λ .

Observe that $q_\lambda(V_\lambda)$ is a circled convex set in E/F_λ and $\mathbf{0} \in \text{core } q_\lambda(V_\lambda)$. Moreover, $\mathbb{R}_+\xi \not\subset q_\lambda(V_\lambda)$ for each $\xi \in (E/F_\lambda) \setminus \{\mathbf{0}\}$. Indeed, take $x \in E \setminus F_\lambda$ so that $q_\lambda(x) = \xi$. By the maximality of F_λ , $\mathbb{R}x + F_\lambda \not\subset V_\lambda$, i.e., $tx + y \notin V_\lambda$ for some $t \in \mathbb{R}$ and $y \in F_\lambda$, where we can take $t > 0$ because V_λ is circled. For each $z \in F_\lambda$,

$$tx + y = \frac{1}{2}(2tx + z) + \frac{1}{2}(2y - z).$$

Since $2y - z \in F_\lambda \subset V_\lambda$, it follows that $2tx + z \notin V_\lambda$. Then, $2t\xi = q_\lambda(2tx) \notin q_\lambda(V_\lambda)$.

By Proposition 3.3.4, the Minkowski functional $p_\lambda = p_{q_\lambda(V_\lambda)} : E/F_\lambda \rightarrow \mathbb{R}$ for $q_\lambda(V_\lambda)$ is a norm. Thus, we have a normed linear space $E_\lambda = (E/F_\lambda, p_\lambda)$. Observe that

$$\begin{aligned} \mathbf{0} \in \text{int } V_\lambda &= \text{core } V_\lambda \subset q_\lambda^{-1}(\text{core } q_\lambda(V_\lambda)) \\ &= q_\lambda^{-1}(p_{q_\lambda(V_\lambda)}^{-1}([0, 1])) = (p_\lambda q_\lambda)^{-1}([0, 1]). \end{aligned}$$

By Proposition 3.4.11, the sublinear functional $p_\lambda q_\lambda : E \rightarrow \mathbb{R}$ is continuous, which implies that $q_\lambda : E \rightarrow E_\lambda$ is continuous.

Let $h : E \rightarrow \prod_{\lambda \in \Lambda} E_\lambda$ be the linear map¹⁹ defined by $h(x) = (q_\lambda(x))_{\lambda \in \Lambda}$. If $x \neq \mathbf{0} \in E$ then $x \notin V_\lambda$ (so $x \notin F_\lambda$) for some $\lambda \in \Lambda$, which implies $q_\lambda(x) \neq 0$, hence $h(x) \neq \mathbf{0}$. Thus, h is a continuous linear injection. To see that h is an embedding, it suffices to show that

$$h(V_\lambda) \supset h(E) \cap \text{pr}_\lambda^{-1}(p_\lambda^{-1}([0, \frac{1}{2}])) \text{ for each } \lambda \in \Lambda.$$

¹⁹That is, a continuous linear function.

If $p_\lambda(\text{pr}_\lambda h(x)) < \frac{1}{2}$ then

$$q_\lambda(2x) = \text{pr}_\lambda h(2x) \in p_\lambda^{-1}([0, 1]) \subset q_\lambda(V_\lambda),$$

hence $2x - y \in F_\lambda$ for some $y \in V_\lambda$. Then, it follows that

$$x = \frac{1}{2}(2x - y) + \frac{1}{2}y \in V_\lambda,$$

so $h(x) \in h(V_\lambda)$. This completes the proof. \square

Combining Theorem 3.6.10 with Proposition 3.6.5 and Corollary 3.6.7, we have the following:

Corollary 3.6.11. *A topological linear space E is a Fréchet space if and only if E is linearly homeomorphic to a closed linear subspace of the countable product $\prod_{i \in \mathbb{N}} E_i$ of Banach spaces E_i .* \square

3.7 The Closed Graph and Open Mapping Theorems

This section is devoted to two very important theorems, the Closed Graph Theorem and the Open Mapping Theorem. They are proved using the Baire Category Theorem 2.5.1.

Theorem 3.7.1 (CLOSED GRAPH THEOREM). *Let E and F be completely metrizable topological linear spaces and $f : E \rightarrow F$ be a linear function. If the graph of f is closed in $E \times F$, then f is continuous.*

Proof. It suffices to show the continuity of f at $\mathbf{0} \in E$. Let d and ρ be admissible complete invariant metrics for E and F , respectively (cf. Proposition 3.6.4).

First, we show that for each $\varepsilon > 0$, there is some $\delta(\varepsilon) > 0$ such that $B_d(\mathbf{0}, \delta(\varepsilon)) \subset \text{cl } f^{-1}(B_\rho(\mathbf{0}, \varepsilon))$. Since $F = \bigcup_{n \in \mathbb{N}} nB_\rho(\mathbf{0}, \varepsilon/2)$ and f is linear, it follows that $E = \bigcup_{n \in \mathbb{N}} nf^{-1}(B_\rho(\mathbf{0}, \varepsilon/2))$. By the Baire Category Theorem 2.5.1, $\text{int cl } nf^{-1}(B_\rho(\mathbf{0}, \varepsilon/2)) \neq \emptyset$ for some $n \in \mathbb{N}$, which implies that $\text{int cl } f^{-1}(B_\rho(\mathbf{0}, \varepsilon/2)) \neq \emptyset$. Let $z \in \text{int cl } f^{-1}(B_\rho(\mathbf{0}, \varepsilon/2))$ and choose $\delta(\varepsilon) > 0$ so that

$$z + B_d(\mathbf{0}, \delta(\varepsilon)) = B_d(z, \delta(\varepsilon)) \subset \text{cl } f^{-1}(B_\rho(\mathbf{0}, \varepsilon/2)).$$

Then, it follows that

$$B_d(\mathbf{0}, \delta(\varepsilon)) \subset \text{cl } f^{-1}(B_\rho(\mathbf{0}, \varepsilon/2)) - z \subset \text{cl } f^{-1}(B_\rho(\mathbf{0}, \varepsilon)).$$

The second inclusion can be proved as follows: for each $y \in \text{cl } f^{-1}(B_\rho(\mathbf{0}, \varepsilon/2))$ and $\eta > 0$, we have $y', z' \in f^{-1}(B_\rho(\mathbf{0}, \varepsilon/2))$ such that $d(y, y'), d(z, z') < \eta/2$, which implies $d(y - z, y' - z') < \eta$. Observe that

$$\rho(f(y' - z'), \mathbf{0}) = \rho(f(y'), f(z')) \leq \rho(f(y'), \mathbf{0}) + \rho(f(z'), \mathbf{0}) < \varepsilon,$$

which means $y' - z' \in f^{-1}(B_\rho(\mathbf{0}, \varepsilon))$. Therefore, $y - z \in \text{cl } f^{-1}(B_\rho(\mathbf{0}, \varepsilon))$.

