Functional Analysis

Peter Philip*

Lecture Notes

Originally Created for the Class of Summer Semester 2017 at LMU Munich
Includes Subsequent Corrections and Revisions†

March 22, 2024

Contents

*E-Mail: philip@math.lmu.de

†Resources used in the preparation of this text include [Els07, Rud73, Wer11].

1 Topological Vector Spaces

1.1 Basic Definitions and Properties

As in Analysis I–III, we write $\mathbb{K}$ to denote $\mathbb{R}$ or $\mathbb{C}$.

The central topic of (linear) *Functional Analysis* is the investigation and representation of *continuous linear functionals*, i.e. of continuous linear functions $f : X \longrightarrow \mathbb{K}$, where X is a vector space over $\mathbb{K}$. To know what continuity of f means, we need to specify topologies on X and $\mathbb{K}$. On $\mathbb{K}$, we will always consider the standard topology (induced by $|\cdot|$), unless another topology is explicitly specified. While one will, in general, want to study many different vector spaces X with many different topologies $\mathcal{T}$, $\mathcal{T}$ should at least be compatible with the linear structure on X, giving rise to the following definition:

Definition 1.1. Let X be a vector space over $\mathbb{K}$ and let $\mathcal{T}$ be a topology on X. Then the topological space $(X, \mathcal{T})$ is called a *topological vector space* if, and only if, addition and scalar multiplication are continuous, i.e. if, and only if, the maps

$$+ : X \times X \longrightarrow X, \quad (x, y) \mapsto x + y, \tag{1.1a}$$
$$\cdot : \mathbb{K} \times X \longrightarrow X, \quad (\lambda, x) \mapsto \lambda x, \tag{1.1b}$$

are continuous (with respect to the respective product topology).

Example 1.2. (a) Every normed vector space $(X, \|\cdot\|)$ over $\mathbb{K}$ is a topological vector space: Let $(x_k)_{k\in\mathbb{N}}$ and $(y_k)_{k\in\mathbb{N}}$ be sequences in X with $\lim_{k\to\infty} x_k = x \in X$ and $\lim_{k\to\infty} y_k = y \in X$. Then $\lim_{k\to\infty}(x_k + y_k) = x + y$ by [Phi16b, (2.20a)], showing continuity of addition. Now let $(\lambda_k)_{k\in\mathbb{N}}$ be a sequence in $\mathbb{K}$ such that $\lim_{k\to\infty} \lambda_k = \lambda \in \mathbb{K}$. Then $(|\lambda_k|)_{k\in\mathbb{N}}$ is bounded by some $M \in \mathbb{R}_0^+$ and

$$\|\lambda_k x_k - \lambda x\| \leq \|\lambda_k x_k - \lambda_k x\| + \|\lambda_k x - \lambda x\| \leq M\,\|x_k - x\| + |\lambda_k - \lambda|\,\|x\| \to 0 \text{ for } k \to \infty,$$

showing $\lim_{k\to\infty}(\lambda_k x_k) = \lambda x$ and the continuity of scalar multiplication. We will see in Sec. 1.2 below that, for $\dim X < \infty$, the norm topology on X is the *only* T_1 topology on X that makes X into a topological vector space (but cf. (b) below).

(b) Let X be a vector space over $\mathbb{K}$. With the indiscrete topology, X is always a topological vector space (the continuity of addition and scalar multiplication is trivial). If $X \neq \{0\}$, then the indiscrete space is not T_1 and, hence, not metrizable (cf. [Phi16b, Sec. 3.1]). If $X \neq \{0\}$, then, with the discrete topology, X is never a topological vector space: While addition is continuous (since $X \times X$ is also discrete), scalar multiplication is not: Let $0 \neq x \in X$. Then, while $\{x\} \subseteq X$ is open, the preimage $P := (\cdot)^{-1}(\{x\})$ is not open in $\mathbb{K}\times X$: Let $(\lambda, y) \in P$, i.e. $\lambda y = x$. If P were open, then there had to be an open neighborhood O of λ such that $O \times \{y\} \subseteq P$, in contradiction to λ being the *unique* element of $\mathbb{K}$ such that $\lambda y = x$.

(c) Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. Then, for each $0 < p \leq \infty$, both $\mathcal{L}^p(\mu)$ and $L^p(\mu)$ are topological vector spaces: First, let $1 \leq p \leq \infty$. Then $L^p(\mu)$ is a special case of (a), and, for, $\mathcal{L}^p(\mu)$ one observes that the arguments of (a) still work if the norm is replaced by a seminorm (since seminormed spaces are still first countable, cf. [Phi16b, Th. 2.8]). If there exists a nonempty μ-null set, then $\mathcal{L}^p(\mu)$ is not T_1 (in particular, not metrizable), cf. [Phi17, Def. and Rem. 2.41]. Now consider $0 < p < 1$. We know from [Phi17, Def. and Rem. 2.41] that $\mathcal{L}^p(\mu)$ is a pseudometric space, where the pseudometric d_p is defined by

$$d_p : \mathcal{L}^p(\mu) \times \mathcal{L}^p(\mu) \longrightarrow \mathbb{R}_0^+, \quad d_p(f,g) := N_p^p(f-g), \quad N_p(f) := \left(\int_\Omega |f|^p \,\mathrm{d}\mu \right)^{1/p},$$

and $L^p(\mu)$ is the corresponding (factor) metric space. Like metric spaces, pseudometric spaces are first countable and we can show continuity using sequences according to [Phi16b, Th. 2.8]. Let $(f_k)_{k\in\mathbb{N}}$ and $(g_k)_{k\in\mathbb{N}}$ be sequences in $\mathcal{L}^p(\mu)$ with $\lim_{k\to\infty} f_k = f \in \mathcal{L}^p(\mu)$ and $\lim_{k\to\infty} g_k = g \in \mathcal{L}^p(\mu)$. Then

$$\begin{aligned} d_p(f_k + g_k, f + g) &= N_p^p\big(f_k + g_k - (f+g)\big) \overset{\text{[Phi17, (2.51a)]}}{\leq} N_p^p(f_k - f) + N_p^p(g_k - g) \\ &= d_p(f_k, f) + d_p(g_k, g) \to 0 \text{ for } k \to \infty, \end{aligned}$$

showing continuity of addition. Now let $(\lambda_k)_{k\in\mathbb{N}}$ be a sequence in $\mathbb{K}$ such that $\lim_{k\to\infty} \lambda_k = \lambda \in \mathbb{K}$. Then $(|\lambda_k|)_{k\in\mathbb{N}}$ is bounded by some $M \in \mathbb{R}_0^+$ and

$$\begin{aligned} d_p(\lambda_k f_k, \lambda f) &\leq d_p(\lambda_k f_k, \lambda_k f) + d_p(\lambda_k f, \lambda f) \\ &= \int_\Omega \big|\lambda_k f_k - \lambda_k f\big|^p \,\mathrm{d}\mu + \int_\Omega \big|\lambda_k f - \lambda f\big|^p \,\mathrm{d}\mu \\ &\leq M^p \, d_p(f_k, f) + |\lambda_k - \lambda|^p \, N_p^p(f) \to 0 \text{ for } k \to \infty, \end{aligned}$$

showing $\lim_{k\to\infty}(\lambda_k f_k) = \lambda f$ and the continuity of scalar multiplication. As for $p \geq 1$, if there exists a nonempty μ-null set, then $\mathcal{L}^p(\mu)$ is not T_1 (in particular, not metrizable), again cf. [Phi17, Def. and Rem. 2.41]. We will see in Ex. 1.11(b) that, for $0 < p < 1$, balls in $L^p([0,1], \mathcal{L}^1, \lambda^1)$, where $\mathcal{L}^1$ denotes the usual σ-algebra of Lebesgue-measurable sets and λ^1 denotes Lebesgue measure, are *not* convex. In particular, the metric d_p is not generated by any norm on $L^p([0,1], \mathcal{L}^1, \lambda^1)$.

(d) Consider $X := \mathbb{K}^{\mathbb{R}} = \mathcal{F}(\mathbb{R}, \mathbb{K})$, i.e. the set of functions $f : \mathbb{R} \longrightarrow \mathbb{K}$, with the product topology $\mathcal{T}$ (i.e. the topology of pointwise convergence). We know from [Phi16b, Ex. 1.53(c)] that $\mathcal{T}$ is not metrizable (however, $\mathcal{T}$ is T_2 by [Phi16b, Prop. 3.5(b)]). We show that $(X, \mathcal{T})$ is a topological vector space over $\mathbb{K}$: According to [Phi16b, Ex. 2.12(c)], we have to show that, for each $f, g : \mathbb{R} \longrightarrow \mathbb{K}$, each

$\lambda \in \mathbb{K}$, and each $t \in \mathbb{R}$, $(f,g) \mapsto f(t)+g(t)$ is continuous at (f,g), $(\lambda,f) \mapsto \lambda f(t)$ is continuous at (λ, f). Due to the continuity of the maps $+ : \mathbb{K} \times \mathbb{K} \longrightarrow \mathbb{K}$, $\cdot : \mathbb{K} \times \mathbb{K} \longrightarrow \mathbb{K}$, given $\epsilon \in \mathbb{R}^+$, there exist neighborhoods U_f, U_g, U_λ of $f(t)$, $g(t)$, and λ, respectively, such that

$$\forall_{(z,w)\in U_f\times U_g} \quad z+w \in B_\epsilon\big(f(t)+g(t)\big), \qquad \forall_{(\mu,z)\in U_\lambda\times U_f} \quad \mu z \in B_\epsilon\big(\lambda\, f(t)\big).$$

Letting $V_f := \pi_t^{-1}(U_f)$, $V_g := \pi_t^{-1}(U_g)$, we obtain

$$\forall_{(\tilde f,\tilde g)\in V_f\times V_g} \quad \tilde f(t)+\tilde g(t) \in B_\epsilon\big(f(t)+g(t)\big),$$

proving continuity of addition on X. We also have

$$\forall_{(\mu,\tilde f)\in U_\lambda\times V_f} \quad \mu\tilde f(t) \in B_\epsilon\big(\lambda\, f(t)\big),$$

proving continuity of scalar multiplication.

—

In certain situations, the following notation has already been used in both Analysis and Linear Algebra:

Notation 1.3. Let X be a vector space over $\mathbb{K}$, let $\mathcal{A} \subseteq \mathcal{P}(X)$ (where $\mathcal{P}(X)$ denotes the power set of X), $A, B \subseteq X$, $x \in X$, and $\lambda \in \mathbb{K}$. Define

$$x + A := \{x+a :\ a \in A\}, \tag{1.2a}$$
$$A + B := \{a+b :\ a \in A,\ b \in B\}, \tag{1.2b}$$
$$\lambda A := \{\lambda a :\ a \in A\}, \tag{1.2c}$$
$$x + \mathcal{A} := \{x+A :\ A \in \mathcal{A}\}. \tag{1.2d}$$

Note that, in general, the familiar arithmetic laws do *not* hold for set arithmetic: For example, if $X \neq \{0\}$, then $X - X = X \neq \{0\}$; if $0 \neq x \in X$, $A := \{-x, x\}$, then $0 \in A + A$, but $0 \notin 2A$, i.e. $2A \neq A + A$.

Proposition 1.4. *Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{K}$.*

(a) *For each $a \in X$ and each $\lambda \in \mathbb{K} \setminus \{0\}$, the maps*

$$T_a, M_\lambda :\ X \longrightarrow X, \quad T_a(x) := x + a, \quad M_\lambda(x) := \lambda x, \tag{1.3}$$

are homeomorphisms, where $T_a^{-1} = T_{-a}$, $M_\lambda^{-1} = M_{\lambda^{-1}}$.

(b) *$\mathcal{T}$ is both* translation-invariant *and* scaling-invariant*, i.e. the following holds for each $O \subseteq X$:*

$$O \text{ open} \quad \Leftrightarrow \quad \left(\forall_{a \in X} \; O + a \text{ open} \right) \quad \Leftrightarrow \quad \left(\forall_{\lambda \in \mathbb{K} \setminus \{0\}} \; \lambda O \text{ open} \right).$$

(c) *Let $x, a \in X$, $U \subseteq X$, $\lambda \in \mathbb{K} \setminus \{0\}$. Then U is a neighborhood of x if, and only if, $a + U$ is a neighborhood of $a + x$, and if, and only if, λU is a neighborhood of λx.*

(d) *Let $x, a \in X$, $\mathcal{B} \subseteq \mathcal{P}(X)$. Then $\mathcal{B}$ is a local base at x (see* [Phi16b, Def. 1.38]*) if, and only if, $a + \mathcal{B}$ is a local base at $a + x$.*

Proof. (a): T_a and M_λ are clearly bijective with the provided inverses. The continuity of the maps and their inverses is due to (1.1).

(b) is immediate from (a).

(c) follows from (a) and (b), since $T_a(U) = a + U$ and $M_\lambda(U) = \lambda U$. Now (d) is another immediate consequence. ■

Proposition 1.5. *Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{K}$.*

(a) *If $W \in \mathcal{U}(0)$, then there exists an open $U \in \mathcal{U}(0)$ such that U satisfies the following two properties:*

$$U = -U, \tag{1.4a}$$

$$U + U \subseteq W \tag{1.4b}$$

(here, and in the following, $\mathcal{U}(x)$ denotes the set of all neighborhoods of $x \in X$ (cf. [Phi16b, Def. 1.1]*).*

(b) *Disjoint sets $A, K \subseteq X$, where A is closed and K is compact, can be separated by disjoint open sets (in particular, topological vector spaces are always T_3, cf.* [Phi16b, Def. 3.1(c)]*):*

$$\forall_{\substack{A,K \subseteq X, \\ A \text{ closed}, \\ K \text{ compact}}} \left(A \cap K = \emptyset \;\Rightarrow\; \exists_{O_1, O_2 \in \mathcal{T}} \left(A \subseteq O_1 \;\wedge\; K \subseteq O_2 \;\wedge\; O_1 \cap O_2 = \emptyset \right) \right). \tag{1.5a}$$

The following reformulation uses the linear structure of X:

$$\forall_{\substack{A,K \subseteq X, \\ A \text{ closed}, \\ K \text{ compact}}} \left(A \cap K = \emptyset \;\Rightarrow\; \exists_{U \in \mathcal{U}(0)} \left((A + U) \cap (K + U) = \emptyset \right) \right). \tag{1.5b}$$

(c) *Every neighborhood contains a closed neighborhood:*

$$\forall_{x\in X} \quad \forall_{U\in\mathcal{U}(x)} \quad \exists_{A\in\mathcal{U}(x)} \quad \big(x \in A \subseteq U \ \wedge \ A \text{ closed}\big).$$

(d) *If $(X,\mathcal{T})$ is T_1, then $(X,\mathcal{T})$ is regular (i.e. T_1 and T_3). In particular, $(X,\mathcal{T})$ is then also T_2 (i.e. Hausdorff).*

Proof. (a): Since addition is continuous and $0+0=0$, given $W \in \mathcal{U}(0)$, there exist open $U_1, U_2 \in \mathcal{U}(0)$ such that $U_1 + U_2 \subseteq W$. According to Prop. 1.4, $-U_1, -U_2$ are open and $-U_1, -U_2 \in \mathcal{U}(0)$ as well. Thus, $U := U_1 \cap U_2 \cap (-U_1) \cap (-U_2) \in \mathcal{U}(0)$, U open, $U + U \subseteq U_1 + U_2 \subseteq W$, and $x \in U$ if, and only if, $-x \in U$, showing $U = -U$.

(b),(c): Exercise.

(d) is immediate from (c) (and since $T_1 + T_3$ implies T_2, cf. [Phi16b, Lem. 3.2(b)]). ∎

Definition 1.6. Let X be a vector space over $\mathbb{K}$, $A \subseteq X$.

(a) A is called *convex* if, and only if,

$$\forall_{a,b\in A} \quad \forall_{0\leq\lambda\leq 1} \quad \lambda a + (1-\lambda)b \in A$$

(i.e. if, and only if, for each $0 \leq \lambda \leq 1$, $\lambda A + (1-\lambda)A \subseteq A$ in terms of Not. 1.3).

(b) A is called *balanced* if, and only if,

$$\forall_{a\in A} \quad \forall_{|\lambda|\leq 1} \quad \lambda a \in A$$

(i.e. if, and only if, for each $\lambda \in \mathbb{K}$ with $|\lambda| \leq 1$, $\lambda A \subseteq A$ in terms of Not. 1.3).

Example 1.7. If $0 \neq x \in X$ (X vector space over $\mathbb{K}$), then $\{x\}$ is convex, but not balanced. The set $A := ([-1,1] \times \{0\}) \cup (\{0\} \times [-1,1]) \subseteq \mathbb{R}^2$ is balanced, but not convex. Moreover,

$$\begin{aligned}
\{A \subseteq \mathbb{R} : A \text{ balanced}\} &= \big\{\mathbb{R}\big\} \cup \big\{]-r,r[: r \in \mathbb{R}^+\big\} \cup \big\{[-r,r] : r \in \mathbb{R}_0^+\big\},\\
\{A \subseteq \mathbb{C} : A \text{ balanced}\} &= \big\{\mathbb{C}\big\} \cup \big\{B_r(0) : r \in \mathbb{R}^+\big\} \cup \big\{\overline{B}_r(0) : r \in \mathbb{R}_0^+\big\}.
\end{aligned}$$

Proposition 1.8. *Let X be a vector space over $\mathbb{K}$. Let $(A_i)_{i\in I}$ be a family of subsets of X; $A, B \subseteq X$.*

(a) *If each A_i is convex, then so is $C := \bigcap_{i\in I} A_i$. If A, B are convex, then so are $A+B$ and αA for each $\alpha \in \mathbb{K}$.*

(b) *If each A_i is balanced, then so are $C := \bigcap_{i\in I} A_i$ and $D := \bigcup_{i\in I} A_i$. If A, B are balanced, then so are $A + B$ and αA for each $\alpha \in \mathbb{K}$.*

(c) *If A is balanced and $0 \leq s \leq t$, then $sA \subseteq tA$.*

(d) *If X is a Cartesian product $X = \prod_{i\in I} X_i$ of vector spaces X_i over $\mathbb{K}$ and $B_i \subseteq X_i$ is balanced (resp. convex), then $B := \prod_{i\in I} B_i$ is also balanced (resp. convex).*

Proof. Exercise. ■

Definition 1.9. Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{K}$.

(a) Let $\mathcal{B} \subseteq \mathcal{P}(X)$ be a local base at $x \in X$. We call $\mathcal{B}$ *convex* (resp. *balanced*) if, and only if, each $B \in \mathcal{B}$ is convex (resp. balanced).

(b) We call $(X, \mathcal{T})$ *locally convex* if, and only if, 0 has a convex local base (then, by Prop. 1.4(d), every $x \in X$ has a convex local base).

(c) A set $B \subseteq X$ is called *bounded* if, and only if,

$$\underset{U\in\mathcal{U}(0)}{\forall} \quad \underset{s\in\mathbb{R}^+}{\exists} \quad B \subseteq s\,U.$$

(d) We call $(X, \mathcal{T})$ *locally bounded* if, and only if, 0 has a bounded neighborhood.

Proposition 1.10. *Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{K}$. Let $A, B \subseteq X$.*

(a) *One has*

$$\overline{A} = \bigcap_{U\in\mathcal{U}(0)} (A + U).$$

(b) $\overline{A} + \overline{B} \subseteq \overline{A + B}$.

(c) *If A is a vector subspace of X, then so is $\overline{A}$.*

(d) *If B is open, then so is $A + B$.*

(e) *If A is convex, then so are $\overline{A}$ and A°.*

(f) *If A is balanced, then so is $\overline{A}$. If A is balanced with $0 \in A^\circ$, then A° is balanced as well.*

(g) *If A is bounded, then so is $\overline{A}$.*

Proof. (a): Given $x \in X$, according to Prop. 1.4(c), $U \subseteq X$ is a neighborhood of 0 if, and only if, $x + U$ is a neighborhood of x. Thus,

$$x \in \overline{A} \quad \Leftrightarrow \quad \left(\forall_{U \in \mathcal{U}(0)} \ (x+U) \cap A \neq \emptyset \right) \quad \Leftrightarrow \quad \left(\forall_{U \in \mathcal{U}(0)} \ x \in A - U \right),$$

where the last equivalence holds since $a \in (x+U) \cap A$ if, and only if, there exists $u \in U$ such that $a = x + u \in A$. Using Prop. 1.4(c) again, U is a neighborhood of 0 if, and only if, $-U$ is a neighborhood of 0, showing the above equivalences prove (a).

(b): Let $a \in \overline{A}$, $b \in \overline{B}$, $U \in \mathcal{U}(a+b)$. By the continuity of addition, there exist $U_1 \in \mathcal{U}(a)$ and $U_2 \in \mathcal{U}(b)$ such that $U_1 + U_2 \subseteq U$. Since $a \in \overline{A}$, $b \in \overline{B}$, there exist $x \in A \cap U_1$ and $y \in B \cap U_2$. Then $x+y \in (A+B) \cap (U_1+U_2) \subseteq (A+B) \cap U \neq \emptyset$, showing $a+b \in \overline{A+B}$.

(c): Let $a, b \in \overline{A}$. Then

$$a + b \in \overline{A} + \overline{A} \overset{\text{(b)}}{\subseteq} \overline{A+A} \overset{A \text{ subsp.}}{=} \overline{A}. \tag{1.6a}$$

According to Prop. 1.4(a),

$$\forall_{\lambda \in \mathbb{K} \setminus \{0\}} \quad \lambda \overline{A} = M_\lambda(\overline{A}) = \overline{M_\lambda(A)} = \overline{\lambda A}. \tag{1.6b}$$

As we also have $0\overline{A} = \{0\} \subseteq \overline{\{0\}} \subseteq \overline{A}$, since $0 \in A$ if A is a subspace (note that $\{0\} \subsetneq \overline{\{0\}}$ if $(X, \mathcal{T})$ is not T_1), $\overline{A}$ is a subspace of X.

(d): Let B be open, $a \in A$, $b \in B$. Then $a + B$ is an open neighborhood of $a + b$ and $a + B \subseteq A + B$, showing $a + b$ to be an interior point of $A + B$, i.e. $A + B$ is open.

(e): Let A be convex. Then

$$\forall_{0<\lambda<1} \quad \lambda\overline{A} + (1-\lambda)\overline{A} \overset{\text{(1.6b),(b)}}{\subseteq} \overline{\lambda A + (1-\lambda)A} \overset{A \text{ convex}}{=} \overline{A},$$

showing $\overline{A}$ to be convex. Furthermore, since $A^\circ \subseteq A$, we have

$$\forall_{0<\lambda<1} \quad A_\lambda := \lambda A^\circ + (1-\lambda)A^\circ \subseteq \lambda A + (1-\lambda)A \overset{A \text{ convex}}{=} A.$$

Since A_λ is open by (d) and Prop. 1.4(a), and since A° is the union of all open subsets of A, we obtain $A_\lambda \subseteq A^\circ$, showing A° to be convex.

(f): Let A be balanced. As in the proof of (c), we obtain $\lambda\overline{A} \subseteq \overline{\lambda A}$ for each $\lambda \in \mathbb{K}$. Thus,

$$\forall_{|\lambda| \leq 1} \quad \lambda\overline{A} \subseteq \overline{\lambda A} \overset{A \text{ bal.}}{\subseteq} \overline{A},$$

showing $\overline{A}$ to be balanced. According to Prop. 1.4(a),

$$\forall_{\lambda\in\mathbb{K}\setminus\{0\}} \quad \lambda A^\circ = M_\lambda(A^\circ) = (M_\lambda(A))^\circ = (\lambda A)^\circ,$$

implying

$$\forall_{0<|\lambda|\leq 1} \quad \lambda A^\circ = (\lambda A)^\circ \overset{A \text{ bal.}}{\subseteq} A^\circ.$$

Since $0A^\circ = \{0\} \subseteq A^\circ$ holds by hypothesis, A° is balanced.

(g): Let A be bounded and $U \in \mathcal{U}(0)$. According to Prop. 1.5(c), there exists $C \in \mathcal{U}(0)$ such that $C \subseteq U$ and C is closed. Since A is bounded, there exists $s \in \mathbb{R}^+$ such that $A \subseteq sC$, where sC is still closed. Thus, $\overline{A} \subseteq sC$, showing $\overline{A}$ to be bounded. ■

It is an exercise to find counterexamples that show that, in general, Prop. 1.10(b) does not hold with equality (there exist examples with $(X, \mathcal{T})$ being $\mathbb{R}$ with the norm topology) and that, in general, the second part of Prop. 1.10(f) becomes false if $0 \notin A^\circ$ is omitted from the hypothesis.

Example 1.11. (a) If $(X, \|\cdot\|)$ is a normed vector space over $\mathbb{K}$, then it is both locally bounded and locally convex: Each ball $B_r(x)$, $x \in X$, $r \in \mathbb{R}^+$, is both convex and bounded: Indeed, if $a, b \in B_r(x)$ and $\lambda \in [0,1]$, then

$$\|\lambda a+(1-\lambda)b-x\| = \|\lambda a-\lambda x+(1-\lambda)b-(1-\lambda)x\| \leq \lambda\|a-x\|+(1-\lambda)\|b-x\| < r,$$

showing $\lambda a + (1-\lambda)b \in B_r(x)$. Now let $\epsilon \in \mathbb{R}^+$, $y \in B_r(x)$, $\alpha := \|x\|$. Then $\|y\| \leq \|y-x\| + \|x\| < r + \alpha$, showing $y \in B_{r+\alpha}(0)$. Thus $\frac{\epsilon}{r+\alpha} y \in B_\epsilon(0)$ and $y \in \frac{r+\alpha}{\epsilon} B_\epsilon(0)$, showing $B_r(x)$ to be bounded in the sense of Def. 1.9(c). As a caveat, it is pointed out that, in general, the topology $\mathcal{T}$ of a topological vector space $(X, \mathcal{T})$ can be induced by a metric d on X without the corresponding metric balls being convex (see (b) below), balanced (see (d) below), or even bounded in the sense of Def. 1.9(c) (in Ex. 1.43 below, we will construct topological vector spaces that are metrizable, but not locally bounded).

(b) We come back to the spaces $X := L^p([0,1], \mathcal{L}^1, \lambda^1)$, $0 < p < 1$, with the metric d_p (cf. Ex. 1.2(c)). It is an exercise to show (X, d_p) is locally bounded, but not locally convex (and, actually, $\emptyset$ and X are the *only* convex open subsets of X). As mentioned earlier, a main goal of Functional Analysis is the representation of continuous linear functionals. In the present case, it turns out that $A \equiv 0$ is the *only* continuous linear functional on X: Indeed, let $(Y, \mathcal{T})$ be an arbitrary locally convex topological vector space that is T_1 (for example, $Y = \mathbb{K}$) and let $A : X \longrightarrow Y$ be linear and continuous. Let $\mathcal{B}$ be a local base of $\mathcal{T}$ at 0, consisting of convex open sets. If $C \in \mathcal{B}$, then $A^{-1}(C)$ is convex, open, and nonempty, i.e. $A^{-1}(C) = X$. If

$y \in Y$, $y \neq 0$, then, since $(Y, \mathcal{T})$ is T_1, there exists $U \in \mathcal{U}(0)$ such that $y \notin U$. Then there is $C \in \mathcal{B}$ such that $C \subseteq U$. Since $A^{-1}(C) = X$, $y \notin A(X)$, showing $A \equiv 0$.

(c) The topological vector space $(\mathbb{K}^{\mathbb{R}}, \mathcal{T})$ (where $\mathcal{T}$ is the product topology) is locally convex, but not locally bounded: The set

$$\left\{ \bigcap_{j \in J} \pi_j^{-1}(B_{\epsilon_j}(0)) : \; J \subseteq \mathbb{R},\, 0 < \#J < \infty,\, \underset{j \in J}{\forall}\, \epsilon_j \in \mathbb{R}^+ \right\}$$

constitutes a local base at 0 and, by Prop. 1.8(d), each element of this local base is convex. We now show that 0 does not have a bounded neighborhood: Indeed, if $U \in \mathcal{U}(0)$, then U contains a set from the above local base, say $B := \bigcap_{j \in J} \pi_j^{-1}(B_{\epsilon_j}(0))$, where J is a nonempty finite subset of $\mathbb{R}$ and $\epsilon_j > 0$ for each $j \in J$. Let $\alpha \in \mathbb{R} \setminus J$. Then $V := \pi_\alpha^{-1}(B_1(0))$ is another neighborhood of 0. However,

$$\underset{s \in \mathbb{R}^+}{\forall} \quad sV = \pi_\alpha^{-1}(B_s(0)),$$

showing that $B \not\subseteq sV$ and that neither B nor U is bounded.

(d) If we consider $\mathbb{R}^2$ as a vector space over $\mathbb{R}$ and $\|\cdot\|$ is some norm on $\mathbb{R}^2$, then the balls $B_r(0)$ and $\overline{B}_r(0)$, $r \in \mathbb{R}^+$, with respect to this norm are $\mathbb{R}$-balanced. However, if we consider $\mathbb{R}^2 = \mathbb{C}$ as a vector space over $\mathbb{C}$, then each norm on $\mathbb{R}^2$ still induces a metric and the standard topology on $\mathbb{C}$. However, the balls are not necessarily $\mathbb{C}$-balanced: For example, consider $\mathbb{R}^2$ with the max-norm: $d(z, w) := \max\{|\operatorname{Re}(z-w)|, |\operatorname{Im}(z-w)|\}$. Then $1 + i \in \overline{B}_1(0)$, $|1+i| = \sqrt{2}$, i.e. $1 + i = \sqrt{2}\zeta$ with $|\zeta| = 1$. Thus, $\sqrt{2} = (1+i)\zeta^{-1} \notin \overline{B}_1(0)$, showing that $\overline{B}_1(0)$ is not $\mathbb{C}$-balanced.

Proposition 1.12. *Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{K}$.*

(a) *If $U \in \mathcal{U}(0)$, then there exists $B \in \mathcal{U}(0)$ such that $0 \in B \subseteq U$ and B is balanced and open. In particular, $(X, \mathcal{T})$ has a balanced local base at 0.*

(b) *If $U \in \mathcal{U}(0)$ is convex, then there exists $C \in \mathcal{U}(0)$ such that $0 \in C \subseteq U$ and C is convex, balanced, and open. In particular, if $(X, \mathcal{T})$ is locally convex, then $(X, \mathcal{T})$ has a balanced convex local base at 0.*

(c) *$B \subseteq X$ is bounded if, and only if,*

$$\underset{U \in \mathcal{U}(0)}{\forall} \quad \underset{s \in \mathbb{R}^+}{\exists} \quad \underset{t \geq s}{\forall} \quad B \subseteq t\,U.$$

(d) *$A \subseteq X$ is bounded if, and only if, every countable subset of A is bounded.*

(e) *If $A, B \subseteq X$ are bounded, then so are $A \cup B$ and $A + B$.*

(f) *If $U \in \mathcal{U}(0)$ and $(r_k)_{k\in\mathbb{N}}$ is a sequence in $\mathbb{R}^+$ such that $\lim_{k\to\infty} r_k = \infty$, then*

$$X = \bigcup_{k\in\mathbb{N}} (r_k U).$$

(g) *If $K \subseteq X$ is compact, then K is bounded.*

(h) *$(X, \mathcal{T})$ is locally bounded if, and only if,*

$$\underset{x\in X}{\forall} \quad \underset{U\in\mathcal{U}(x)}{\exists} \quad U \text{ bounded.}$$

(i) *If $U \in \mathcal{U}(0)$ is bounded and $(r_k)_{k\in\mathbb{N}}$ is a sequence in $\mathbb{R}^+$ such that $\lim_{k\to\infty} r_k = 0$, then $\mathcal{B} := \{r_k U : k \in \mathbb{N}\}$ is a local base at 0.*

Proof. (a): Since scalar multiplication is continuous and $0 \cdot 0 = 0$, given $U \in \mathcal{U}(0)$, there exists $\delta > 0$ and $V \in \mathcal{U}(0)$ open such that $\alpha V \subseteq U$ for each $\alpha \in B_\delta(0) = \{\lambda \in \mathbb{K} : |\lambda| < \delta\}$. Set $B := \bigcup_{\alpha\in B_\delta(0)}(\alpha V)$. Then, clearly, B is open, $B \in \mathcal{U}(0)$, and $B \subseteq U$. Moreover, B is balanced: Let $x \in B$, $\lambda \in \mathbb{K}$, $|\lambda| \leq 1$. Then $x = \alpha v$ with $\alpha \in B_\delta(0)$ and $v \in V$. Then $|\lambda\alpha| \leq |\alpha| < \delta$, showing $\lambda x = \lambda\alpha v \in B$.