Now, for each $\varepsilon > 0$ and $x \in B_d(\mathbf{0}, \delta(\varepsilon/2))$, we can inductively choose $x_n \in E$, $n \in \mathbb{N}$, so that $x_n \in f^{-1}(B_\rho(\mathbf{0}, 2^{-n}\varepsilon))$ and

$$d(x, \sum_{i=1}^n x_i) = d(x - \sum_{i=1}^n x_i, \mathbf{0}) < \min \{2^{-n}, \delta(2^{-n-1}\varepsilon)\}.$$

Indeed, if x_1, \dots, x_{n-1} have been chosen, then

$$x - \sum_{i=1}^{n-1} x_i \in B_d(\mathbf{0}, \delta(2^{-n}\varepsilon)) \subset \text{cl } f^{-1}(B_\rho(\mathbf{0}, 2^{-n}\varepsilon)),$$

hence we can choose $x_n \in f^{-1}(B_\rho(\mathbf{0}, 2^{-n}\varepsilon))$ so that

$$d(x, \sum_{i=1}^n x_i) = d(x - \sum_{i=1}^{n-1} x_i, x_n) < \min \{2^{-n}, \delta(2^{-n-1}\varepsilon)\}.$$

Since $\rho(f(x_n), \mathbf{0}) < 2^{-n}\varepsilon$ for each $n \in \mathbb{N}$, it follows that $(f(\sum_{i=1}^n x_i))_{n \in \mathbb{N}}$ is a Cauchy sequence, which converges to some $y \in F$. For each $n \in \mathbb{N}$,

$$\rho(f(\sum_{i=1}^n x_i), \mathbf{0}) \leq \sum_{i=1}^n \rho(f(x_i), \mathbf{0}) < \sum_{i=1}^n 2^{-i}\varepsilon < \varepsilon,$$

hence $y \in \overline{B}_\rho(\mathbf{0}, \varepsilon)$. On the other hand, $\sum_{i=1}^n x_i$ converges to x . Since the graph of f is closed in $E \times F$, the point (x, y) belongs to the graph of f , which means $f(x) = y \in \overline{B}_\rho(\mathbf{0}, \varepsilon)$. Thus, we have $f(B_d(\mathbf{0}, \delta(\varepsilon/2))) \subset \overline{B}_\rho(\mathbf{0}, \varepsilon)$. Therefore, f is continuous. \square

Corollary 3.7.2. *Let E and F be completely metrizable topological linear spaces. Then, every continuous linear isomorphism $f : E \rightarrow F$ is a homeomorphism.*

Proof. In general, the continuity of f implies the closedness of the graph of f in $E \times F$. By changing coordinates, the graph of f can be regarded as the graph of f^{-1} . Then, it follows that the graph of f^{-1} is closed in $F \times E$, which implies the continuity of f^{-1} by Theorem 3.7.1. \square

Theorem 3.7.3 (OPEN MAPPING THEOREM). *Let E and F be completely metrizable topological linear spaces. Then, every continuous linear surjection $f : E \rightarrow F$ is open.*

Proof. Since $f^{-1}(\mathbf{0})$ is a closed linear subspace of E , the quotient linear space $E/f^{-1}(\mathbf{0})$ is completely metrizable by Proposition 3.6.8. Then, f induces the continuous linear isomorphism $\tilde{f} : E/f^{-1}(\mathbf{0}) \rightarrow F$. By Corollary 3.7.2, \tilde{f} is a homeomorphism. Note that the quotient map $q : E \rightarrow E/f^{-1}(\mathbf{0})$ is open. Indeed, for every open set U in E , $q^{-1}(q(U)) = U + f^{-1}(\mathbf{0})$ is open in E , which means that $q(U)$ is open in $E/f^{-1}(\mathbf{0})$. Hence, f is also open. \square

Note. In the above, the Closed Graph Theorem is first proved and then the Open Mapping Theorem is obtained as a corollary of the Closed Graph Theorem. Conversely, we can directly prove the Open Mapping Theorem and then obtain the Closed Graph Theorem as a corollary of the Open Mapping Theorem.

Direct Proof of the Open Mapping Theorem. Let d and ρ be admissible complete invariant metrics for E and F , respectively.

First, we show that for each $\varepsilon > 0$, there is some $\delta(\varepsilon) > 0$ such that $B_\rho(\mathbf{0}, \delta(\varepsilon)) \subset \text{cl } f(B_d(\mathbf{0}, \varepsilon))$. Since $E = \bigcup_{n \in \mathbb{N}} nB_d(\mathbf{0}, \varepsilon/2)$, it follows that $F = f(E) = \bigcup_{n \in \mathbb{N}} nf(B_d(\mathbf{0}, \varepsilon/2))$. By the Baire Category Theorem 2.5.1, $\text{int } \text{cl } nf(B_d(\mathbf{0}, \varepsilon/2)) \neq \emptyset$ for some $n \in \mathbb{N}$, which implies that $\text{int } \text{cl } f(B_d(\mathbf{0}, \varepsilon/2)) \neq \emptyset$. Let $z \in \text{int } \text{cl } f(B_d(\mathbf{0}, \varepsilon/2))$ and choose $\delta(\varepsilon) > 0$ so that

$$z + B_\rho(\mathbf{0}, \delta(\varepsilon)) = B_\rho(z, \delta(\varepsilon)) \subset \text{cl } f(B_d(\mathbf{0}, \varepsilon/2)).$$

Then, it follows that

$$B_\rho(\mathbf{0}, \delta(\varepsilon)) \subset \text{cl } f(B_d(\mathbf{0}, \varepsilon/2)) - z \subset \text{cl } f(B_d(\mathbf{0}, \varepsilon)),$$

where the second inclusion can be seen as follows: for $y \in \text{cl } f(B_d(\mathbf{0}, \varepsilon/2))$ and $\eta > 0$, choose $y', z' \in B_d(\mathbf{0}, \varepsilon/2)$ so that $\rho(y, f(y')), \rho(z, f(z')) < \eta/2$. Then, observe that $\rho(y - z, f(y' - z')) < \eta$ and $d(y' - z', \mathbf{0}) = d(y', z') < \varepsilon$, hence $y - z \in \text{cl } f(B_d(\mathbf{0}, \varepsilon))$.

Next, we prove that $\text{cl } f(B_d(\mathbf{0}, \varepsilon/2)) \subset f(\overline{B}_d(\mathbf{0}, \varepsilon))$ for each $\varepsilon > 0$. For each $y \in \text{cl } f(B_d(\mathbf{0}, \varepsilon/2))$, choose $x_1 \in B_d(\mathbf{0}, \varepsilon/2)$ so that

$$\rho(y, f(x_1)) < \min\{2^{-1}, \delta(2^{-2}\varepsilon)\}.$$

By induction, we can choose $x_n \in B_d(\mathbf{0}, 2^{-n}\varepsilon)$, $n \in \mathbb{N}$, so that

$$\rho(y, f(\sum_{i=1}^n x_i)) = \rho(y - \sum_{i=1}^n f(x_i), \mathbf{0}) < \min\{2^{-n}, \delta(2^{-n-1}\varepsilon)\}.$$

Indeed, if x_1, \dots, x_{n-1} have been chosen, then

$$y - \sum_{i=1}^{n-1} f(x_i) \in B_\rho(\mathbf{0}, \delta(2^{-n}\varepsilon)) \subset \text{cl } f(B_d(\mathbf{0}, 2^{-n}\varepsilon)),$$

hence we can choose $x_n \in B_d(\mathbf{0}, 2^{-n}\varepsilon)$ so that

$$\begin{aligned} \rho(y, f(\sum_{i=1}^n x_i)) &= \rho(y - \sum_{i=1}^{n-1} f(x_i), f(x_n)) \\ &< \min\{2^{-n}, \delta(2^{-n-1}\varepsilon)\}. \end{aligned}$$

Since $(\sum_{i=1}^n x_i)_{n \in \mathbb{N}}$ is a Cauchy sequence in E , it converges to some $x \in E$. On the other hand, $(f(\sum_{i=1}^n x_i))_{n \in \mathbb{N}}$ converges to y . By the continuity of f , we have $f(x) = y$. For each $n \in \mathbb{N}$,

$$d(\sum_{i=1}^n x_i, \mathbf{0}) \leq \sum_{i=1}^n d(x_i, \mathbf{0}) < \sum_{i=1}^n 2^{-i}\varepsilon < \varepsilon,$$

hence $x \in \overline{B}_d(\mathbf{0}, \varepsilon)$. Thus, it follows that $\text{cl } f(B_d(\mathbf{0}, \varepsilon/2)) \subset f(\overline{B}_d(\mathbf{0}, \varepsilon))$.

To see that f is open, let U be an open set in E . For each $x \in U$, choose $\varepsilon > 0$ so that $\overline{B}_d(\mathbf{0}, \varepsilon) \subset -x + U$. Since

$$B_\rho(\mathbf{0}, \delta(\varepsilon/2)) \subset \text{cl } f(B_d(\mathbf{0}, \varepsilon/2)) \subset f(\overline{B}_d(\mathbf{0}, \varepsilon)) \subset -f(x) + f(U),$$

it follows that $B_\rho(f(x), \delta(\varepsilon/2)) \subset f(U)$. Hence, $f(U)$ is open in F . □

Now, using the Open Mapping Theorem, we shall prove the Closed Graph Theorem.

Proof of the Closed Graph Theorem. The product space $E \times F$ is a completely metrizable topological linear space. The graph G of f is a linear subspace of $E \times F$ that is completely metrizable because it is closed in $E \times F$. Since $p = \text{pr}_E|_G : G \rightarrow E$ is a homeomorphism by the Open Mapping Theorem, $f = \text{pr}_F \circ p^{-1}$ is continuous. □

Remark 13. In both the Closed Graph Theorem and the Open Mapping Theorem, the completeness is essential. Let $E = (\ell_1, \|\cdot\|_2)$, where $\ell_1 \subset \ell_2$ as sets and $\|\cdot\|_2$ is the norm inherited from ℓ_2 . Then, E is not completely metrizable. Indeed, if so, it would be closed in ℓ_2 by Corollary 3.6.7, but E is dense in ℓ_2 and $E \neq \ell_2$. The linear bijection $f = \text{id} : \ell_1 \rightarrow E$ is continuous, but is not a homeomorphism, so it is not an open map. It follows from the continuity of f that the graph of f is closed in $\ell_1 \times E$, hence the graph of f^{-1} is closed in $E \times \ell_1$. However, $f^{-1} : E \rightarrow \ell_1$ is not continuous.

3.8 Continuous Selections

Let X and Y be spaces and $\varphi : X \rightarrow \mathfrak{P}(Y)$ be a set-valued function, where $\mathfrak{P}(Y)$ is the power set of Y . We denote $\mathfrak{P}_0(Y) = \mathfrak{P}(Y) \setminus \{\emptyset\}$. A **(continuous) selection** for φ is a map $f : X \rightarrow Y$ such that $f(x) \in \varphi(x)$ for each $x \in X$. For a topological linear space Y , we denote by $\text{Conv}(Y)$ the set of all non-empty convex sets in Y . In this section, we consider the problem of when a convex-valued function $\varphi : X \rightarrow \text{Conv}(Y)$ has a selection.

It is said that $\varphi : X \rightarrow \mathfrak{P}(Y)$ is **lower semi-continuous (l.s.c.)** (resp. **upper semi-continuous (u.s.c.)**) if, for each open set V in Y ,

$$\{x \in X \mid \varphi(x) \cap V \neq \emptyset\} \text{ (resp. } \{x \in X \mid \varphi(x) \subset V\}) \text{ is open in } X;$$

equivalently, for each open set V in Y and $x_0 \in X$ such that $\varphi(x_0) \cap V \neq \emptyset$ (resp. $\varphi(x_0) \subset V$), there exists a neighborhood U of x_0 in X such that $\varphi(x) \cap V \neq \emptyset$ (resp. $\varphi(x) \subset V$) for every $x \in U$. We say that φ is **continuous** if $\varphi : X \rightarrow \mathfrak{P}(Y)$ is l.s.c. and u.s.c. The continuity of φ coincides with that in the usual sense when $\mathfrak{P}(Y)$ is regarded as a space with the topology generated by the following sets:

$$U^- = \{A \in \mathfrak{P}(Y) \mid A \cap U \neq \emptyset\} \text{ and } U^+ = \{A \in \mathfrak{P}(Y) \mid A \subset U\},$$

where U is non-empty and open in Y . This topology is called the **Vietoris topology**, where \emptyset is isolated because $\{\emptyset\} = \emptyset^+$ ($\emptyset \notin U^-$ for any open set U in Y). The Vietoris topology has an open basis consisting of the following sets: $V(\emptyset) = \{\emptyset\}$ and

$$V(U_1, \dots, U_n) = \{A \subset Y \mid A \subset \bigcup_{i=1}^n U_i, \forall i = 1, \dots, n, A \cap U_i \neq \emptyset\}$$

$$= \left(\bigcup_{i=1}^n U_i \right)^+ \cap \bigcap_{i=1}^n U_i^-,$$

where $n \in \mathbb{N}$ and U_1, \dots, U_n are open in Y . In fact, $U^- = \mathbb{V}(U, X)$ and $U^+ = \mathbb{V}(U) \cup \mathbb{V}(\emptyset)$. The subspace $\mathbb{F}_1(Y) = \{\{y\} \mid y \in Y\}$ of $\mathfrak{P}_0(Y)$ consisting of all singletons is homeomorphic to Y because $U^+ \cap \mathbb{F}_1(Y) = U^- \cap \mathbb{F}_1(Y) = \mathbb{F}_1(U)$ for each open set U in Y . It should be noted that $\mathfrak{P}_0(Y)$ with the Vietoris topology is not T_1 in general.