(b): Let $U \in \mathcal{U}(0)$ be convex. Set

$$V := \bigcap_{|\lambda|=1} (\lambda U).$$

Then V is convex by Prop. 1.8(a). According to (a), there exists $B \in \mathcal{U}(0)$ such that $B \subseteq U$ and B is balanced and open. Since B is balanced, $B \subseteq V$, showing $V \in \mathcal{U}(0)$ as well as $C := V^\circ \in \mathcal{U}(0)$. Since C is open and convex (by Prop. 1.10(e)), it only remains to show C is balanced. By Prop. 1.10(f), it suffices to show V is balanced. Given $\alpha \in \mathbb{K}$ with $|\alpha| \leq 1$, we write $\alpha = r\zeta$ with $r = |\alpha|$ and $|\zeta| = 1$. Then

$$\alpha V = r\zeta V = \bigcap_{|\lambda|=1} (r\zeta\lambda U) = \bigcap_{|\lambda|=1} (r\lambda U) \subseteq \bigcap_{|\lambda|=1} (\lambda U) = V,$$

where the inclusion holds, as $0 \in \lambda U$, λU convex, i.e. $rx = (1-r)0 + rx \in \lambda U$ for each $x \in \lambda U$. We have, thus, shown V to be balanced.

(c): That the condition in (c) implies B to be bounded is immediate from Def. 1.9(c). Conversely, let B be bounded and $U \in \mathcal{U}(0)$. According to (a), U contains a balanced neighborhood V of 0. Since B is bounded, there exists $s \in \mathbb{R}^+$ such that $B \subseteq sV \subseteq sU$.

Let $t \geq s$. Since V is balanced, Prop. 1.8(c) implies $B \subseteq sV \subseteq tV \subseteq tU$, proving B to satisfy the condition in (c).

(d): If A is bounded, then every subset (in particular, every countable subset) is bounded. Conversely, assume A to be unbounded. We construct an unbounded countable subset: As A is not bounded,

$$\underset{U\in\mathcal{U}(0)}{\exists} \quad \underset{n\in\mathbb{N}}{\forall} \quad \underset{a_n\in A}{\exists} \quad a_n \notin nU.$$

Then $\tilde{A} := \{a_n : n \in \mathbb{N}\}$ is a countable subset of A and $\tilde{A}$ is unbounded: If $s \in \mathbb{R}^+$ and $n \in \mathbb{N}$, $n > s$, then $a_n \notin nU$, i.e. $\tilde{A} \not\subseteq nU$, i.e. $\tilde{A}$ is unbounded according to (c).

(e): Let A, B be bounded. Given $U \in \mathcal{U}(0)$, according to (c),

$$\underset{s_1,s_2\in\mathbb{R}^+}{\exists} \quad \underset{t\geq\max\{s_1,s_2\}}{\forall} \quad A \cup B \subseteq t\,U,$$

showing $A \cup B$ to be bounded. Next, we use Prop. 1.5 to obtain $V \in \mathcal{U}(0)$ such that $V + V \subseteq U$. By (a), we may assume V to be balanced as well. Now choose $s_1, s_2 \in \mathbb{R}^+$ such that $A \subseteq s_1 V$, $B \subseteq s_2 V$. Then, for $s := \max\{s_1, s_2\}$,

$$A + B \subseteq sV + sV \subseteq sU,$$

showing $A + B$ to be bounded.

(f): Fix $x \in X$. Since $\lambda \mapsto \lambda x$ is continuous, $V := \{\alpha \in \mathbb{K} : \alpha x \in U\}$ is a neighborhood of $0 \in \mathbb{K}$. Thus, there exists $N \in \mathbb{N}$ such that $r_k^{-1} \in V$ for each $k > N$. In consequence, for $k > N$, $r_k^{-1} x \in U$ and $x \in r_k U$.

(g): Let K be compact and $U \in \mathcal{U}(0)$. Moreover, let $B \subseteq U$ be a balanced open neighborhood of 0. Then, by (f) $(nB)_{n\in\mathbb{N}}$ is an open cover of K. Thus, since K is compact and B is balanced, there exists $n_0 \in \mathbb{N}$ such that $K \subseteq n_0 B \subseteq n_0 U$, showing K to be bounded.

(h): If $U \in \mathcal{U}(0)$ is bounded and $x \in X$, then $x + U \in \mathcal{U}(x)$ and $x + U$ is bounded by (e) (since $\{x\}$ is compact and, thus, bounded by (g)).

(i): Let $U \in \mathcal{U}(0)$ be bounded, $V \in \mathcal{U}(0)$. Then there exists $s \in \mathbb{R}^+$ such that $U \subseteq tV$ for each $t \geq s$. Choose $n \in \mathbb{N}$ such that $sr_n \leq 1$. Then $U \subseteq r_n^{-1} V$, i.e. $r_n U \subseteq V$, showing $\mathcal{B} = \{r_k\,U : k \in \mathbb{N}\}$ to be a local base at 0. ■

We conclude this section with some basic properties of continuous linear maps between topological vector spaces ([Phi16b, Th. 2.22]):

Theorem 1.13. *For a $\mathbb{K}$-linear function $A : X \longrightarrow Y$ between topological vector spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ over $\mathbb{K}$, the following statements are equivalent:*

(i) *A is continuous.*

(ii) *There exists $\xi \in X$ such that A is continuous in ξ.*

(iii) *A is* uniformly *continuous, i.e.*

$$\underset{U\in\mathcal{U}(0)\subseteq\mathcal{P}(Y)}{\forall} \quad \underset{V\in\mathcal{U}(0)\subseteq\mathcal{P}(X)}{\exists} \quad \underset{x,y\in X}{\forall} \quad \Big(y - x \in V \;\Rightarrow\; Ay - Ax \in U\Big).$$

Proof. (i) trivially implies (ii).
"(ii) $\Rightarrow$ (iii)": Let A be continuous in $\xi \in X$. Then

$$f : X \longrightarrow Y, \quad f(x) := A(x + \xi) - A(\xi),$$

is continuous in 0 with $f(0) = 0$. Let $U \in \mathcal{U}(0)$. Then there exists $V \in \mathcal{U}(0)$ such that $f(V) \subseteq U$. Thus, if $x, y \in X$ are such that $y - x \in V$, then

$$Ay - Ax = A(y - x + \xi) - A(\xi) = f(y - x) \in U,$$

proving (iii).
"(iii) $\Rightarrow$ (i)": Let $x \in X$. We show A is continuous at x: If $W \in \mathcal{U}(Ax)$, then $U := W - Ax \in \mathcal{U}(0)$ and $W = Ax + U$. Choose $V \in \mathcal{U}(0)$ according to (iii). Then $x+V \in \mathcal{U}(x)$ and if $y \in x+V$, then $y-x \in V$. Thus, $Ay-Ax \in U$ and $Ay \in Ax+U = W$, showing A to be continuous at x. ■

The following notions of *kernel* and *image* of a linear map $A : X \longrightarrow Y$ between vector spaces are familiar from Linear Algebra:

$$\ker A := A^{-1}\{0\} = \{x \in X : A(x) = 0\}, \tag{1.7a}$$

$$\operatorname{im} A := A(X) = \{A(x) : x \in X\}. \tag{1.7b}$$

Theorem 1.14. *Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{K}$. For a $\mathbb{K}$-linear functional $A : X \longrightarrow \mathbb{K}$ the following statements are equivalent:*

(i) *A is continuous.*

(ii) $\ker A$ *is closed.*

(iii) $\ker A = X$ *or* $\ker A$ *is not dense in X.*

(iv) *There exists $U \in \mathcal{U}(0)$ such that $A(U)$ is bounded.*

Proof. "(i)⇒(ii)": If A is continuous, then preimages of closes sets are closed. Since $\{0\} \subseteq \mathbb{K}$ is closed, so is $\ker A = A^{-1}\{0\}$.

"(ii)⇒(iii)": If $X \neq \ker A$ and $\ker A$ is closed, then $\overline{\ker A} = \ker A \neq X$, showing $\ker A$ is not dense in X.

"(iii)⇒(iv)": If $\ker A = X$, then $A(X) = \{0\}$, which is bounded. Now suppose $\ker A$ is not dense in X. Then $O := X \setminus \overline{\ker A}$ is nonempty and open. Let $x \in O$. Since O is open, there exists $U \in \mathcal{U}(0)$ balanced, such that

$$(x + U) \cap \ker A = \emptyset. \tag{1.8}$$

Since A is linear, $A(U) \subseteq \mathbb{K}$ is a also balanced. If $A(U)$ is bounded, then (iv) holds. If $A(U)$ is unbounded, then $A(U) = \mathbb{K}$ (since $A(U)$ is balanced). Thus, in this case, there exists $y \in U$ such that $Ay = -Ax$, i.e. $x + y \in \ker A \cap (x + U)$ in contradiction to (1.8).

"(iv)⇒(i)": By Th. 1.13, it suffices to show A is continuous at 0. Given $\epsilon > 0$, we have to find $V \in \mathcal{U}(0)$ such that $A(V) \subseteq B_\epsilon(0) \subseteq \mathbb{K}$. Using (iv), we know $A(U) \subseteq B_M(0)$ for some $U \in \mathcal{U}(0)$ and some $M > 0$. If $V := \frac{\epsilon}{M} U$, then, for each $\frac{\epsilon}{M} u$, $u \in U$, we have $A(\frac{\epsilon}{M} u) = \frac{\epsilon}{M} A(u) \in B_\epsilon(0)$, since $|A(u)| < M$. Thus, $A(V) \subseteq B_\epsilon(0)$ as needed. ■

1.2 Finite-Dimensional Spaces

In [Phi16b, Th. 3.24], we already stated the important result that the closed unit ball in a normed vector space X is compact if, and only if, X is finite-dimensional. In the present section, we will obtain this again as a corollary to the more general statement that a T_1 topological vector space has finite dimension if, and only if, it is locally compact (cf. Th. 1.19 below). In the process, we will prove in Th. 1.16(b) that finite-dimensional subspaces of T_1 topological vector spaces are always closed (again in generalization of the corresponding result for normed spaces, cf. [Phi16b, Th. E.2]). The remaining main result of this section is that the norm topology on $\mathbb{K}^n$ is the *only* T_1 topology that makes $\mathbb{K}^n$ into a topological vector space (see Cor. 1.17).

Lemma 1.15. *Let $(Y, \mathcal{T})$ be a topological vector space over $\mathbb{K}$. If $A : \mathbb{K}^n \longrightarrow Y$ is $\mathbb{K}$-linear, then A is continuous.*

Proof. For each $i \in \{1, \dots, n\}$, let $a_i := A(e_i)$ be the image of the standard unit vector $e_i \in \mathbb{K}^n$. Then

$$\underset{z=(z_1,\dots,z_n)\in\mathbb{K}^n}{\forall} \quad A(z) = \sum_{i=1}^n z_i a_i = \sum_{i=1}^n \pi_i(z) a_i,$$

and, since the projections $\pi_i : \mathbb{K}^n \longrightarrow \mathbb{K}$ are continuous as well as the constant functions $z \mapsto a_i$ and addition and scalar multiplication, so is A. ■

Theorem 1.16. *Let $(X, \mathcal{T})$ be a T_1 topological vector space over $\mathbb{K}$ and let $Y \subseteq X$ be a finite-dimensional subspace,* $\dim Y = n \in \mathbb{N}$. *Then the following statements hold true:*

(a) *If $A : \mathbb{K}^n \longrightarrow Y$ is a linear isomorphism, then A is also a homeomorphism.*

(b) *Y is closed.*

Proof. (a): The linear isomorphism A is continuous by Lem. 1.15. Let $S := S_1(0) = \{z \in \mathbb{K}^n : \|z\|_2 = 1\}$ be the unit sphere in $\mathbb{K}^n$. Then S is compact and, by [Phi16b, Th. 3.18], so is $K := A(S)$. Since $A(0) = 0$ and A is injective, $0 \notin K$. Since $(X, \mathcal{T})$ is T_2 by Prop. 1.5(d), the compact set K is closed by [Phi16b, Prop. 3.14(b)]. Thus, as $(X, \mathcal{T})$ is T_3 by Prop. 1.5(b), there exists an open $O \in \mathcal{U}(0)$ such that $O \cap K = \emptyset$. By Prop. 1.12(a), we may also assume O to be balanced as well. Then

$$U := A^{-1}(O) = A^{-1}(O \cap Y) \subseteq \mathbb{K}^n$$

satisfies $0 \in U$, $U \cap S = \emptyset$, and U balanced (as A is linear). We claim $U \subseteq B := B_1(0) = \{z \in \mathbb{K}^n : \|z\|_2 < 1\}$: Seeking a contradiction, assume $z \in U$ with $\|z\|_2 > 1$. Since U is balanced and $\|z\|_2^{-1} < 1$, this implies $\frac{z}{\|z\|_2} \in U$, in contradiction to $U \cap S = \emptyset$. But $U \subseteq B$ shows $U = A^{-1}(O)$ to be bounded. Since the n coordinate functions of A^{-1} are $\mathbb{K}$-linear functionals, the continuity of A^{-1} is now a consequence of Th. 1.14(iv).

(b): We show $\overline{Y} = Y$: Let $x \in \overline{Y}$. Let $A : \mathbb{K}^n \longrightarrow Y$, $O \subseteq X$, and B be as in the proof of (a). By Prop. 1.12(f), we may choose $r \in \mathbb{R}^+$ such that $x \in rO$. Then x is an element of the closure of each of the three sets

$$Y \cap (rO) \subseteq A\big(rB\big) \subseteq C := A\big(r\overline{B}\big).$$

As $r\overline{B}$ is compact, so is C. As above, we conclude C to be closed as well (as $(X, \mathcal{T})$ is T_2). Thus, $x \in C \subseteq Y$, showing $\overline{Y} = Y$ as desired. ■

Corollary 1.17. *If X is a finite-dimensional vector space over $\mathbb{K}$, then the norm topology on X is the unique topology on X that makes X into a T_1 topological vector space.*

Proof. This is an immediate consequence of applying Th. 1.16(a) with $Y := X$. ■

Definition 1.18. A topological space $(X, \mathcal{T})$ is called *locally compact* if, and only if, for each $x \in X$, there exists a compact neighborhood of x.

Theorem 1.19. **(a)** *If $(X, \mathcal{T})$ is a T_1 topological vector space over $\mathbb{K}$, then X has finite dimension if, and only if, $(X, \mathcal{T})$ is locally compact.*

(b) *A normed vector space $(X, \|\cdot\|)$ over $\mathbb{K}$ is finite-dimensional if, and only if, its closed unit ball $\overline{B}_1(0)$ is compact.*

Proof. (a): If X has finite dimension and $\mathcal{T}$ is T_1, then, by Cor. 1.17, $\mathcal{T}$ is the norm topology on X and that means balls are compact. In particular, $(X,\mathcal{T})$ is locally compact. Conversely, assume $(X,\mathcal{T})$ to be a locally compact T_1 space. Then there exist $O, K \in \mathcal{U}(0)$ such that $O \subseteq K$, O is open, K is compact. For each $x \in X$, one has $x \in x + \frac{1}{2}O$. Thus, $(x + \frac{1}{2}O)_{x\in K}$ is an open cover of K and there exist $x_1, \dots, x_m \in K$, $m \in \mathbb{N}$, such that

$$O \subseteq K \subseteq \bigcup_{i=1}^{m} \left(x_i + \frac{1}{2}O \right). \tag{1.9}$$

Let $Y := \operatorname{span}\{x_1, \dots, x_m\}$. Then $\dim Y \leq m$. We will show $Y = X$. As an intermediate step, we use an induction to prove

$$\underset{k\in\mathbb{N}}{\forall} \quad O \subseteq Y + 2^{-k}O : \tag{1.10}$$

The case $k = 1$ holds due to (1.9). Now let $k \in \mathbb{N}$ and assume $O \subseteq Y + 2^{-k}O$ to hold by induction hypothesis. Now if $x \in O$, then $\frac{1}{2}x \in \frac{1}{2}Y + \frac{1}{2} \cdot 2^{-k}O = Y + 2^{-(k+1)}O$ (as Y is a vector space). Thus,

$$O \subseteq Y + \frac{1}{2}O \subseteq Y + Y + 2^{-(k+1)}O = Y + 2^{-(k+1)}O,$$

completing the induction. As a consequence of (1.10), we now obtain

$$O \subseteq \bigcap_{k\in\mathbb{N}} (Y + 2^{-k}O). \tag{1.11}$$

Since K is compact, K (and, thus O) is bounded by Prop. 1.12(g), and $\{2^{-k}O : k \in \mathbb{N}\}$ is a local base at 0 according to Prop. 1.12(i). In consequence, we conclude

$$Y \overset{\text{Th. 1.16(b)}}{=} \overline{Y} \overset{\text{Prop. 1.10(a)}}{=} \bigcap_{U\in\mathcal{U}(0)} (Y + U) = \bigcap_{k\in\mathbb{N}} (Y + 2^{-k}O)$$

and (1.11) implies $O \subseteq Y$. Thus, $kO \subseteq Y$ for each $k \in \mathbb{N}$ (since Y is a vector space), and Prop. 1.12(f) yields $Y = X$ as desired.