For example, the space $\mathfrak{P}_0(\mathbf{I})$ is not T_1 . Indeed, for any neighborhood \mathcal{U} of $\mathbf{I} \in \mathfrak{P}_0(\mathbf{I})$, there are open sets U_1, \dots, U_n in \mathbf{I} such that $\mathbf{I} \in \mathbb{V}(U_1, \dots, U_n) \subset \mathcal{U}$. Then, $D \in \mathbb{V}(U_1, \dots, U_n) \subset \mathcal{U}$ for every dense subset $D \subset \mathbf{I}$. In particular, $\mathbf{I} \cap \mathcal{Q} \in \mathcal{U}$.

The subspace $\text{Comp}(Y)$ of $\mathfrak{P}(Y)$ consisting of all non-empty compact sets is Hausdorff.²⁰ Indeed, for each $A \neq B \in \text{Comp}(Y)$, we may assume that $A \setminus B \neq \emptyset$. Take $y_0 \in B \setminus A$. Because of the compactness of A , we have disjoint open sets U and V in Y such that $A \subset U$ and $y_0 \in V$. Then, $A \in U^+$, $B \subset V^-$, and $U^+ \cap V^- \neq \emptyset$. It will be proved that $\text{Comp}(Y)$ is metrizable if Y is metrizable (Proposition 5.12.4). Moreover, $\text{Cld}(Y)$ is metrizable if and only if Y is compact and metrizable (cf. Note after Proposition 5.12.4).

By the same argument as above, it follows that if Y is regular then the subspace $\text{Cld}(Y)$ of $\mathfrak{P}(Y)$ consisting of all non-empty closed sets is Hausdorff. One should note that the converse is also true, that is, if $\text{Cld}(Y)$ is Hausdorff then Y is regular. When Y is not regular, we have a closed set $A \subset Y$ and $y_0 \in Y \setminus A$ such that if U and V are open sets with $A \subset U$ and $y_0 \in V$ then $U \cap V \neq \emptyset$. Let $B = A \cup \{y_0\} \in \text{Cld}(Y)$ and let $U_1, \dots, U_n, U'_1, \dots, U'_n$ be open sets in Y such that

$$A \in \mathbb{V}(U_1, \dots, U_n) \text{ and } B \in \mathbb{V}(U'_1, \dots, U'_n).$$

Let $U_0 = \bigcap \{U'_i \mid U'_i \cap A = \emptyset\}$. Since $y_0 \in U_0$, we have $y_1 \in U_0 \cap \bigcup_{i=1}^n U_i$. It follows that

$$A \cup \{y_1\} \in \mathbb{V}(U_1, \dots, U_n) \cap \mathbb{V}(U'_1, \dots, U'_n).$$

Thus, $\text{Cld}(Y)$ is not Hausdorff.

Proposition 3.8.1. *For a function $g : Y \rightarrow X$, the set-valued function $g^{-1} : X \rightarrow \mathfrak{P}(Y)$ is l.s.c. (resp. u.s.c.) if and only if g is open (resp. closed).*

Proof. This follows from the fact that, for $V \subset Y$,

$$\begin{aligned} \{x \in X \mid g^{-1}(x) \cap V \neq \emptyset\} &= g(V) \quad \text{and} \\ \{x \in X \mid g^{-1}(x) \subset V\} &= X \setminus \{x \in X \mid g^{-1}(x) \cap (X \setminus V) \neq \emptyset\} \\ &= X \setminus g(X \setminus V). \end{aligned} \quad \square$$

²⁰Recall that Y is assumed to be Hausdorff.

Because of the following proposition, we consider the selection problem for l.s.c. set-valued functions.

Proposition 3.8.2. *Let $\varphi : X \rightarrow \mathfrak{P}_0(Y)$ be a set-valued function. Assume that, for each $x_0 \in X$ and $y_0 \in \varphi(x_0)$, there exists a neighborhood U of x_0 in X and a selection $f : U \rightarrow Y$ for $\varphi|_U$ such that $f(x_0) = y_0$. Then, φ is l.s.c.*

Proof. Let V be an open set in Y and $x_0 \in X$ such that $\varphi(x_0) \cap V \neq \emptyset$. Take any $y_0 \in \varphi(x_0) \cap V$. From the assumption, there is a neighborhood U of x_0 in X with a selection $f : U \rightarrow Y$ for $\varphi|_U$ such that $f(x_0) = y_0$. Then, $f^{-1}(V)$ is a neighborhood of x_0 in X and $f(x) \in \varphi(x) \cap V$ for each $x \in f^{-1}(V)$. \square

Lemma 3.8.3. *Let $\varphi, \psi : X \rightarrow \mathfrak{P}(Y)$ be set-valued functions such that $\text{cl } \varphi(x) = \text{cl } \psi(x)$ for each $x \in X$. If φ is l.s.c. then so is ψ .*

Sketch of Proof. This follows from the fact that, for each open set V in Y and $B \subset Y$, $V \cap B \neq \emptyset$ if and only if $V \cap \text{cl } B \neq \emptyset$.

Lemma 3.8.4. *Let $\varphi : X \rightarrow \mathfrak{P}(Y)$ be l.s.c., A be a closed set in X , and $f : A \rightarrow Y$ be a selection for $\varphi|_A$. Define $\psi : X \rightarrow \mathfrak{P}(Y)$ by*

$$\psi(x) = \begin{cases} \{f(x)\} & \text{if } x \in A, \\ \varphi(x) & \text{otherwise.} \end{cases}$$

Then, ψ is also l.s.c.

Proof. For each open set V in Y , $f^{-1}(V)$ is open in A and

$$f^{-1}(V) \subset \{x \in X \mid \varphi(x) \cap V \neq \emptyset\},$$

where the latter set is open in X because φ is l.s.c. Then, we can choose an open set U in X so that $f^{-1}(V) = U \cap A$ and $U \subset \{x \in X \mid \varphi(x) \cap V \neq \emptyset\}$. Observe that

$$\{x \in X \mid \psi(x) \cap V \neq \emptyset\} = U \cup (\{x \in X \mid \varphi(x) \cap V \neq \emptyset\} \setminus A).$$

Thus, it follows that ψ is l.s.c. \square

For each $W \subset Y^2$ and $y_0 \in Y$, we denote

$$W(y_0) = \{y \in Y \mid (y_0, y) \in W\}.$$

If W is a neighborhood of the diagonal $\Delta_Y = \{(y, y) \mid y \in Y\}$ in Y^2 , then $W(y_0)$ is a neighborhood of y_0 in Y .