(b): If X is finite-dimensional, then it is linearly homeomorphic to $\mathbb{K}^n$, $n \in \mathbb{N}$. Thus, $\overline{B}_1(0) \subseteq X$ is compact. Conversely, if $\overline{B}_1(0) \subseteq X$ is compact, then the space is locally compact, hence, finite-dimensional by (a). ■

1.3 Metrization

Definition 1.20. Let X be a vector space and let d be a metric on X. Then d is called *translation-invariant* if, and only if,

$$\underset{x,y,z\in X}{\forall} \quad d(x + z, y + z) = d(x, y). \tag{1.12}$$

Theorem 1.21. *Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{K}$. Then $(X, \mathcal{T})$ is metrizable if, and only if, it is both T_1 and C_1 (i.e. first countable). In that case, there exists a metric d on X that induces $\mathcal{T}$ and has the following additional properties:*

(i) *$B_r(0)$ is balanced for each $r \in \mathbb{R}^+$.*

(ii) *d is translation-invariant.*

If $(X, \mathcal{T})$ is also locally convex, then one can choose d to have properties (i), (ii), *and* (iii), *where*

(iii) *$B_r(x)$ is convex for each $x \in X$ and each $r \in \mathbb{R}^+$.*

Proof. If $(X, \mathcal{T})$ is metrizable, then it is T_1 (and even normal) by [Phi16b, Ex. 3.4(c)] and C_1 by [Phi16b, Rem. 1.39(a)]. To prove the converse needs some work. Assuming $(X, \mathcal{T})$ to be T_1 and C_1, we need to construct a suitable metric d on X.

Claim 1. There exists a sequence $(V_k)_{k\in\mathbb{N}}$ in $\mathcal{U}(0)$ such that each V_k is open and balanced (and convex if $(X, \mathcal{T})$ is locally convex), $\mathcal{B} := \{V_k : k \in \mathbb{N}\}$ is a local base at 0, and

$$\forall_{k\in\mathbb{N}} \quad V_{k+1} + V_{k+1} + V_{k+1} + V_{k+1} \subseteq V_k. \tag{1.13}$$

Proof. One starts with some countable local base at 0, given by sets $(W_j)_{j\in\mathbb{N}}$ (which exists as $(X, \mathcal{T})$ is C_1). Letting, for each $j \in \mathbb{N}$, $U_j := \bigcap_{k=1}^{j} W_k$, the sets $(U_j)_{j\in\mathbb{N}}$ still form a local base at zero, but the sets $(U_j)_{j\in\mathbb{N}}$ are also decreasing in the sense that $U_{j+1} \subseteq U_j$. From the U_j, one inductively constructs a sequence $(O_n)_{n\in\mathbb{N}}$ that satsfies everything the V_k are supposed to satisfy, except, instead of (1.13), one has $O_{n+1} \subseteq O_n$: Suppose $O_1, \dots, O_N$ are already constructed, $N \in \mathbb{N}_0$. Choose $j > N$ such that $U_j \subseteq O_N$ (and $j \in \mathbb{N}$ arbitrary for $N = 0$). By Prop. 1.12(a),(b), there exists $O \in \mathcal{U}(0)$ such that $O \subseteq U_j$, and O is open and balanced (and also convex for $(X, \mathcal{T})$ locally convex). Thus, we may set $O_{N+1} := O$, completing the definition of the O_n. From the O_n, we can now inductively construct the V_j: Set $V_1 := O_1$. Now suppose $V_1, \dots, V_J$ have already been constructed, $J \in \mathbb{N}$. Using Prop. 1.5(a) twice, we obtain $U \in \mathcal{U}(0)$ such that $U + U + U + U \subseteq V_J$. Then there exists $n > J$ such that $O_n \subseteq U$ (since the O_n are decreasing and form a local base at 0). Setting $V_{J+1} := O_n$, V_{J+1} has all the required properties, including $V_{J+1} + V_{J+1} + V_{J+1} + V_{J+1} \subseteq V_J$. ▲

We now define

$$D := \left\{ q \in \mathbb{Q} : \; q = \sum_{i=1}^{N(q)} c_i(q)\, 2^{-i},\; c_1(q), \dots, c_{N(q)}(q) \in \{0, 1\},\; N(q) \in \mathbb{N} \right\}.$$

Then, clearly, $D \subseteq [0,1[$, and the coefficients $c_1(q), \dots, c_{N(q)}(q)$ are uniquely determined by $q \in D$ if we require $c_{N(q)}(q) = 1$ for $q \neq 0$ (e.g., due to [Phi16a, Th. 7.99]). We now define the following functions:

$$A : D \cup \{1\} \longrightarrow \mathcal{P}(X), \qquad A(q) := \begin{cases} \sum_{i=1}^{N(q)} c_i(q)\, V_i & \text{if } q \in D, \\ X & \text{if } q = 1, \end{cases} \tag{1.14a}$$

$$f : X \longrightarrow [0,1], \qquad f(x) := \inf \big\{ q \in D \cup \{1\} : x \in A(q) \big\}, \tag{1.14b}$$

$$d : X \times X \longrightarrow [0,1], \qquad d(x,y) := f(x-y). \tag{1.14c}$$

Before we can prove that d constitutes a metric with the desired properties, we need to establish some properties of the function A:

Claim 2. Let $q \in D$. Then,

$$\underset{n \in \{1,\dots,N(q)\}}{\forall} \quad A(q) \subseteq \left(\sum_{i=1}^{n} c_i(q)\, V_i \right) + V_n.$$

Proof. We prove the claim by an induction on $n = N(q), \dots, 1$. Since the V_i contain 0, the claimed inclusion trivially holds for $n = N(q)$. Now assume the inclusion to hold for some $1 < n \leq N(q)$. Then

$$A(q) \subseteq \left(\sum_{i=1}^{n-1} c_i(q)\, V_i \right) + V_n + V_n \overset{(1.13)}{\subseteq} \left(\sum_{i=1}^{n-1} c_i(q)\, V_i \right) + V_{n-1},$$

completing the induction. ▲

Claim 3. Each $A(q)$, $q \in D \cup \{1\}$ is balanced (also convex it the V_i are convex). One has

$$A(0) = \{0\}. \tag{1.15a}$$

Moreover, for each $q, r \in D$, the following holds:

$$q + r \leq 1 \quad \Rightarrow \quad A(q) + A(r) \subseteq A(q+r), \tag{1.15b}$$

$$q < r \quad \Rightarrow \quad A(q) \subseteq A(r). \tag{1.15c}$$

Proof. According to (1.14a) and Prop. 1.8(a),(b), the $A(q)$ are balanced (resp. convex) if the V_i are. Next, $A(0) = 0 \cdot V_1 = \{0\}$ proves (1.15a). We now establish (1.15b), which is not quite as obvious. If $q + r = 1$, then $A(q+r) = X$ and (1.15b) holds. Thus, let $q + r < 1$. We extend the coefficients $c_i(q), c_i(r), c_i(q+r)$ to all $i \in \mathbb{N}$ by setting them 0 for $i > N(q)$, $i > N(r)$, $i > N(q+r)$, respectively. If $c_i(q) + c_i(r) = c_i(q+r)$ hold for each $i \in \mathbb{N}$, then (1.15b) (even with equality) is immediate from (1.14a). Otherwise, there exists a smallest $n \in \mathbb{N}$ such that $c_n(q) + c_n(r) \neq c_n(q+r)$ (due to $q + r < 1$ no carry can be left at $n = 1$). Then $c_n(q) = c_n(r) = 0$ and $c_n(q+r) = 1$. Thus,

$$A(q) \overset{\text{Cl. 2}}{\subseteq} \left(\sum_{i=1}^{n+1} c_i(q)\, V_i\right) + V_{n+1} \overset{c_n(q)=0}{\subseteq} \left(\sum_{i=1}^{n-1} c_i(q)\, V_i\right) + V_{n+1} + V_{n+1}.$$

Completely analogously, we also obtain

$$A(r) \subseteq \left(\sum_{i=1}^{n-1} c_i(r)\, V_i\right) + V_{n+1} + V_{n+1}.$$

Since, for $1 \leq i \leq n-1$, $c_i(q) + c_i(r) = c_i(q+r)$, we now obtain

$$\begin{aligned} A(q) + A(r) \quad &\subseteq \quad \left(\sum_{i=1}^{n-1} c_i(q+r)\, V_i\right) + V_{n+1} + V_{n+1} + V_{n+1} + V_{n+1} \\ &\overset{(1.13),\, c_n(q+r)=1}{\subseteq} \quad \sum_{i=1}^{N(q+r)} c_i(q+r)\, V_i = A(q+r), \end{aligned}$$

proving (1.15b). Finally, (1.15c) follows from (1.15b), since, for $r > q$, $r - q \in D$ and $A(q) \subseteq A(q) + A(r-q) \subseteq A(r)$. ▲

Claim 4. The function d, as defined in (1.14c), constitutes a translation-invariant metric on X.

Proof. We have $f(0) = 0$ due to (1.14b) and (1.15a), implying $d(x,x) = 0$ for each $x \in X$. Now let $x \in X$, $x \neq 0$. Since $(X, \mathcal{T})$ is T_1 and the V_k form a local base at 0, there exists $k \in \mathbb{N}$ such that $x \notin V_k$, i.e. $x \notin A(2^{-k}) = V_k$. Then (1.15c) implies $f(x) \geq 2^{-k} > 0$, also showing $d(x,y) > 0$ for $x \neq y$. From Cl. 3, we know the $A(q)$ to be balanced. Thus, $x \in A(q)$ if, and only if $-x \in A(q)$, implying $f(x) = f(-x)$ and $d(x,y) = d(y,x)$. Next, we show

$$\forall_{x,y \in X} \quad f(x+y) \leq f(x) + f(y): \tag{1.16}$$

Since (1.16) trivially holds for $f(x) + f(y) \geq 1$, it remains to consider the case $f(x) + f(y) < 1$. Clearly, D is dense in $[0,1]$, and, thus,

$$\forall_{\epsilon \in \mathbb{R}^+} \quad \exists_{q,r \in D} \quad \Big(f(x) < q \quad \wedge \quad f(y) < r \quad \wedge \quad q + r < \min\{1,\, f(x) + f(y) + \epsilon\}\Big).$$

In consequence, $x \in A(q)$, $y \in A(r)$, and $x + y \in A(q+r)$ by (1.15b). Thus,

$$f(x+y) \leq q + r < f(x) + f(y) + \epsilon,$$

proving (1.16), since $\epsilon > 0$ was arbitrary. Now, if $x, y, z \in X$, then

$$d(x,y) = f(x-y) = f(x-z+z-y) \leq f(x-z) + f(z-y) = d(x,z) + d(z,y),$$

proving the triangle inequality for d and that d is a metric. Translation invariance of d is immediate from (1.14c). ▲

Claim 5. One has

$$\underset{r\in\mathbb{R}^+}{\forall} \quad B_r(0) = \{x \in X : f(x) < r\} = \begin{cases} X & \text{for } r > 1, \\ \bigcup_{q\in D,\, q<r} A(q) & \text{for } r \leq 1. \end{cases} \tag{1.17}$$

Moreover, d induces $\mathcal{T}$ and all $B_r(0)$, $r \in \mathbb{R}^+$, are balanced. All $B_r(x)$, $r \in \mathbb{R}^+$, $x \in X$, are also convex if $(X, \mathcal{T})$ is locally convex.

Proof. The first equality in (1.17) is immediate from (1.14c). For $r > 1$, the second equality is also clear. For $0 < r \leq 1$, the second equality is due to

$$x \in B_r(0) \quad \Leftrightarrow \quad f(x) < r \quad \Leftrightarrow \quad \underset{q\in D, q<r}{\exists} x \in A(q) \quad \Leftrightarrow \quad x \in \bigcup_{q\in D,\, q<r} A(q).$$

Let $\mathcal{T}_d$ denote the topology induced by d. As d is translation-invariant, we have $B_r(x) = x + B_r(0)$ for each $r \in \mathbb{R}^+$, $x \in X$. Due to (1.17), $B_r(0) \in \mathcal{T}$ for each $r \in \mathbb{R}^+$ (as each $A(q)$, $q > 0$, is open due to Prop. 1.10(d)). Then, for each $x \in X$, $B_r(x) = x + B_r(0) \in \mathcal{T}$ as well, showing $\mathcal{T}_d \subseteq \mathcal{T}$. For the remaining inclusion, we recall $V_k = A(2^{-k})$, i.e. (1.17) implies $B_r(0) \subseteq V_k$ for $r < 2^{-k}$, showing $\mathcal{B}_0 := \{B_r(0) : r \in \mathbb{R}^+\}$ to be a local base for $\mathcal{T}$ at 0. In consequence, for each $x \in X$, $\mathcal{B}_x := \{B_r(x) : r \in \mathbb{R}^+\}$ is a local base for $\mathcal{T}$ at x, showing $\mathcal{T} \subseteq \mathcal{T}_d$. Each $A(q)$ is balanced according (1.14a) and Prop. 1.8(b). Then, according to (1.17) and Prop. 1.8(b), each $B_r(0)$ is balanced as well. If $(X, \mathcal{T})$ is locally convex, then each $A(q)$ is convex by Cl. 3. If $x, y \in B_r(0)$, then, by (1.17) they must be in the same (convex) $A(q) \subseteq B_r(0)$ (for some suitable $q \in D \cup \{1\}$), showing $B_r(0)$ to be convex. Then each $B_r(x) = x + B_r(0)$ is convex as well. ▲

With the proof of Cl. 5, we have concluded the proof of the theorem. ■

Definition 1.22. Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{K}$.

(a) $(X, \mathcal{T})$ is called an *F-space* if, and only if, $\mathcal{T}$ is induced by a complete translation-invariant metric on X.

(b) $(X, \mathcal{T})$ is called a *Fréchet space* if, and only if, it is a locally convex F-space.

(c) $(X, \mathcal{T})$ is called a *normable* if, and only if, $\mathcal{T}$ is induced by a norm on X.

Remark 1.23. For a normed vector space X, we, clearly, have the equivalences

$$X \text{ Fréchet space} \quad \Leftrightarrow \quad X\ F\text{-space} \quad \Leftrightarrow \quad X \text{ Banach space}.$$

The spaces $L^p([0,1],\mathcal{L}^1,\lambda^1)$, $0 < p < 1$, of Ex. 1.11(b) are examples of nonnormable F-spaces that are not Fréchet: They are complete by [Phi17, Th. 2.44] (i.e. F), but they are not locally convex by Ex. 1.11(b) (i.e. neither Fréchet nor normable). We will see in Th. 1.41 below that a T_1 topological vector space is normable if, and only if, it is both locally bounded and locally convex.

—

We conclude the section with a lemma we will use in the proof of Th. 1.32 below:

Lemma 1.24. **(a)** *Let X be a vector space and let d be a translation-invariant metric on X (here, we do not assume that d makes X into a topological vector space). Then*

$$\underset{x\in X}{\forall} \quad \underset{n\in\mathbb{N}}{\forall} \quad d(nx,0) \leq n d(x,0).$$

(b) *Let $(X,\mathcal{T})$ be a metrizable topological vector space over $\mathbb{K}$ and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X such that $\lim_{n\to\infty} x_n = 0$. Then there exists an increasing sequence $(N_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that $\lim_{n\to\infty} N_n = \infty$ and $\lim_{n\to\infty} N_n x_n = 0$.*

Proof. Exercise. ■

1.4 Boundedness, Cauchy Sequences, Continuity

Boundedness and Cauchy sequences are both notions that, in general, do not make sense in an arbitrary topological space. However, both notions are familiar in metric spaces and both can be defined in arbitrary topological vector spaces (see Def. 1.9(c) and Def. 1.27, respectively). Unfortunately, if the topology $\mathcal{T}$ of a topological vector space is induced by a metric d, then the resulting notions of boundedness and Cauchy sequences with respect to $\mathcal{T}$ and d are, in general, not the same (cf. Rem. 1.26 and Ex. 1.30 below). Thus, it is necessary to use some care regarding these notions. Some related results will be presented in the current section.