Lemma 3.8.5. *Let $\varphi : X \rightarrow \mathfrak{P}(Y)$ be l.s.c., $f : X \rightarrow Y$ be a map, and W be a neighborhood of Δ_Y in Y^2 . Define a set-valued function $\psi : X \rightarrow \mathfrak{P}(Y)$ by $\psi(x) = \varphi(x) \cap W(f(x))$ for each $x \in X$. Then, ψ is l.s.c.*

Proof. Let V be an open set in Y and $x_0 \in X$ such that $\psi(x_0) \cap V \neq \emptyset$. Take any $y_0 \in \varphi(x_0) \cap W(f(x_0)) \cap V$. Since $(f(x_0), y_0) \in W$, there are open sets V_1 and V_2 in Y such that $(f(x_0), y_0) \in V_1 \times V_2 \subset W$. Then, x_0 has the following open neighborhood in X :

$$U = f^{-1}(V_1) \cap \{x \in X \mid \varphi(x) \cap V_2 \cap V \neq \emptyset\}.$$

For each $x \in U$, we have $y \in \varphi(x) \cap V_2 \cap V$. Since $(f(x), y) \in V_1 \times V_2 \subset W$, it follows that $y \in \varphi(x) \cap W(f(x)) \cap V$, hence $\psi(x) \cap V \neq \emptyset$. Therefore, ψ is l.s.c. \square

Let E be a linear space. The set of all non-empty convex sets in E is denoted by $\text{Conv}(E)$. Recall that $\langle A \rangle$ denotes the convex hull of $A \subset E$.

Lemma 3.8.6. *Let E be a topological linear space and $\varphi : X \rightarrow \mathfrak{P}_0(E)$ be an l.s.c. set-valued function. Define a convex-valued function $\psi : X \rightarrow \text{Conv}(E)$ by $\psi(x) = \langle \varphi(x) \rangle$ for each $x \in X$. Then, ψ is also l.s.c.*

Proof. Let V be an open set in E and $x_0 \in X$ such that $\psi(x_0) \cap V \neq \emptyset$. Choose any $y_0 = \sum_{i=1}^n t_i y_i \in \psi(x_0) \cap V$, where $y_1, \dots, y_n \in \varphi(x_0)$ and $t_1, \dots, t_n \geq 0$ with $\sum_{i=1}^n t_i = 1$. Then, each y_i has an open neighborhood V_i such that $t_1 V_1 + \dots + t_n V_n \subset V$. Since φ is l.s.c.,

$$U = \bigcap_{i=1}^n \{x \in X \mid \varphi(x) \cap V_i \neq \emptyset\}$$

is an open neighborhood of x_0 in X . For each $x \in U$, let $z_i \in \varphi(x) \cap V_i$, $i = 1, \dots, n$. Then, $\sum_{i=1}^n t_i z_i \in \psi(x) \cap V$, hence $\psi(x) \cap V \neq \emptyset$. Therefore, ψ is l.s.c. \square

Lemma 3.8.7. *Let X be paracompact, E be a topological linear space, and $\varphi : X \rightarrow \text{Conv}(E)$ be an l.s.c. convex-valued function. Then, for each convex open neighborhood V of $\mathbf{0}$ in E , there exists a map $f : X \rightarrow E$ such that $f(x) \in \varphi(x) + V$ for each $x \in X$.*

Proof. For each $y \in E$, let

$$U_y = \{x \in X \mid \varphi(x) \cap (y - V) \neq \emptyset\}.$$

Since φ is l.s.c., we have $\mathcal{U} = \{U_y \mid y \in E\} \in \text{cov}(X)$. From paracompactness, X has a locally finite partition of unity $(f_\lambda)_{\lambda \in \Lambda}$ subordinated to \mathcal{U} . For each $\lambda \in \Lambda$, choose $y_\lambda \in E$ so that $\text{supp } f_\lambda \subset U_{y_\lambda}$. We define a map $f : X \rightarrow E$ by $f(x) = \sum_{\lambda \in \Lambda} f_\lambda(x) y_\lambda$. If $f_\lambda(x) \neq 0$ then $x \in U_{y_\lambda}$, which means that $\varphi(x) \cap (y_\lambda - V) \neq \emptyset$, i.e., $y_\lambda \in \varphi(x) + V$. Since each $\varphi(x) + V$ is convex, $f(x) \in \varphi(x) + V$. \square

Now, we can prove the following:

Theorem 3.8.8 (MICHAEL SELECTION THEOREM). *Let X be a paracompact space and $E = (E, d)$ be a locally convex metric linear space.²¹ Then, every l.s.c. convex-valued function $\varphi : X \rightarrow \text{Conv}(E)$ admits a selection if each $\varphi(x)$ is d -complete. Moreover, if A is a closed set in X then each selection $f : A \rightarrow E$ for $\varphi|_A$ can extend to a selection $\tilde{f} : X \rightarrow E$ for φ .*

Proof. Let $\{V_i \mid i \in \mathbb{N}\}$ be a neighborhood basis of $\mathbf{0}$ in E such that each V_i is symmetric, convex, and $\text{diam } V_i < 2^{-(i+1)}$. By induction, we construct maps $f_i : X \rightarrow E, i \in \mathbb{N}$, so that, for each $x \in X$ and $i \in \mathbb{N}$,

- (1) $f_i(x) \in \varphi(x) + V_i$ and
- (2) $d(f_{i+1}(x), f_i(x)) < 2^{-i}$.

The existence of f_1 is guaranteed by Lemma 3.8.7. Assume we have maps f_1, \dots, f_n satisfying (1) and (2). Define $\psi : X \rightarrow \text{Conv}(E)$ by

$$\psi(x) = \varphi(x) \cap (f_n(x) + V_n) \quad \text{for each } x \in X.$$

Since V_n is symmetric, we have $\psi(x) \neq \emptyset$ by (1). Consider the neighborhood $W = \{(x, y) \in E^2 \mid y - x \in V_n\}$ of Δ_E in E^2 . Then, $W(f_n(x)) = f_n(x) + V_n$. By Lemma 3.8.5, ψ is l.s.c. We can apply Lemma 3.8.7 to obtain a map $f_{n+1} : X \rightarrow E$ such that

$$f_{n+1}(x) \in \psi(x) + V_{n+1} \quad \text{for each } x \in X.$$

Then, as is easily observed, f_{n+1} satisfies (1) and (2). Thus, we have the desired sequence of maps $f_i, i \in \mathbb{N}$.

Using maps $f_i : X \rightarrow E, i \in \mathbb{N}$, we shall define a selection $f : X \rightarrow E$ for φ . For each $x \in X$ and $i \in \mathbb{N}$, we have $x_i \in \varphi(x)$ such that $d(f_i(x), x_i) < 2^{-(i+1)}$ by (1). Then, $(x_i)_{i \in \mathbb{N}}$ is Cauchy in $\varphi(x)$. Since $\varphi(x)$ is complete, $(x_i)_{i \in \mathbb{N}}$ converges to $f(x) \in \varphi(x)$. Thus, we have $f : X \rightarrow E$. Note that $(f_i)_{i \in \mathbb{N}}$ uniformly converges to f , so f is continuous. Hence, f is a selection for φ .