If we call a subset of a topological vector space *bounded*, we will always mean bounded in the sense of Def. 1.9(c). When we need to distinguish this boundedness from the boundedness with respect to a metric d, then we will speak of d-boundedness in regard to the metric.

Proposition 1.25. *Let $(X,\mathcal{T})$ be a topological vector space over $\mathbb{K}$, $A \subseteq X$. Then the following statements are equivalent:*

(i) *A is bounded.*

(ii) *For each sequence $(x_n)_{n\in\mathbb{N}}$ in A and each sequence $(\lambda_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$ one has*

$$\lim_{n\to\infty} \lambda_n = 0 \quad \Rightarrow \quad \lim_{n\to\infty} \lambda_n \, x_n = 0.$$

Proof. Suppose, A is bounded. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in A and let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{K}$ such that $\lim_{n\to\infty} \lambda_n = 0$. Let $U \in \mathcal{U}(0)$. Then there exists $B \in \mathcal{U}(0)$ such that $B \subseteq U$ and B is balanced. Since A is bounded, there exists $t \in \mathbb{R}^+$ such that $A \subseteq tB$. Then

$$\underset{N\in\mathbb{N}}{\exists} \quad \underset{n>N}{\forall} \quad t\,|\lambda_n| < 1.$$

We also know $t^{-1}x_n \in B$ for each $n \in \mathbb{N}$ since $t^{-1}A \subseteq B$. Thus, since B is balanced, for each $n > N$, $t\lambda_n t^{-1}x_n = \lambda_n x_n \in B \subseteq U$, showing $\lim_{n\to\infty} \lambda_n \, x_n = 0$. Conversely, if A is not bounded, then there exists $U \in \mathcal{U}(0)$ such that no nU, $n \in \mathbb{N}$, contains A. For each $n \in \mathbb{N}$, let $x_n \in A \setminus nU$. Then $\lim_{n\to\infty} \frac{1}{n} = 0$, but $\frac{1}{n}x_n \notin U$ for each $n \in \mathbb{N}$ (otherwise, $n\frac{1}{n}x_n = x_n \in nU$), showing $\frac{1}{n}x_n \not\to 0$ for $n \to \infty$. ∎

Remark 1.26. Let $(X, \mathcal{T})$ be a T_1 topological vector space over $\mathbb{K}$. If $Y \neq \{0\}$ is a vector subspace of X, then Y is not bounded (in general, this is not true without the T_1 hypothesis – e.g., if $(X, \mathcal{T})$ is indiscrete, then, clearly, *every* subset is bounded): Let $0 \neq y \in Y$. Then $(ny)_{n\in\mathbb{N}}$ is a sequence in Y such that $\lim_{n\to\infty}(\frac{1}{n}\,n\,y) = y \neq 0$. If $(X, \mathcal{T})$ is T_1, then it is T_2 and limits are unique, showing Y not to be bounded by Prop. 1.25. On the other hand, the metric d constructed in the proof of Th. 1.21 is such that X itself is d-bounded (by 1), showing that each metrizable topological vector space is metrizable by a translation-invariant *bounded* metric, even though $X \neq \{0\}$ can not be bounded.

Definition 1.27. Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{K}$, and let $\mathcal{B}$ be a local base at 0. Then a sequence $(x_n)_{n\in\mathbb{N}}$ in X is called a *Cauchy sequence* if, and only if,

$$\underset{B\in\mathcal{B}}{\forall} \quad \underset{N\in\mathbb{N}}{\exists} \quad \underset{m,n>N}{\forall} \quad x_n - x_m \in B. \tag{1.18}$$

Moreover, $(X, \mathcal{T})$ is called *complete* if, and only if, every Cauchy sequence in X converges in X.

Proposition 1.28. *Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{K}$.*

(a) *Then the notion of Cauchy sequence as defined in Def. 1.27 does not depend on the chosen local base at 0.*

(b) *If $(x_n)_{n\in\mathbb{N}}$ is a convergent sequence in X, then it is a Cauchy sequence.*

(c) *If $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in X (in particular, if it is a convergent sequence in X), then it is a bounded sequence (i.e. $\{x_n : n \in \mathbb{N}\}$ is bounded).*

Proof. (a): Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X, let $\mathcal{B}, \mathcal{C}$ be local bases at 0. We show $(x_n)_{n\in\mathbb{N}}$ is $\mathcal{C}$-Cauchy if it is $\mathcal{B}$-Cauchy: Let $C \in \mathcal{C}$. Then there exists $B \in \mathcal{B}$ such that $B \subseteq C$. To B choose $N \in \mathbb{N}$ according to (1.18). Then

$$\forall_{m,n>N} \quad x_n - x_m \in B \subseteq C,$$

showing $(x_n)_{n\in\mathbb{N}}$ to be $\mathcal{C}$-Cauchy.

(b): Suppose, $\lim_{n\to\infty} x_n = x \in X$. Then $\lim_{n\to\infty}(x_n - x) = 0$. Let $W \in \mathcal{U}(0)$. According to Prop. 1.5, there exists $U \in \mathcal{U}(0)$ such that $U = -U$ and $U + U \subseteq W$. Let $N \in \mathbb{N}$ such that $x_n - x \in U$ for each $n > N$. Then

$$\forall_{m,n>N} \quad x_n - x_m = x_n - x - (x_m - x) \in U - U = U + U \subseteq W,$$

showing $(x_n)_{n\in\mathbb{N}}$ to be Cauchy.

(c): Let $U \in \mathcal{U}(0)$. Then there exists $V \in \mathcal{U}(0)$ such that V is balanced and $V + V \subseteq U$. Since $(x_n)_{n\in\mathbb{N}}$ is Cauchy, there exists $N \in \mathbb{N}$ such that $x_n - x_m \in V$ for each $m, n > N$. Thus, $x_n \in x_{N+1} + V$ for each $n > N$. Choose $s > 1$ such that $x_{N+1} \in sV$. Then

$$\forall_{n>N} \quad x_n \in sV + V \overset{\text{Prop. 1.8(c)}}{\subseteq} sV + sV \subseteq sU.$$

Since $\{x_n : n \leq N\}$ is bounded by Prop. 1.12(e), it follows that $\{x_n : n \in \mathbb{N}\}$ is bounded. ∎

Theorem 1.29. *Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{K}$.*

(a) *If d_1, d_2 are translation-invariant metrics on X that both induce $\mathcal{T}$, then, given a sequence $(x_n)_{n\in\mathbb{N}}$ in X, the following statements are equivalent:*

(i) *$(x_n)_{n\in\mathbb{N}}$ is d_1-Cauchy.*

(ii) *$(x_n)_{n\in\mathbb{N}}$ is d_2-Cauchy.*

(iii) *$(x_n)_{n\in\mathbb{N}}$ is Cauchy in the sense of Def. 1.27.*

(b) *If d_1, d_2 are translation-invariant metrics on X that both induce $\mathcal{T}$, then the following statements are equivalent:*

(i) *(X, d_1) is a complete metric space.*

(ii) (X, d_2) *is a complete metric space.*

(iii) *Every Cauchy sequence in* X *converges (i.e.* $(X, \mathcal{T})$ *is complete).*

(c) *Let* $(X, \mathcal{T})$ *be* T_1*. If* $Y \subseteq X$ *is a vector subspace (with the relative topology) and* Y *is an* F*-space, then* Y *is closed in* X*.*

Proof. Exercise. ■

Example 1.30. Consider

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \quad f(x) := \frac{x}{1 + |x|}.$$

It is an exercise to verify that

$$d : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}_0^+, \quad d(x, y) := |f(x) - f(y)|,$$

defines a metric on $\mathbb{R}$ that is equivalent to the metric induced by $|\cdot|$ (i.e. it induces the same topology), but, in contrast to $(\mathbb{R}, |\cdot|)$, $(\mathbb{R}, d)$ is *not* complete.

—

Due to Rem. 1.26, a nontrivial linear map $A : X \longrightarrow Y$ between T_1 topological vector spaces can never be bounded in the sense that $A(X)$ is a bounded subset of Y. However, for such maps, the following definition turns out to be useful:

Definition 1.31. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological vector spaces over $\mathbb{K}$. Then a $\mathbb{K}$-linear map $A : X \longrightarrow Y$ is called *bounded* if, and only if,

$$\underset{B \subseteq X}{\forall} \Big(B \text{ bounded} \quad \Rightarrow \quad A(B) \text{ bounded} \Big).$$

—

We can now supplement Th. 1.13 with the following result:

Theorem 1.32. *Regarding the following statements* (i) – (iii) *for a* $\mathbb{K}$*-linear function* $A : X \longrightarrow Y$ *between topological vector spaces* $(X, \mathcal{T}_X)$ *and* $(Y, \mathcal{T}_Y)$ *over* $\mathbb{K}$*, one has the implications*

$$\text{(i)} \quad \Rightarrow \quad \text{(ii)} \quad \Rightarrow \quad \text{(iii)}.$$

If $(X, \mathcal{T}_X)$ *is metrizable, then one also has the additional implications*

$$\text{(iii)} \quad \Rightarrow \quad \text{(iv)} \quad \Rightarrow \quad \text{(i)}$$

and all statements (i) – (iv) *are equivalent.*

(i) *A is continuous.*

(ii) *A is bounded.*

(iii) *For each sequence $(x_n)_{n\in\mathbb{N}}$ in X such that $\lim_{n\to\infty} x_n = 0$, the sequence $(A(x_n))_{n\in\mathbb{N}}$ is bounded in Y.*

(iv) *For each sequence $(x_n)_{n\in\mathbb{N}}$ in X, one has*

$$\lim_{n\to\infty} x_n = 0 \quad \Rightarrow \quad \lim_{n\to\infty} A(x_n) = 0.$$

Proof. "(i)⇒(ii)": Let A be continuous, let $B \subseteq X$ be bounded, and $U \subseteq Y$, $U \in \mathcal{U}(0)$. As A is continuous, there exists $V \subseteq X$, $V \in \mathcal{U}(0)$ such that $A(V) \subseteq U$. Since B is bounded,

$$\underset{s\in\mathbb{R}^+}{\exists} \quad B \subseteq sV,$$

implying

$$A(B) \subseteq A(sV) = sA(V) \subseteq sU,$$

showing $A(B)$ to be bounded.

"(ii)⇒(iii)": Assume A to be bounded and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X such that $\lim_{n\to\infty} x_n = 0$. Then $(x_n)_{n\in\mathbb{N}}$ is bounded according to Prop. 1.28(c). Thus, $(A(x_n))_{n\in\mathbb{N}}$ is bounded due to the boundedness of A.

For the remaining implications, we now assume $(X, \mathcal{T}_X)$ to be metrizable.

"(iii)⇒(iv)": Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X such that $\lim_{n\to\infty} x_n = 0$. According to Lem. 1.24(b), there exists a sequence $(N_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that $\lim_{n\to\infty} N_n = \infty$ and $\lim_{n\to\infty} N_n x_n = 0$. Using (iii), $(A(N_n x_n))_{n\in\mathbb{N}}$ is a bounded sequence in Y. Since $\lim_{n\to\infty} N_n^{-1} = 0$, Prop. 1.25(ii) implies

$$\lim_{n\to\infty} A(x_n) = \lim_{n\to\infty} \left(N_n^{-1} N_n A(x_n)\right) = 0.$$

"(iv)⇒(i)": Since $(X, \mathcal{T}_X)$ is metrizable, (iv) implies the continuity of A at 0 by [Phi16b, Th. 2.8], which, in turn, implies the (global) continuity of A according to Th. 1.13. ∎

We will see in Ex. 1.42 below that it is possible for a linear map between topological vector spaces to be bounded without being continuous.

1.5 Seminorms and Local Convexity

We are already familiar with norms on vector spaces over $\mathbb{K}$. We have also encountered the seminorms $\|\cdot\|_p$ on the spaces $\mathcal{L}^p(\mu)$, given a measure space $(\Omega, \mathcal{A}, \mu)$, $p \in [1, \infty]$ (cf. [Phi17, Def. and Rem. 2.41(a)] and [Phi16b, Sec. C.3]). In the present section, we investigate the relation between seminorms and topological vector spaces more thoroughly. As a main result, we will see in Th. 1.40 that a topological vector space $(X, \mathcal{T})$ is locally convex if, and only if, $\mathcal{T}$ is induced by a family of seminorms. Moreover, according to Th. 1.41, $(X, \mathcal{T})$ is normable if, and only if, it is T_1 and both locally convex and locally bounded.

Definition 1.33. Let X be a vector space over $\mathbb{K}$. Then a function $p : X \longrightarrow \mathbb{R}_0^+$ is called a *seminorm* on X if, and only if, the following three conditions are satisfied:

(i) $p(0) = 0$.

(ii) p is *homogeneous of degree* 1, i.e.

$$p(\lambda x) = |\lambda|\, p(x) \qquad \text{for each } \lambda \in \mathbb{K},\ x \in X.$$

(iii) p is *subadditive*, i.e. p satisfies the *triangle inequality*, i.e.

$$p(x + y) \leq p(x) + p(y) \qquad \text{for each } x, y \in X.$$

If p constitutes a seminorm on X, then the pair (X, p) is called a *seminormed vector space* or just *seminormed space*. Given a seminormed space (X, p), we denote open and closed balls by

$$\underset{x\in X}{\forall}\quad \underset{r\in\mathbb{R}^+}{\forall}\quad \begin{aligned} B_{p,r}(x) &:= \{y \in X : p(x - y) < r\}, \\ \overline{B}_{p,r}(x) &:= \{y \in X : p(x - y) \leq r\}. \end{aligned} \tag{1.19}$$

Remark 1.34. Let (X, p) be a seminormed vector space over $\mathbb{K}$.

(a) Clearly, p is a norm if, and only if, $p(x) \neq 0$ for each $x \neq 0$, $x \in X$.

(b) p induces the pseudometric $d : X \times X \longrightarrow \mathbb{R}_0^+$, $d(x, y) := p(x - y)$, on X (cf. [Phi16b, Def. C.9]). Since $d(x + z, y + z) = p(x + z - y - z) = d(x, y)$, it is always translation-invariant. Related results are

$$\underset{x\in X}{\forall}\quad \underset{r\in\mathbb{R}^+}{\forall}\quad \begin{aligned} B_{p,r}(x) &= x + B_{p,r}(0), \\ \overline{B}_{p,r}(x) &= x + \overline{B}_{p,r}(0), \end{aligned} \tag{1.20}$$

since

$$y \in x + B_{p,r}(0) \;\Leftrightarrow\; y = x + z,\ p(z) < r \;\Leftrightarrow\; p(y - x) < r \;\Leftrightarrow\; y \in B_{p,r}(x)$$

(and analogously for the closed balls).

(c) Let $x \in X$, $r \in \mathbb{R}^+$. Then the balls $B_{p,r}(0)$, $\overline{B}_{p,r}(0)$ are balanced; the balls $B_{p,r}(x)$, $\overline{B}_{p,r}(x)$ are convex: If $p(y) < r$ (resp. $p(y) \leq r$), and $\lambda \in \mathbb{K}$, $|\lambda| \leq 1$, then $p(\lambda y) = |\lambda| p(y) < r$ (resp. $\leq r$). If $y, z \in \overline{B}_{p,r}(x)$, then

$$\underset{\alpha\in[0,1]}{\forall} \quad \begin{aligned} p\big(\alpha y + (1-\alpha)z - x\big) &\leq p\big(\alpha(y-x)\big) + p\big((1-\alpha)(z-x)\big) \\ &= \alpha p(y-x) + (1-\alpha)p(z-x) \leq r, \end{aligned}$$

with strict inequality for $y, z \in B_{p,r}(x)$.

Definition and Remark 1.35. Let X be vector space over $\mathbb{K}$, $A \subseteq X$.

(a) A is called *absorbing* if, and only if,

$$X = \bigcup_{t\in\mathbb{R}^+} (t\,A). \tag{1.21}$$

It is then immediate that A absorbing implies $0 \in A$. If $(X, \mathcal{T})$ is a topological vector space and $U \in \mathcal{U}(0)$, then we know from Prop. 1.12(f) that U is absorbing.

(b) If A is absorbing, then we define

$$\mu_A : X \longrightarrow \mathbb{R}_0^+, \quad \mu_A(x) := \inf\{t \in \mathbb{R}^+ : t^{-1}x \in A\}, \tag{1.22}$$

and call μ_A the *Minkowski functional* of A. Since A is absorbing, for each $x \in X$, there exists $t \in \mathbb{R}^+$ such that $t^{-1}x \in A$ and μ_A is well-defined. It is also immediate from (1.22) that $\mu_A(0) = 0$.

—

We will see in Cor. 1.38(a) below that seminorms are precisely the Minkowski functionals of balanced convex absorbing sets. The definition of a seminorm that we gave in Def. 1.33 is quite common and it underlines the relation between seminorms and norms. However, it turns out that nonnegativity and Def. 1.33(i) do not have to be required in the definition, as they *follow* from Def. 1.33(ii),(iii), as we will now see as part of Th. 1.36:

Theorem 1.36. *Let X be a vector space over $\mathbb{K}$. Then a function $p : X \longrightarrow \mathbb{R}$ is a seminorm on X if, and only if, p satisfies Def.* 1.33(ii),(iii)*. Moreover, if p is a seminorm, then it further satisfies*

(a) $|p(x) - p(y)| \leq p(x-y)$ *for each* $x, y \in X$.

(b) $p^{-1}(\{0\})$ *is a vector subspace of* X.

(c) *The unit ball* $B := \{x \in X : p(x) < 1\}$ *is convex, balanced, and absorbing, with* $p = \mu_B$.

Proof. Clearly, it suffices to show that each $p : X \longrightarrow \mathbb{R}$ with the properties of Def. 1.33(ii),(iii) is a seminorm and satisfies (a) – (c). Thus, assume $p : X \longrightarrow \mathbb{R}$ to have the properties of Def. 1.33(ii),(iii). We then have

$$p(0) = p(0 \cdot 0) \overset{\text{Def. 1.33(ii)}}{=} |0| \cdot p(0) = 0,$$

proving Def. 1.33(i). Next, we note that

$$\begin{aligned} p(x) = p(x - y + y) &\overset{\text{Def. 1.33(iii)}}{\leq} p(x - y) + p(y), \\ p(y) = p(y - x + x) &\overset{\text{Def. 1.33(iii)}}{\leq} p(y - x) + p(x) \overset{\text{Def. 1.33(ii)}}{=} p(x - y) + p(x), \end{aligned}$$

proves (a). Applying (a) with $y := 0$, proves p to be $\mathbb{R}_0^+$-valued and, thus, a seminorm. Now let $x, y \in X$ such that $p(x) = p(y) = 0$ and $\lambda, \mu \in \mathbb{K}$. Then

$$0 \leq p(\lambda x + \mu y) \overset{\text{Def. 1.33(ii),(iii)}}{\leq} |\lambda| p(x) + |\mu| p(y) = 0,$$

showing $p(\lambda x + \mu y) = 0$ and, thus, (b). We know B to be convex and balanced by Rem. 1.34(c). For each $x \in X$,

$$\forall_{t > p(x)} \quad p(t^{-1}x) = t^{-1}p(x) < 1,$$

showing $t^{-1}x \in B$. Thus, B is absorbing and $\mu_B(x) \leq p(x)$. Since $x \in X$ was arbitrary, we have $\mu_B \leq p$. Conversely, if $0 < t \leq p(x)$, then $p(t^{-1}x) = t^{-1}p(x) \geq 1$, showing $p(x) \leq \mu_B(x)$ and $p \leq \mu_B$. ∎

Theorem 1.37. *Let* X *be a vector space over* $\mathbb{K}$ *and let* $A, B \subseteq X$ *be absorbing with Minkowski functionals* μ_A, μ_B. *Then the following statements hold true:*

(a) *If* $A \subseteq B$, *then* $\mu_B \leq \mu_A$.

(b) $\mu_A(sx) = s\mu_A(x)$ *holds for each* $x \in X$ *and each* $s \in \mathbb{R}_0^+$.

(c) *If* A *is balanced, then* $\mu_A(\lambda x) = |\lambda| \mu_A(x)$ *holds for each* $x \in X$ *and each* $\lambda \in \mathbb{K}$.

(d) *If* A *is convex, then*

$$\forall_{x \in X} \ \forall_{s \in \mathbb{R}^+} \ \Big(s > \mu_A(x) \ \Rightarrow \ s^{-1}x \in A \Big).$$

(e) *If A is convex, then*

$$\forall_{x,y\in X} \quad \mu_A(x+y) \leq \mu_A(x) + \mu_A(y).$$

(f) *If A is convex and balanced, then μ_A is a seminorm on X.*

(g) *If A is convex, then, for each $r \in \mathbb{R}^+$, the sets $B_{A,r} := \{x \in X : \mu_A(x) < r\}$ and $C_{A,r} := \{x \in X : \mu_A(x) \leq r\}$ are convex.*

(h) *Using the notation from* (g)*, one always has $A \subseteq C_{A,1}$ and, if A is convex, then $B_{A,1} \subseteq A \subseteq C_{A,1}$ and $\mu_{B_{A,1}} = \mu_A = \mu_{C_{A,1}}$.*

Proof. (a): If $A \subseteq B$, then

$$\forall_{x\in X} \quad \{t \in \mathbb{R}^+ : t^{-1}x \in A\} \subseteq \{t \in \mathbb{R}^+ : t^{-1}x \in B\},$$

proving (a).

(b): $\mu_A(0) = 0$ yields the case $s = 0$. If $s > 0$, then, for each $x \in X$,

$$\begin{aligned}\mu_A(sx) &= \inf\{t \in \mathbb{R}^+ : t^{-1}sx \in A\} \overset{(*)}{=} \inf\{st \in \mathbb{R}^+ : t^{-1}x \in A\}\\ &= s\inf\{t \in \mathbb{R}^+ : t^{-1}x \in A\} = s\mu_A(x),\end{aligned}$$

where the equality at $(*)$ holds since, for

$$M := \{t \in \mathbb{R}^+ : t^{-1}sx \in A\} \quad \text{and} \quad N := \{st \in \mathbb{R}^+ : t^{-1}x \in A\},$$

one has

$$r \in M \;\Leftrightarrow\; r^{-1}sx \in A \;\Leftrightarrow\; \left(\frac{r}{s}\right)^{-1} x \in A \;\Leftrightarrow\; r \in N.$$

(c): If A is balanced, $x \in X$, $0 \neq \lambda \in \mathbb{K}$, then $x \in A$ if, and only if, $\frac{|\lambda|}{\lambda}x \in A$. Thus,

$$\begin{aligned}\mu_A(\lambda x) &= \inf\{t \in \mathbb{R}^+ : t^{-1}\lambda x \in A\} = \inf\left\{t \in \mathbb{R}^+ : t^{-1}\lambda\frac{|\lambda|}{\lambda}x \in A\right\} = \mu_A(|\lambda|x)\\ &\overset{\text{(b)}}{=} |\lambda|\mu_A(x).\end{aligned}$$

(d),(e): Exercise.

(f) is now a consequence of combining (c) and (e) with Th. 1.36.

(g): Let A be convex and fix $x, y \in C_{A,r}$, $r > 0$. We use (e) to obtain

$$\forall_{\alpha\in[0,1]} \quad \mu_A(\alpha x + (1-\alpha)y) \overset{\text{(e)}}{\leq} \mu_A(\alpha x) + \mu_A((1-\alpha)y) \overset{\text{(b)}}{=} \alpha\mu_A(x) + (1-\alpha)\mu_A(y) \leq r,$$

with strict inequality for $x, y \in B_{A,r}$, showing $\alpha x + (1-\alpha)y \in C_{A,r}$, and $\alpha x + (1-\alpha)y \in B_{A,r}$ for $x, y \in B_{A,r}$.

(h): If $x \in A$, then $\frac{x}{1} \in A$, showing $\mu_A(x) \leq 1$ and $x \in C_{A,1}$. Now let A be convex. If $x \in B_{A,1}$, then $x = \frac{x}{1} \in A$ by (d). As we now have $B_{A,1} \subseteq A \subseteq C_{A,1}$, $\mu_{C_{A,1}} \leq \mu_A \leq \mu_{B_{A,1}}$ follows from (a). To prove equality, we show $\mu_{B_{A,1}} \leq \mu_{C_{A,1}}$: Let $x \in X$. As $C_{A,1}$ is convex by (g), we can use (d) again to obtain, for each $\mu_{C_{A,1}}(x) < s < t$, $s^{-1}x \in C_{A,1}$, i.e. $\mu_A(\frac{x}{s}) \leq 1$, i.e. $\mu_A(\frac{x}{t}) = \frac{s}{t}\mu_A(\frac{x}{s}) \leq \frac{s}{t} < 1$. Thus, $t^{-1}x \in B_{A,1}$, showing $\mu_{B_{A,1}}(x) \leq t$. Taking the limit $t \downarrow \mu_{C_{A,1}}(x)$ yields $\mu_{B_{A,1}}(x) \leq \mu_{C_{A,1}}(x)$ as desired. ■

Corollary 1.38. *Let X be a vector space over $\mathbb{K}$.*

(a) *A function $p : X \longrightarrow \mathbb{R}$ is a seminorm on X if, and only if, there exists an absorbing, convex, balanced set $A \subseteq X$ such that $p = \mu_A$.*

(b) *Let $p : X \longrightarrow \mathbb{R}$ be a seminorm on X. Then p is a norm on X if, and only if, the unit ball $B_{p,1}(0)$ does not contain a nontrivial vector subspace of X.*

Proof. (a): We know "⇒" from Th. 1.36(c) and "⇐" from Th. 1.37(f).

(b) is an immediate consequence of Th. 1.36(b). ■

Definition 1.39. Let X be a vector space over $\mathbb{K}$. A family $(p_i)_{i\in I}$ (I some index set) of seminorms on X is called *separating* if, and only if,

$$\forall_{0\neq x\in X} \quad \exists_{i\in I} \quad p_i(x) \neq 0. \tag{1.23}$$

Theorem 1.40. *Let X be a vector space over $\mathbb{K}$.*

(a) *Let $I \neq \emptyset$ be an index set and let $\mathcal{P} := (p_i)_{i\in I}$ be a family of seminorms on X. If*

$$\mathcal{S} := \{B_{p_i,\frac{1}{n}}(0) : n \in \mathbb{N}, i \in I\},$$

then

$$\mathcal{B} := \{x + B : x \in X, B \text{ finite intersection of sets from } \mathcal{S}\}$$

forms the base of a topology $\mathcal{T}$ on X – we call $\mathcal{T}$ the topology induced *by the family of seminorms $\mathcal{P}$. Then, for each $x \in X$,*

$$\mathcal{B}_x := \{x + B : B \text{ finite intersection of sets from } \mathcal{S}\}$$

is a local base for $\mathcal{T}$ at x and, moreover, $(X, \mathcal{T})$ is a locally convex topological vector space with the additional properties:

(i) *Each* p_i, $i \in I$, *is continuous.*

(ii) $E \subseteq X$ *is bounded if, and only if, each* p_i, $i \in I$, *is bounded on* E.

The space $(X, \mathcal{T})$ *is* T_1 *if, and only if, the family* $\mathcal{P}$ *is separating. If* $(X, \mathcal{T})$ *is* T_1 *and* $I = \mathbb{N}$, *then* $(X, \mathcal{T})$ *is metrizable by each metric*

$$d : X \times X \longrightarrow \mathbb{R}_0^+, \quad d(x,y) := \max\left\{ \frac{c_i\, p_i(x-y)}{1 + p_i(x-y)} : i \in \mathbb{N} \right\}, \tag{1.24}$$

where $(c_i)_{i\in\mathbb{N}}$ *is some sequence in* $\mathbb{R}^+$ *converging to* 0 *(e.g.,* $c_i := \frac{1}{i}$*). Furthermore,* d *is translation-invariant and the set of open metric balls* $\mathcal{D} := \{B_{d,r}(0) : r \in \mathbb{R}^+\}$ *forms a convex balanced local base for* $\mathcal{T}$ *at* 0.

(b) *Let* $(X, \mathcal{T})$ *be a topological vector space and let* $B \in \mathcal{U}(0)$ *be open and convex. Since each* B *is absorbing by Prop.* 1.12(f), *the corresponding Minkowski functional* μ_B *is well-defined. Then*

$$B = B_1 := \{x \in X : \mu_B(x) < 1\}. \tag{1.25}$$

Moreover, if $\mathcal{B}$ *is a local base for* $\mathcal{T}$ *at* 0, *consisting of convex balanced open sets (we know from Prop.* 1.12(b) *that such a* $\mathcal{B}$ *exists if* $(X, \mathcal{T})$ *is locally convex), then* $\mathcal{P} := (\mu_B)_{B\in\mathcal{B}}$ *constitutes a family of (continuous) seminorms on* X, *and this family induces* $\mathcal{T}$. *Moreover,* $(X, \mathcal{T})$ *is* T_1 *if, and only if,* $\mathcal{P}$ *is separating.*

Proof. (a): To apply [Phi16b, Prop. 1.48] to show $\mathcal{B}$ constitutes the base for a topology $\mathcal{T}$ on X, we have to show that $\mathcal{B}$ is a cover of X (i.e. $X = \bigcup_{B\in\mathcal{B}} B$) and

$$\underset{B_1,B_2\in\mathcal{B}}{\forall} \quad \underset{x\in B_1\cap B_2}{\forall} \quad \underset{B_3\in\mathcal{B}}{\exists} \quad x \in B_3 \subseteq B_1 \cap B_2.$$

If $x \in X$, then $x \in B_{p_i,1}(x) \in \mathcal{B}$ for $i \in I$ shows $\mathcal{B}$ to be a cover of X. Now let $B_1, B_2 \in \mathcal{B}$ and $x \in B_1 \cap B_2$. There exist $x_1, x_2 \in X$; $N_1, N_2 \in \mathbb{N}$; $m_1, \dots, m_{N_1}, n_1, \dots, n_{N_2} \in \mathbb{N}$; and $i_1, \dots, i_{N_1}, j_1, \dots, j_{N_2} \in I$ such that

$$B_1 = \bigcap_{k=1}^{N_1} B_{p_{i_k}, \frac{1}{m_k}}(x_1), \quad B_2 = \bigcap_{k=1}^{N_2} B_{p_{j_k}, \frac{1}{n_k}}(x_2).$$