For the additional statement, apply Lemma 3.8.4. □

Concerning factors of a metric linear space, we have the following:

Corollary 3.8.9 (BARTLE–GRAVES–MICHAEL). *Let E be a locally convex metric linear space and F be a linear subspace of E that is complete (so a Fréchet space). Then, $E \approx F \times E/F$. In particular, $E \approx \mathbb{R} \times G$ for some metric linear space G .*

Proof. Note that the quotient space E/F is metrizable (Proposition 3.6.8) and the natural map $g : E \rightarrow E/F$ is open, hence $g^{-1} : E/F \rightarrow \text{Conv}(E)$ is l.s.c. by Proposition 3.8.1. Since $g^{-1}g(x) = x + F$ is complete for each $x \in E$, we apply the Michael Selection Theorem 3.8.8 to obtain a map $f : E/F \rightarrow E$ that is a

²¹Recall that a metric linear space is a linear space with a linear metric (cf. Sect. 3.5).

selection for g^{-1} , i.e., $gf = \text{id}$. Then, $x - fg(x) \in F$ for each $x \in E$. Hence, a homeomorphism $h : E \rightarrow F \times (E/F)$ can be defined by

$$h(x) = (x - fg(x), g(x)) \text{ for each } x \in E.$$

In fact, $h^{-1}(y, z) = y + f(z)$ for each $(y, z) \in F \times E/F$. □

By combining the Michael Selection Theorem 3.8.8 and the Open Mapping Theorem 3.7.3, the following Bartle–Graves Theorem can be obtained as a corollary:

Theorem 3.8.10 (BARTLE–GRAVES). *Let E and F be Fréchet spaces and $f : E \rightarrow F$ be a continuous linear surjection. Then, there is a map $g : F \rightarrow E$ such that $fg = \text{id}$. Therefore, $E \approx F \times \ker f$ by the homeomorphism h defined as follows:*

$$h(x) = (f(x), x - gf(x)) \text{ for each } x \in E. \quad \square$$

We show that each Banach space is a (topological) factor of $\ell_1(\Gamma)$. To this end, we need the following:

Theorem 3.8.11 (BANACH–MAZUR, KLEE). *For every Banach space E , there is a continuous linear surjection $q : \ell_1(\Gamma) \rightarrow E$, where $\text{card } \Gamma = \text{dens } E$.*

Proof. The unit closed ball \mathbf{B}_E of E has a dense set $\{\mathbf{e}_\gamma \mid \gamma \in \Gamma\}$. Since $\sum_{\gamma \in \Gamma} |x(\gamma)| = \|x\| < \infty$ for each $x \in \ell_1(\Gamma)$ and E is complete, we can define a linear map $q : \ell_1(\Gamma) \rightarrow E$ as follows:

$$q(x) = \sum_{\gamma \in \Gamma} x(\gamma)\mathbf{e}_\gamma \text{ for each } x \in E.^{22}$$

Since $\|q(x)\| \leq \sum_{\gamma \in \Gamma} |x(\gamma)| = \|x\|$, it follows that q is continuous.

To see that q is surjective, it suffices to show $\mathbf{B}_E \subset q(\ell_1(\Gamma))$. For each $y \in \mathbf{B}_E$, we can inductively choose \mathbf{e}_{γ_i} , $i \in \mathbb{N}$, so that $\gamma_i \neq \gamma_j$ if $i \neq j$, and

$$\begin{aligned} \|y - \mathbf{e}_{\gamma_1}\| &< 2^{-1}, \quad \|y - \mathbf{e}_{\gamma_1} - 2^{-1}\mathbf{e}_{\gamma_2}\| < 2^{-2}, \\ \|y - \mathbf{e}_{\gamma_1} - 2^{-1}\mathbf{e}_{\gamma_2} - 2^{-2}\mathbf{e}_{\gamma_3}\| &< 2^{-3}, \quad \dots \end{aligned}$$

We have $x \in \ell_1(\Gamma)$ defined by

$$x(\gamma) = \begin{cases} 2^{1-i} & \text{if } \gamma = \gamma_i, \\ 0 & \text{otherwise.} \end{cases}$$

Then, it follows that $y = \sum_{i=1}^{\infty} 2^{1-i}\mathbf{e}_{\gamma_i} = q(x)$. This completes the proof. □

²²See Proposition 1.2.3.

As a combination of the Bartle–Graves Theorem 3.8.10 and Theorem 3.8.11 above, we have the following:

Corollary 3.8.12. *For any Banach space E , there exists a Banach space F such that $E \times F \approx \ell_1(\Gamma)$, where $\text{card } \Gamma = \text{dens } E$. \square*

In the Michael Selection Theorem 3.8.8, the paracompactness of X is necessary. Actually, we have the following characterization:

Theorem 3.8.13. *A space X is paracompact if and only if the following holds for any Banach space E : if $\varphi : X \rightarrow \text{Conv}(E)$ is l.s.c. and each $\varphi(x)$ is closed, then φ has a selection.*

Proof. Since the “only if” part is simply Theorem 3.8.8, it suffices to prove the “if” part. For each $\mathcal{U} \in \text{cov}(X)$, we define $\varphi : X \rightarrow \mathfrak{F}_0(\ell_1(\mathcal{U}))$ as follows:

$$\varphi(x) = \{z \in \ell_1(\mathcal{U}) \mid \|z\| = 1, \forall U \in \mathcal{U}, z(U) \geq 0, z(U) = 0 \text{ if } x \notin U\}.$$

Clearly, each $\varphi(x)$ is a closed convex set.

To see that φ is l.s.c., let W be an open set in $\ell_1(\mathcal{U})$ and $z \in \varphi(x) \cap W$. Choose $\delta > 0$ so that $B(z, 2\delta) \subset W$. Then, we have $V_1, \dots, V_n \in \mathcal{U}[x]$ such that $\sum_{i=1}^n z(V_i) > 1 - \delta$, where $\bigcap_{i=1}^n V_i$ is a neighborhood of x in X . We define $z' \in \ell_1(\mathcal{U})$ as follows:

$$z'(V_i) = \frac{z(V_i)}{\sum_{j=1}^n z(V_j)} \text{ and } z'(U) = 0 \text{ for } U \neq V_1, \dots, V_n.$$

It is easy to see that $z' \in \varphi(x') \cap W$ for every $x' \in \bigcap_{i=1}^n V_i$. Thus, φ is l.s.c.

By the assumption, φ has a selection $f : X \rightarrow \ell_1(\mathcal{U})$. For each $U \in \mathcal{U}$, let $f_U : X \rightarrow \mathbf{I}$ be the map defined by $f_U(x) = f(x)(U)$ for $x \in X$. Then, $(f_U)_{U \in \mathcal{U}}$ is a partition of unity such that $f_U^{-1}((0, 1]) \subset U$ for every $U \in \mathcal{U}$. The result follows from Theorem 2.7.5. \square

Remark 14. Let $g, h : X \rightarrow \mathbb{R}$ be real-valued functions on a space X such that g is u.s.c., h is l.s.c., and $g(x) \leq h(x)$ for each $x \in X$. We define the convex-valued function $\varphi : X \rightarrow \text{Conv}(\mathbb{R})$ by $\varphi(x) = [g(x), h(x)]$ for each $x \in X$. Then, φ is l.s.c. Indeed, for each open set V in \mathbb{R} , let $\varphi(x) \cap V \neq \emptyset$. Take $y \in \varphi(x) \cap V$ and $a < y < b$ so that $[a, b] \subset V$. Since g is u.s.c. and h is l.s.c., x has a neighborhood U in X such that $x' \in U$ implies $g(x') < b$ and $h(x') > a$. Since $g(x') \leq h(x')$, it follows that

$$\varphi(x') \cap V \supset [g(x'), h(x')] \cap [a, b] = [\max\{a, g(x')\}, \min\{b, h(x')\}] \neq \emptyset.$$

Now, we can apply the Michael Selection Theorem 3.8.8 to obtain a map $f : X \rightarrow \mathbb{R}$ such that $g(x) \leq f(x) \leq h(x)$ for each $x \in X$. This is analogous to Theorem 2.7.6.