For each i_k, there exists $a_k \in \mathbb{N}$ such that $B_{p_{i_k}, \frac{1}{a_k}}(x) \subseteq B_{p_{i_k}, \frac{1}{m_k}}(x_1)$, and for each j_k, there exists $b_k \in \mathbb{N}$ such that $B_{p_{j_k}, \frac{1}{b_k}}(x) \subseteq B_{p_{j_k}, \frac{1}{n_k}}(x_2)$. Letting

$$N := \max\Big(\{a_k : k \in \{1, \dots, N_1\}\} \cup \{b_k : k \in \{1, \dots, N_2\}\} \Big),$$

$$B_3 := \bigcap_{k=1}^{N_1} B_{p_{i_k}, \frac{1}{N}}(x) \cap \bigcap_{k=1}^{N_2} B_{p_{j_k}, \frac{1}{N}}(x) \in \mathcal{B},$$

we have $x \in B_3 \subseteq B_1 \cap B_2$ as desired. If $x \in X$ and $U \in \mathcal{U}(x)$, then there exist $B \in \mathcal{B}$ such that $x \in B \subseteq U$. The previous argument shows there is $B_3 \in \mathcal{B}_x$ such that $x \in B_3 \subseteq B \subseteq U$, proving $\mathcal{B}_x$ to be a local base for $\mathcal{T}$ at x. We now verify the continuity of addition: Let $x, y \in X$ and $z := x + y$. If $U \in \mathcal{U}(z)$, then there exist $N, n \in \mathbb{N}$ and $i_1, \dots, i_N \in I$ such that $z \in B := \bigcap_{k=1}^{N} B_{p_{i_k}, \frac{1}{n}}(z) \subseteq U$. If

$$V_x := \bigcap_{k=1}^{N} B_{p_{i_k}, \frac{1}{2n}}(x), \quad V_y := \bigcap_{k=1}^{N} B_{p_{i_k}, \frac{1}{2n}}(y),$$

then

$$\forall_{(a,b) \in V_x \times V_y} \quad \forall_{k \in \{1,\dots,N\}} \quad p_{i_k}(a + b - (x + y)) \leq p_{i_k}(a - x) + p_{i_k}(b - y) < \frac{1}{2n} + \frac{1}{2n} = \frac{1}{n},$$

showing $a + b \in B \subseteq U$ and the continuity of addition. We proceed to the continuity of scalar multiplication: Let $x \in X$, $\lambda \in \mathbb{K}$, $z := \lambda x$. Given $U \in \mathcal{U}(z)$, let $B \subseteq U$ be as above. Let $R := 1 + \max\{p_{i_k}(x) : k \in \{1, \dots, N\}\} \in \mathbb{R}^+$, $\epsilon := 1/(2nR) \in \mathbb{R}^+$, $M \in \mathbb{N}$ such that $M \geq |\lambda| + \epsilon$,

$$V_x := \bigcap_{k=1}^{N} B_{p_{i_k}, \frac{1}{2nM}}(x).$$

Consider $(a, \alpha) \in V_x \times B_\epsilon(\lambda)$. Then $|\alpha| < M$ and, for each $k \in \{1, \dots, N\}$,

$$p_{i_k}(\alpha a - \lambda x) \leq p_{i_k}(\alpha a - \alpha x) + p_{i_k}(\alpha x - \lambda x) \leq M\, p_{i_k}(a - x) + \epsilon\, p_{i_k}(x) < \frac{M}{2nM} + \frac{R}{2nR} = \frac{1}{n},$$

showing $\alpha a \in B \subseteq U$ and the continuity of scalar multiplication. Due to Rem. 1.34(c) and Prop. 1.8(a), the elements of $\mathcal{B}_0$ are convex, showing $(X, \mathcal{T})$ to be locally convex. We now fix $i \in I$ and show p_i to be continuous: Let $x \in X$ be arbitrary, $\epsilon \in \mathbb{R}^+$, $n \in \mathbb{N}$ such that $n^{-1} < \epsilon$, $V := B_{p_i, n^{-1}}(0)$. Then $x + V \in \mathcal{U}(x)$ and, for each $y \in x + V \in \mathcal{U}(x)$, we have $y - x \in V$, i.e.

$$|p_i(y) - p_i(x)| \overset{\text{Th. 1.36(a)}}{\leq} p_i(y - x) \in [0, \epsilon[,$$

showing p_i to be continuous at x, proving (i). To prove (ii), let $E \subseteq X$ be bounded and fix $i \in I$. Since $B_{p_i,1}(0) \in \mathcal{U}(0)$, there exists $M_i \in \mathbb{R}^+$ such that $E \subseteq M_i B_{p_i,1}(0) = B_{p_i, M_i}(0)$, showing $p_i{\restriction}_E$ to be bounded by M_i. Conversely, suppose each p_i, $i \in I$, to be (w.l.o.g. strictly) bounded by some $M_i \in \mathbb{R}^+$ on E, i.e. $E \subseteq B_{p_i, M_i}(0)$. Let $U \in \mathcal{U}(0)$ and choose $B \subseteq U$ as before (with $z := 0$). If $s > n M_{i_k}$ for each $k \in \{1, \dots, N\}$, then $E \subseteq sB \subseteq sU$, showing E to be bounded, proving (ii). Assume $\mathcal{P}$ to be a separating family. We show $(X, \mathcal{T})$ to be T_1: If $x \in X \setminus \{0\}$, then there is $i \in I$ with $p_i(x) \neq 0$. Thus, there exists $n \in \mathbb{N}$ such that $x \notin B_{p_i, n^{-1}}(0) \in \mathcal{T}$, $0 \notin B_{p_i, n^{-1}}(x) \in \mathcal{T}$. Now, if $y, z \in X$ with $y \neq z$, and $U_1 \in \mathcal{U}(0)$, $U_2 \in \mathcal{U}(z-y)$ are such that $z-y \notin U_1$, $0 \notin U_2$, then $y+U_1 \in \mathcal{U}(y)$, $y+U_2 \in \mathcal{U}(z)$ with $z \notin y+U_1$ and $y \notin y+U_2$, showing $(X,\mathcal{T})$ to be T_1. Conversely, assume $(X,\mathcal{T})$ to be T_1 and let $x \in X \setminus \{0\}$. Then there exists $U \in \mathcal{U}(0)$ such that $x \notin U$. Once again, we let $B \subseteq U$ be as above. Since $x \notin B$, there exists $k \in \{1,\dots,N\}$ such that $x \notin B_{p_{i_k},n^{-1}}(0)$, i.e. $p_{i_k}(x) \neq 0$, proving $\mathcal{P}$ to be separating. If I is countable, then both $\mathcal{S}$ and $\mathcal{B}_0$ are countable, i.e. $(X,\mathcal{T})$ is C_1. Thus, if $(X,\mathcal{T})$ is also T_1, then it is metrizable by Th. 1.21. The proof that, for $I := \mathbb{N}$ and $(X,\mathcal{T})$ T_1, d as defined in (1.24) constitutes a metric on X that induces $\mathcal{T}$ with $\mathcal{D} = \{B_{d,r}(0) : r \in \mathbb{R}^+\}$ a convex balanced local base at 0, is left as an exercise.

(b): The inclusion $B_1 \subseteq B$ is due to Th. 1.37(h). For the remaining inclusion, let $x \in B$. As B is open, there exists $U \in \mathcal{U}(x)$ such that $U \subseteq B$. Since $1 \cdot x = x$, the continuity of scalar multiplication yields neighborhoods V of x and O of 1 such that $ty \in U \subseteq B$ for each $t \in O$, $y \in V$. In particular, there exists $t \in]0,1[$ such that $t^{-1}x \in B$, implying $\mu_B(x) < 1$ and $x \in B_1$, proving (1.25). Moreover, if the local base $\mathcal{B}$ for $\mathcal{T}$ at 0 consists of convex balanced sets, then $\mathcal{P}$ is a family of seminorms according to Th. 1.37(f), which induces a topology $\mathcal{T}_1$ on X according to (a), such that $(X,\mathcal{T}_1)$ is a topological vector space. We then also know from (a) that each μ_B, $B \in \mathcal{B}$, is $\mathcal{T}_1$-continuous and that $(X,\mathcal{T}_1)$ is T_1 if, and only if, $\mathcal{P}$ is separating. Thus, it remains to prove $\mathcal{T}_1 = \mathcal{T}$. If $B \in \mathcal{B}$, then, using the notation from (a) as well as (1.25), we obtain $B = B_{\mu_B,1}(0)$, showing $\mathcal{T} \subseteq \mathcal{T}_1$. On the other hand, if $n \in \mathbb{N}$ and $B \in \mathcal{B}$, then $nB_{\mu_B,\frac{1}{n}}(0) = B_{\mu_B,1}(0) = B \in \mathcal{T}$, showing $B_{\mu_B,\frac{1}{n}}(0) = \frac{1}{n}B \in \mathcal{T}$ and $\mathcal{T}_1 \subseteq \mathcal{T}$, concluding the proof. ■

Theorem 1.41. *Let $(X,\mathcal{T})$ be a T_1 topological vector space over $\mathbb{K}$. Then $(X,\mathcal{T})$ is normable if, and only if, there exists a bounded and convex $U \in \mathcal{U}(0)$ (i.e. if, and only if, $(X,\mathcal{T})$ is both locally bounded and locally convex).*

Proof. Let $\|\cdot\|$ be a norm on X that induces $\mathcal{T}$. Then each ball $B_r(0)$, $r \in \mathbb{R}^+$, is convex by Th. 1.36(c) and bounded by Th. 1.40(a)(ii). Conversely, let $U \in \mathcal{U}(0)$ be bounded and convex. Then, by Prop. 1.12(b), U contains some $V \in \mathcal{U}(0)$ that is bounded, convex, balanced, and open. Then $\|\cdot\| := \mu_V$ defines a norm on X: By Th. 1.37(f), μ_V is a seminorm. According to Th. 1.40(b), $B_{\|\cdot\|,1}(0) = V$, i.e. $B_{\|\cdot\|,1}(0)$ is bounded. Since $\mu_V^{-1}\{0\} \subseteq B_{\|\cdot\|,1}(0)$, $\mu_V^{-1}\{0\}$ is bounded. Since $\mu_V^{-1}\{0\}$ is also a vector space by Th. 1.36(b) and $(X,\mathcal{T})$ is T_1 by hypothesis, $\mu_V^{-1}\{0\} = \{0\}$ by Rem. 1.26 and, thus, μ_V is a norm by Cor. 1.38(b). Let $\mathcal{S}$ be the topology on X induced by $\|\cdot\|$. We still need to show $\mathcal{S} = \mathcal{T}$. Let $O \in \mathcal{T}$, $0 \in O$. Since $V = B_{\|\cdot\|,1}(0)$ is bounded, $V \subseteq sO$ for suitable $s > 0$, showing $s^{-1}V = B_{\|\cdot\|,s^{-1}}(0) \subseteq O$ and $\mathcal{T} \subseteq \mathcal{S}$. On the other hand, if $r \in \mathbb{R}^+$, then $B_{\|\cdot\|,r}(0) = rV \in \mathcal{T}$, showing $\mathcal{S} \subseteq \mathcal{T}$. ■

1.6 Further Examples

Example 1.42. We provide an example that shows that bounded linear maps between topological vector spaces need not be continuous. Let $X := C([0,1],\mathbb{K})$ be the vector space over $\mathbb{K}$ consisting of all $\mathbb{K}$-valued continuous functions on $[0,1]$. Define

$$d:\, X \times X \longrightarrow \mathbb{R}_0^+, \quad d(f,g) := \int_0^1 \frac{|f-g|}{1+|f-g|}\, \mathrm{d}\lambda^1 . \tag{1.26}$$

It is an exercise to show d constitutes a translation-invariant metric on X and that the induced topology $\mathcal{S}$ makes X into a topological vector space $(X,\mathcal{S})$. Next, define

$$\underset{t\in[0,1]}{\forall} \quad p_t:\, X \longrightarrow \mathbb{R}_0^+, \quad p_t(f) := |f(t)|. \tag{1.27}$$

If $f,g \in X$, $\lambda \in \mathbb{K}$, then $p_t(\lambda f) = |\lambda||f(t)| = |\lambda| p_t(f)$ and $p_t(f+g) \leq |f(t)| + |g(t)| = p_t(f) + p_t(g)$, showing p_t to define a seminorm on X. Thus, according to Th. 1.40(a), the family $(p_t)_{t\in[0,1]}$ induces a topology $\mathcal{T}$ on X such that $(X,\mathcal{T})$ is a topological vector space as well. We show $\mathrm{Id}:\, (X,\mathcal{T}) \longrightarrow (X,\mathcal{S})$ to be bounded, but not continuous: Let $E \subseteq X$ be $\mathcal{T}$-bounded. We show that E is also $\mathcal{S}$-bounded (i.e. $\mathrm{Id}:\, (X,\mathcal{T}) \longrightarrow (X,\mathcal{S})$ is bounded): According to Prop. 1.12(d), it suffices to consider the case that E is countable (and nonempty), say $E = \{e_k:\, k \in \mathbb{N}\}$. As E is bounded, Th. 1.40(a)(ii) implies

$$\underset{t\in[0,1]}{\forall} \quad M_t := \sup\{|e_k(t)|:\, k \in \mathbb{N}\} \in \mathbb{R}_0^+,$$

and we define

$$F:\, [0,1] \longrightarrow \mathbb{R}_0^+, \quad F(t) := M_t.$$

Then $F = \sup\{|e_k|:\, k \in \mathbb{N}\}$ and, as a sup of countably many λ^1-measurable functions, F is λ^1-measurable. In particular, the following integral makes sense (also note that the integrand is uniformly bounded by 1):

$$s := \int_0^1 \frac{F}{1+F}\, \mathrm{d}\lambda^1 .$$

Fix $r \in \mathbb{R}^+$. Since $\frac{F}{1+F} < 1$, we can apply the dominated convergence theorem (DCT, [Phi17, Th. 2.20]) to obtain $\lim_{\epsilon\to 0} \int_0^1 \frac{\epsilon F}{1+\epsilon F}\, \mathrm{d}\lambda^1 = 0$. In particular, we can fix $\epsilon := \epsilon(r) \in]0,1[$ such that $\int_0^1 \frac{\epsilon F}{1+\epsilon F}\, \mathrm{d}\lambda^1 < \frac{r}{2}$. Next, we note that, for $F > 0$ and $\epsilon \in]0,1[$,

$$\underset{k\in\mathbb{N}}{\forall} \quad \frac{|e_k|}{1+\epsilon|e_k|}\, \frac{1+\epsilon F}{F} < 2 \begin{cases} \text{for } 1 < F, \text{ since } \frac{|e_k|}{1+\epsilon|e_k|} < 1 \text{ and } \frac{1+F}{F} < 2 \Leftrightarrow 1 < F, \\ \text{for } F \leq 1, \text{ since } \frac{|e_k|}{F} \leq 1 \text{ and } \frac{1+\epsilon F}{1+\epsilon|e_k|} \leq 1 + \epsilon F < 2. \end{cases}$$

Thus, due to our choice of ϵ above, for each $k \in \mathbb{N}$,

$$d(\epsilon e_k, 0) = \int_0^1 \frac{\epsilon |e_k|}{1+\epsilon |e_k|} \,\mathrm{d}\lambda^1 = \int_{\{F>0\}} \frac{\epsilon |e_k|}{1+\epsilon |e_k|} \frac{1+\epsilon F}{F} \frac{F}{1+\epsilon F} \,\mathrm{d}\lambda^1$$
$$\leq 2 \int_0^1 \frac{\epsilon F}{1+\epsilon F} \,\mathrm{d}\lambda^1 < \frac{2r}{2} = r,$$

showing $\epsilon\, e_k \in B_{d,r}(0)$, i.e. $E \subseteq \epsilon^{-1} B_{d,r}(0)$ and E is $\mathcal{S}$-bounded. However, $\mathrm{Id} : (X, \mathcal{T}) \longrightarrow (X, \mathcal{S})$ is not continuous: We show that $B_{d,\frac{1}{4}}(0) \notin \mathcal{T}$: Otherwise, according to Th. 1.40(a), there must exist $N \in \mathbb{N}$, $\epsilon \in \mathbb{R}^+$, and $t_1, \dots, t_N \in [0,1]$ such that $0 \leq t_1 < \dots < t_N \leq 1$ and $O := \bigcap_{k=1}^N B_{p_{t_k},\epsilon}(0) \subseteq B_{d,\frac{1}{4}}(0)$. Seeking a contradiction, we will show that there must exist some $f \in O \setminus B_{d,\frac{1}{4}}(0)$: The idea is to construct a continuous f such that f is constantly equal to 1 everywhere, except in sufficiently small neighborhoods to the t_k, where it decreases affinely and sufficiently fast to have $f(t_k) = 0$ for each $k \in \{1, \dots, N\}$. To avoid special considerations for $k = 1$ and $k = N$, we first define $g : \mathbb{R} \longrightarrow \mathbb{R}_0^+$ and then let $f := g{\restriction}_{[0,1]}$. In preparation for the definition of g, let $\delta_1 := \frac{1}{2N}$,

$$\delta_2 := \begin{cases} \min\Big\{ t_{k+1} - t_k : \, k \in \{1, \dots, N-1\} \Big\} & \text{for } N \geq 2, \\ \delta_1 & \text{for } N = 1, \end{cases}$$

and $\delta := \min\{\delta_1, \delta_2\}$. We now let

$$g : \mathbb{R} \longrightarrow \mathbb{R}_0^+, \quad g(s) := \begin{cases} -\frac{2s}{\delta} + \frac{2t_k}{\delta} & \text{for } t_k - \frac{\delta}{2} \leq s \leq t_k, \\ \frac{2s}{\delta} - \frac{2t_k}{\delta} & \text{for } t_k \leq s \leq t_k + \frac{\delta}{2}, \\ 1 & \text{otherwise.} \end{cases}$$

In particular, note g to be continuous and piecewise affine with $g(\mathbb{R}) = [0,1]$ and, for each $k \in \{1, \dots, N\}$, $g(t_k - \frac{\delta}{2}) = g(t_k + \frac{\delta}{2}) = 1$, $g(t_k) = 0$. Moreover, letting, for each $k \in \{1, \dots, N\}$, $I_k :=]t_k - \frac{\delta}{2}, t_k + \frac{\delta}{2}[$, we have $I_k \cap I_l = \emptyset$ for $k \neq l$. Letting $I := \bigcup_{k=1}^N I_k$, one has $g{\restriction}_{\mathbb{R}\setminus I} \equiv 1$ and

$$\lambda^1(I) = N \cdot \delta \leq N\delta_1 = \frac{N}{2N} = \frac{1}{2} \quad \Rightarrow \quad \lambda^1\big([0,1] \setminus I\big) \geq \frac{1}{2}.$$

Letting $f := g{\restriction}_{[0,1]}$, we obtain $f \in X$ and, since $f(t_k) = 0$ for each $k \in \{1, \dots, N\}$, we also have $f \in O = \bigcap_{k=1}^N B_{p_{t_k},\epsilon}(0)$. However,

$$d(f, 0) = \int_0^1 \frac{f}{1+f} \,\mathrm{d}\lambda^1 > \int_{[0,1]\setminus I} \frac{f}{1+f} \,\mathrm{d}\lambda^1 = \frac{1}{2} \cdot \lambda^1\big([0,1] \setminus I\big) \geq \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4},$$

showing $f \notin B_{d,\frac{1}{4}}(0)$ and $B_{d,\frac{1}{4}}(0) \notin \mathcal{T}$.