3.9 Free Topological Linear Spaces

The **free topological linear space** over a space X is a topological linear space $L(X)$ that contains X as a subspace and has the following extension property:

(LE) For an arbitrary topological linear space F , every map $f : X \rightarrow F$ of X uniquely extends to a linear map²³ $\tilde{f} : L(X) \rightarrow F$.

$$\begin{array}{ccc} X & \xrightarrow{f} & F \\ \cap & \nearrow \tilde{f} & \\ L(X) & & \end{array}$$

If such a space $L(X)$ exists, then it is uniquely determined up to linear homeomorphism, that is, if E is a topological linear space that contains X and has the property (LE), then E is linearly homeomorphic to $L(X)$.

Indeed, there exist linear maps $\varphi : L(X) \rightarrow E$ and $\psi : E \rightarrow L(X)$ such that $\varphi|_X = \psi|_X = \text{id}_X$. Since $\text{id}_{L(X)}$ is a linear map extending id_X , it follows from the uniqueness that $\psi\varphi = \text{id}_{L(X)}$. Similarly, we have $\varphi\psi = \text{id}_E$. Therefore, φ is a linear homeomorphism with $\psi = \varphi^{-1}$.

Lemma 3.9.1. *If X is a Tychonoff space,*

- (1) X is a Hamel basis for $L(X)$;
- (2) $L(X)$ is regular.

Proof. (1): First, let F be the linear span of X . Applying (LE), we have a linear map $r : L(X) \rightarrow F$ such that $r|_X = \text{id}_X$. Since $r : L(X) \rightarrow L(X)$ is a linear map extending id_X , we have $r = \text{id}_{L(X)}$, which implies $F = L(X)$, that is, $L(X)$ is generated by X .

To see that X is linearly independent in $L(X)$, let $x_1, \dots, x_n \in X$, where $x_i \neq x_j$ if $i \neq j$. For each $i = 1, \dots, n$, there is a map $f_i : X \rightarrow \mathbf{I}$ such that $f_i(x_i) = 1$ and $f_i(x_j) = 0$ for $j \neq i$. Let $f : X \rightarrow \mathbb{R}^n$ be the map defined by $f(x) = (f_1(x), \dots, f_n(x))$. Then, by (LE), f extends to a linear map $\tilde{f} : L(X) \rightarrow \mathbb{R}^n$, where $\tilde{f}(x_i) = f(x_i) = \mathbf{e}_i$ for each $i = 1, \dots, n$. Since $\mathbf{e}_1, \dots, \mathbf{e}_n$ is linearly independent in \mathbb{R}^n , it follows that $x_1, \dots, x_n \in X$ is linearly independent in $L(X)$.

(2): Due to the Fact in Sect. 3.4 and Proposition 3.4.2, it suffices to show that $\{\mathbf{0}\}$ is closed in $L(X)$. Each $z \in L(X) \setminus \{\mathbf{0}\}$ can be uniquely represented as follows:

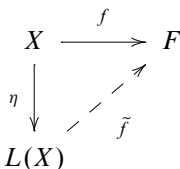
$$z = \sum_{i=1}^n t_i x_i, \quad x_i \in X, \quad t_i \in \mathbb{R} \setminus \{0\},$$

²³That is, a continuous linear function.

where $x_i \neq x_j$ if $i \neq j$. There is a map $f : X \rightarrow \mathbf{I}$ such that $f(x_1) = 1$ and $f(x_i) = 0$ for each $i = 2, \dots, n$. By (LE), f extends to a linear map $\tilde{f} : L(X) \rightarrow \mathbb{R}$. Then, $\tilde{f}(z) = t_1 f(x_1) = t_1 \neq 0 = \tilde{f}(\mathbf{0})$. Hence, $\tilde{f}^{-1}(\mathbb{R} \setminus \{0\})$ is an open neighborhood of z in $L(X)$ that misses $\mathbf{0}$. \square

Remark 15. In the definition of a free topological linear space $L(X)$, specify a map $\eta : X \rightarrow L(X)$ instead of assuming $X \subset L(X)$ and replace the property (LE) with the following universality:

(*) For each map $f : X \rightarrow F$ of X to an arbitrary topological linear space F , there exists a unique linear map $\tilde{f} : L(X) \rightarrow F$ such that $\tilde{f}\eta = f$.



Then, we can show that η is an embedding if X is a Tychonoff space.

To see that η is injective, let $x \neq y \in X$. Then, there is a map $f : X \rightarrow \mathbf{I}$ with $f(x) = 0$ and $f(y) = 1$. By (*), we have a linear map $\tilde{f} : L(X) \rightarrow \mathbb{R}$ such that $\tilde{f}\eta = f$. Then, observe $\eta(x) \neq \eta(y)$.

To show that $\eta : X \rightarrow \eta(X)$ is open, let U be an open set in X . For each $x \in U$, there is a map $g : X \rightarrow \mathbf{I}$ such that $g(x) = 0$ and $g(X \setminus U) = 1$. By (*), we have a linear map $\tilde{g} : L(X) \rightarrow \mathbb{R}$ such that $\tilde{g}\eta = g$. Then, $V = \tilde{g}^{-1}((-\frac{1}{2}, \frac{1}{2}))$ is an open neighborhood $\eta(x)$ in $L(X)$. Since $\eta^{-1}(V) = g^{-1}([0, \frac{1}{2})) \subset U$, it follows that $V \cap \eta(X) \subset \eta(U)$, hence $\eta(U)$ is a neighborhood of $\eta(x)$ in $\eta(X)$. This means that $\eta(U)$ is open in $\eta(X)$. Thus, $\eta : X \rightarrow \eta(X)$ is open.

Since η is an embedding, X can be identified with $\eta(X)$, which is a subspace of $L(X)$. Then, (*) is equivalent to (LE). Here, it should be noted that the uniqueness of \tilde{f} in (*) is not used to prove that η is an embedding. Moreover, the linear map \tilde{f} in (*) is unique if and only if $L(X)$ is generated by $\eta(X)$. (For the “only if” part, refer to the proof of Lemma 3.9.1(1).)

Theorem 3.9.2. *For every Tychonoff space X , there exists the free topological linear space $L(X)$ over X .*

Proof. There exists a collection $\mathcal{F} = \{f_\lambda : X \rightarrow F_\lambda \mid \lambda \in \Lambda\}$ such that, for an arbitrary topological linear space F and each continuous map $f : X \rightarrow F$, there exist $\lambda \in \Lambda$ and a linear embedding $\varphi : F_\lambda \rightarrow F$ such that $\varphi f_\lambda = f$.

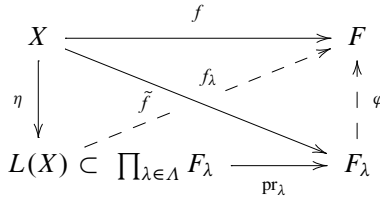
Indeed, for each cardinal $\tau \leq \text{card } X$, let \mathfrak{T}_τ be the topologies \mathcal{T} on \mathbb{R}_f^τ such that $(\mathbb{R}_f^\tau, \mathcal{T})$ is a topological linear space. Then, the desired collection is

$$\mathcal{F} = \bigcup_{\tau \leq \text{card } X} \bigcup_{\mathcal{T} \in \mathfrak{T}_\tau} \mathcal{C}(X, (\mathbb{R}_f^\tau, \mathcal{T})).$$

Consequently, for an arbitrary topological linear space F and each continuous map $f : X \rightarrow F$, let $\tau = \text{card } f(X) \leq \text{card } X$. The linear span F' of $f(X)$ is linearly homeomorphic to $(\mathbb{R}_f^\tau, \mathcal{T})$ for some $\mathcal{T} \in \mathfrak{T}_\tau$. Let $\psi : F' \rightarrow (\mathbb{R}_f^\tau, \mathcal{T})$ be a linear homeomorphism. Accordingly, we have $g = \psi f \in C(X, (\mathbb{R}_f^\tau, \mathcal{T}))$, and thus $f = \psi^{-1}g$.

The product space $\prod_{\lambda \in \Lambda} F_\lambda$ is a topological linear space. Let $\eta : X \rightarrow \prod_{\lambda \in \Lambda} F_\lambda$ be the map defined by $\eta(x) = (f_\lambda(x))_{\lambda \in \Lambda}$. We define $L(X)$ as the linear span of $\eta(X)$ in $\prod_{\lambda \in \Lambda} F_\lambda$. Then, $(L(X), \eta)$ satisfies the condition (*) in the above remark. In fact, for an arbitrary topological linear space F and each map $f : X \rightarrow F$, there exists $\lambda \in \Lambda$ and a linear embedding $\varphi : F_\lambda \rightarrow F$ such that $\varphi f_\lambda = f$. Consequently, we have a linear map $\tilde{f} = \varphi \text{pr}_\lambda | L(X) : L(X) \rightarrow F$ and

$$\tilde{f} \eta(x) = \varphi \text{pr}_\lambda \eta(x) = \varphi f_\lambda(x) = f(x) \text{ for every } x \in X.$$



Because $L(X)$ is generated by $\eta(X)$, a linear map $\tilde{f} : L(X) \rightarrow F$ is uniquely determined by the condition that $\tilde{f} \eta = f$. As observed in the above remark, η is an embedding, hence X can be identified with $\eta(X)$. Then, $L(X)$ satisfies (LE), i.e., $L(X)$ is the free topological linear space over X . □

Let X and Y be Tychonoff spaces. For each map $f : X \rightarrow Y$, we have a unique linear map $f_{\natural} : L(X) \rightarrow L(Y)$ that is an extension of f by (LE).

$$\begin{array}{ccc}
 X & \xrightarrow{f} & Y \\
 \cap & & \cap \\
 L(X) & \xrightarrow{f_{\natural}} & L(Y)
 \end{array}$$

This is functorial, i.e., $(gf)_{\natural} = g_{\natural} f_{\natural}$ for every pair of maps $f : X \rightarrow Y$ and $g : Y \rightarrow Z$, and $\text{id}_{L(X)} = (\text{id}_X)_{\natural}$. Accordingly, we have a covariant functor from the category of Tychonoff spaces into the category of topological linear spaces. Consequently, every homeomorphism $f : X \rightarrow Y$ extends to a linear homeomorphism $f_{\natural} : L(X) \rightarrow L(Y)$.

In Sect. 7.12, we will construct a metrizable linear space that is not an absolute extensor for metrizable spaces. The free topological linear space $L(X)$ over a compactum X has an important role in the construction. The topological and geometrical structures of $L(X)$ will be studied in Sect. 7.11.

Notes for Chap. 3

There are lots of good textbooks for studying topological linear spaces. The following classical book of Köthe is still a very good source on this subject. The textbook by Kelly and Namioka is also recommended by many people. Besides these two books, the textbook by Day is a good reference for normed linear spaces as is Valentine's book for convex sets. Concerning non-locally convex F -spaces and Roberts' example (a compact convex set with no extreme points), one can refer to the book by Kalton, Peck and Roberts.

- G. Köthe, *Topological Vector Spaces, I*, English edition, **GMW 159** (Springer-Verlag, New York, 1969)
- J.L. Kelly and I. Namioka, *Linear Topological Spaces*, Reprint edition, **GTM 36** (Springer-Verlag, New York, 1976)
- M.M. Day, *Normed Linear Spaces*, 3rd edition, **EMG 21** (Springer-Verlag, Berlin, 1973)
- F.A. Valentine, *Convex Sets* (McGraw-Hill Inc., 1964); Reprint of the 1964 original (R.E. Krieger Publ. Co., New York, 1976)
- N.J. Kalton, N.T. Peck and J.W. Roberts, *An F -space Sampler*, London Math. Soc. Lecture Note Ser. **89** (Cambridge Univ. Press, Cambridge, 1984)

For a systematic and comprehensive study on continuous selections, refer to the following book by Repovš and Semenov, which is written in instructive style.

- D. Repovš and P.V. Semenov, *Continuous Selections of Multivalued Mappings*, **MIA 455** (Kluwer Acad. Publ., Dordrecht, 1998)

In Theorem 3.6.4, the construction of a metric d from d_0 is due to Eidelheit and Mazur [1].

The results of Sect. 3.8 are contained in the first part of Michael's paper [2], which consists of three parts. For the finite-dimensional case, refer to the second and third parts of [2] (cf. [3]) and the book of Repovš and Semenov. The finite-dimensional case is deeply related with the concept discussed in Sect. 6.11 but will not be treated in this book. The 0-dimensional case will be treated in Sect. 7.2.

References

1. M. Eidelheit, S. Mazur, Eine Bemerkung über die Räume vom Typus (F). *Stud. Math.* **7**, 159–161 (1938)
2. E. Michael, Continuous selections, I. *Ann. Math.* **63**, 361–382 (1956); Continuous selections, II. *Ann. Math.* **64**, 562–580 (1956); Continuous selections, III. *Ann. Math.* **65**, 375–390 (1957)
3. E. Michael, A generalization of a theorem on continuous selections. *Proc. Am. Math. Soc.* **105**, 236–243 (1989)