We can now conclude from Th. 1.32 that $(X,\mathcal{T})$ is not metrizable (since the linear map $\mathrm{Id}: (X,\mathcal{T}) \longrightarrow (X,\mathcal{S})$ is bounded, but not continuous). Another way to see that $(X,\mathcal{T})$ is not metrizable is to show $\mathrm{Id}: (X,\mathcal{T}) \longrightarrow (X,\mathcal{S})$ to be sequentially continuous: Indeed, if $(f_k)_{k\in\mathbb{N}}$ is a sequence in X a and $f \in X$ such that $\lim_{k\to\mathbb{N}} f_k = f$ with respect to $\mathcal{T}$, then

$$\underset{t\in[0,1]}{\forall} \quad \lim_{k\to\mathbb{N}} |f-f_k|(t) = \lim_{k\to\mathbb{N}} p_t(f-f_k) \overset{p_t \text{ cont.}}{=} p_t(0) = 0,$$

i.e. the $|f-f_k|$ converge pointwise to 0. Since $\frac{|f-f_k|}{1+|f-f_k|} < 1$ and 1 is integrable over $[0,1]$, we can use DCT to obtain

$$\lim_{k\to\infty} d(f_k,f) = \lim_{k\to\infty} \int_0^1 \frac{|f-f_k|}{1+|f-f_k|}\,\mathrm{d}\lambda^1 \overset{\text{DCT}}{=} 0,$$

showing $\lim_{k\to\mathbb{N}} f_k = f$ with respect to $\mathcal{S}$ and sequential continuity of $\mathrm{Id}: (X,\mathcal{T}) \longrightarrow (X,\mathcal{S})$. In consequence, by [Phi16b, Th. 2.8], $(X,\mathcal{T})$ is not metrizable.

Example 1.43. Let $n \in \mathbb{N}$, and let $O \subseteq \mathbb{R}^n$ be open, $X := C(O,\mathbb{K})$. Let $(K_i)_{i\in\mathbb{N}}$ be an exhaustion by compact sets of O, i.e. a sequence of compact subsets of O such that

$$O = \bigcup_{i\in\mathbb{N}} K_i, \tag{1.28a}$$

$$\underset{i\in\mathbb{N}}{\forall} \quad K_i \subseteq K_{i+1}^\circ \tag{1.28b}$$

(cf. Th. A.2 of the Appendix). Clearly, we may also assume, in addition, $K_i \neq \emptyset$ for each $i \in \mathbb{N}$. We define

$$\underset{i\in\mathbb{N}}{\forall} \quad p_i: X \longrightarrow \mathbb{R}_0^+, \quad p_i(f) := \sup\{|f(x)| : x \in K_i\} = \|f{\restriction}_{K_i}\|_\infty. \tag{1.29}$$

Clearly, $\mathcal{F} := (p_i)_{i\in\mathbb{N}}$ constitutes a family of seminorms on X. Thus, according to Th. 1.40(a), $\mathcal{F}$ induces a topology $\mathcal{T}$ on X such that $(X,\mathcal{T})$ is a locally convex topological vector space. Due to (1.28b), we have $p_1 \leq p_2 \leq \dots$ and, thus,

$$\{B_{p_i,\frac{1}{i}}(0) : i \in \mathbb{N}\},$$

forms a convex balanced local base for $\mathcal{T}$ at 0. Moreover, $\mathcal{F}$ is separating due to (1.28a), implying $(X,\mathcal{T})$ to be T_1. Thus, using Th. 1.40(a) once again, $(X,\mathcal{T})$ is metrizable by the metric

$$d: X\times X \longrightarrow \mathbb{R}_0^+, \quad d(f,g) := \max\left\{\frac{2^{-i}\,p_i(f-g)}{1+p_i(f-g)} : i \in \mathbb{N}\right\}. \tag{1.30}$$

We show d to be complete: Let $(f_k)_{k\in\mathbb{N}}$ be a Cauchy sequence in X (with respect to d). Then, for each $i \in \mathbb{N}$, $(f_k\restriction_{K_i})_{k\in\mathbb{N}}$ is Cauchy with respect to $\|\cdot\|_\infty$ (check it!) and, thus, converges (uniformly) to some continuous $F_i : K_i \longrightarrow \mathbb{K}$. Due to (1.28b), $F_j\restriction_{K_i} = F_i$ for $j \geq i$ and $F : O \longrightarrow \mathbb{K}$, $F(x) := F_i(x)$ for $x \in K_i$ well-defines a function $F \in X$. Since $\lim_{k\to\infty} p_i(f_k - F) = \lim_{k\to\infty} p_i(f_k - F_i) = 0$ for each $i \in \mathbb{N}$. This implies $\lim_{k\to\infty} d(f_k, F) = 0$: Given $\epsilon \in \mathbb{R}^+$, choose $i_0 \in \mathbb{N}$ such that $2^{-i_0} < \epsilon$. Then choose $N \in \mathbb{N}$ such that, for each $k > N$ and each $i < i_0$, $2^{-i}\, p_i(f_k - F) < \epsilon$. Then,

$$\underset{k>N}{\forall} \quad d(f_k, F) = \max\left\{\frac{2^{-i}\, p_i(f_k - F)}{1 + p_i(f_k - F)} : i \in \mathbb{N}\right\} < \epsilon,$$

showing $\lim_{k\to\infty} d(f_k, F) = 0$, i.e. d is complete and $(X, \mathcal{T})$ is Fréchet. However, $(X, \mathcal{T})$ is not locally bounded (and, thus, by Th. 1.41, not normable): It suffices to show that no $B_{p_i,\epsilon}(0)$, $i \in \mathbb{N}$, $\epsilon \in \mathbb{R}^+$, is bounded. Let $0 \neq f \in C_c(K_{i+1}^\circ \setminus K_i)$ (cf. [Phi17, Th. 2.49(a)]). Then $p_{i+1}(rf) = rp_{i+1}(f) > 0$ and $p_i(rf) = 0$ (i.e. $rf \in B_{p_i,\epsilon}(0)$ for each $\epsilon \in \mathbb{R}^+$) for each $r \in \mathbb{R}^+$, showing p_{i+1} is not bounded on $B_{p_i,\epsilon}(0)$, i.e. $B_{p_i,\epsilon}(0)$ is not bounded by Th. 1.40(a)(ii). We conclude the example by showing that $\mathcal{T}$ does not depend on the chosen exhaustion by compact sets of O: Suppose $(\tilde{K}_i)_{i\in\mathbb{N}}$ is also an exhaustion by compact sets of O with resulting seminorms $\tilde{p}_i$, metric $\tilde{d}$, and induced topology $\tilde{\mathcal{T}}$. Inductively, we define a sequence of indices $(j_i)_{i\in\mathbb{N}}$ as follows: Given $i \in \mathbb{N}$, since $\tilde{K}_i$ is compact and $(K_j^\circ)_{j\in\mathbb{N}}$ is an open cover of $\tilde{K}_i$, there exists $j(i) \in \mathbb{N}$ such that $\tilde{K}_i \subseteq K_{j(i)}$. If $i = 1$, set $j_i := j(i)$; if $i > 1$, set $j_i := \max\{j_{i-1} + 1, j(i)\}$. Then $(j_i)_{i\in\mathbb{N}}$ is strictly increasing and $\tilde{K}_i \subseteq K_{j_i}$ for each $i \in \mathbb{N}$. For each $i \in \mathbb{N}$ and $f \in X$, we obtain

$$p_{j_i}(f) > 0 \;\Rightarrow\; \frac{\tilde{p}_i(f)}{1 + \tilde{p}_i(f)}\,\frac{1 + p_{j_i}(f)}{p_{j_i}(f)} \leq 2 \begin{cases} \text{for } 1 < p_{j_i}(f), \text{ since } \frac{\tilde{p}_i(f)}{1+\tilde{p}_i(f)} < 1 \\ \qquad \text{and } \frac{1+p_{j_i}(f)}{p_{j_i}(f)} < 2 \;\Leftrightarrow\; 1 < p_{j_i}(f), \\ \text{for } p_{j_i}(f) \leq 1, \text{ since } \frac{\tilde{p}_i(f)}{p_{j_i}(f)} \leq 1 \\ \qquad \text{and } \frac{1+p_{j_i}(f)}{1+\tilde{p}_i(f)} \leq 1 + p_{j_i}(f) \leq 2, \end{cases}$$

and, thus,

$$\frac{2^{-i}\,\tilde{p}_i(f)}{1 + \tilde{p}_i(f)} = \frac{2^{-i}\,\tilde{p}_i(f)}{1 + \tilde{p}_i(f)}\,\frac{1 + p_{j_i}(f)}{p_{j_i}(f)}\,\frac{p_{j_i}(f)}{1 + p_{j_i}(f)} \leq \frac{2^{-i+1+j_i}\, 2^{-j_i}\, p_{j_i}(f)}{1 + p_{j_i}(f)}.$$

In consequence, if $\delta \in \mathbb{R}^+$ and $f \in B_{d,\delta}(0)$, then, for each $l \in \mathbb{N}$, $l > 1$,

$$\tilde{d}(f, 0) \leq \max\left\{2^{-l}, \max\left\{\frac{2^{-i}\,\tilde{p}_i(f)}{1 + \tilde{p}_i(f)} : i < l\right\}\right\} \leq \max\{2^{-l}, c_l\,\delta\}, \tag{1.31}$$

where $c_l := \max\{2^{-i+1+j_i} : i < l\}$. Thus, given $\epsilon \in \mathbb{R}^+$, we may choose $l \in \mathbb{N}$ such that $2^{-l} < \epsilon$ and then $\delta > 0$ such that $c_l\,\delta < \epsilon$ as well. Then (1.31) shows that $B_{d,\delta}(0) \subseteq B_{\tilde{d},\epsilon}(0)$, implying $\tilde{\mathcal{T}} \subseteq \mathcal{T}$. Since the previous argument works exactly the same with the roles of d and $\tilde{d}$ reversed, we obtain $\tilde{\mathcal{T}} = \mathcal{T}$ as desired.

Example 1.44. Let $n \in \mathbb{N}$, and let $O \subseteq \mathbb{R}^n$ be open, $X := C^\infty(O, \mathbb{K})$. Moreover, for each compact $K \subseteq O$, we introduce the notation

$$\mathcal{D}_K := \mathcal{D}_K(O, \mathbb{K}) := \left\{ f \in C^\infty(O, \mathbb{K}) : \operatorname{supp} f \subseteq K \right\}, \tag{1.32}$$

where we recall from [Phi17, Def. 2.48(a)] that the *support* of f is defined by

$$\operatorname{supp} f := \overline{\{x \in O : f(x) \neq 0\}}.$$

The spaces $\mathcal{D}_K$ play an important role in the theory of so-called *distributions* (see, e.g., [Rud73, Ch. 6] – we will also make some further remarks on distributions below). Clearly, each $\mathcal{D}_K$ is a vector subspace of X over $\mathbb{K}$. Applying a procedure that is similar to the one used in the previous Ex. 1.43, we make X into a metrizable topological vector space: We start with a sequence $(K_i)_{i \in \mathbb{N}}$, constituting an exhaustion by compact sets of O, i.e. such that (1.28) holds. As in Ex. 1.43, we also assume $K_i \neq \emptyset$ for each $i \in \mathbb{N}$. This time we define a family of seminorms $\mathcal{F} := (p_i)_{i \in \mathbb{N}}$ by letting, for each $i \in \mathbb{N}$,

$$\begin{aligned} p_i : X &\longrightarrow \mathbb{R}_0^+, \\ p_i(f) &:= \sup \Big\{ |\partial_p f(x)| : x \in K_i,\, p = \emptyset \text{ or } p = (p_1, \dots, p_j) \in \{1, \dots, n\}^j, 1 \leq j \leq i \Big\} \\ &= \max \Big\{ \|(\partial_p f){\restriction_{K_i}}\|_\infty : p = \emptyset \text{ or } p \in \{1, \dots, n\}^j, 1 \leq j \leq i \Big\}, \end{aligned} \tag{1.33}$$

where $\partial_\emptyset f := f$. According to Th. 1.40(a), $\mathcal{F}$ induces a topology $\mathcal{T}$ on X such that $(X, \mathcal{T})$ is a locally convex topological vector space. Due to (1.28b) and (1.33), we have $p_1 \leq p_2 \leq \dots$ and, thus,

$$\{B_{p_i, \frac{1}{i}}(0) : i \in \mathbb{N}\},$$

forms a convex balanced local base for $\mathcal{T}$ at 0. Moreover, $\mathcal{F}$ is separating due to (1.28a), implying $(X, \mathcal{T})$ to be T_1. Thus, as in Ex. 1.43, $(X, \mathcal{T})$ is metrizable by the metric defined in (1.30). Once again, we can show d to be complete: Let $(f_k)_{k \in \mathbb{N}}$ be a Cauchy sequence in X (with respect to d). Then, for each $i \in \mathbb{N}$ and each $p \in \{1, \dots, n\}^N$, $N \in \mathbb{N}$, $(f_k{\restriction_{K_i}})_{k \in \mathbb{N}}$ and $((\partial_p f_k){\restriction_{K_i}})_{k \in \mathbb{N}}$ are Cauchy with respect to $\|\cdot\|_\infty$ (check it!) and, thus, converge (uniformly) to some continuous $F_i : K_i \longrightarrow \mathbb{K}$ and continuous $F_{i,p} : K_i \longrightarrow \mathbb{K}$, respectively. Due to (1.28b), $F_j{\restriction_{K_i}} = F_i$ and $F_{j,p}{\restriction_{K_i}} = F_{i,p}$ for $j \geq i$, such that $F : O \longrightarrow \mathbb{K}$, $F(x) := F_i(x)$ and $F_p : O \longrightarrow \mathbb{K}$, $F_p(x) := F_{i,p}(x)$ for $x \in K_i$ well-define functions $F, F_p \in C(O, \mathbb{K})$. Since the convergences $f_k \to F$ and $(\partial_p f_k) \to F_p$ for each $p \in \{1, \dots, n\}^N$, $N \in \mathbb{N}$, are uniform on each K_i, we can inductively apply Th. B.1 to all partial derivatives of the f_k to obtain $F_p = \partial_p F$ and $F \in X$. Moreover, $\lim_{k\to\infty} p_i(f_k - F) = \lim_{k\to\infty} p_i(f_k - F_i) = 0$ for each $i \in \mathbb{N}$, implies $\lim_{k\to\infty} d(f_k, F) = 0$ as in Ex. 1.43, proving d to be complete and $(X, \mathcal{T})$ to be Fréchet. For each $x \in O$, $e_x : X \longrightarrow \mathbb{K}$, $e_x(f) := f(x)$, constitutes a continuous linear functional (since convergence in $(X, \mathcal{T})$ implies pointwise convergence). Thus, $\ker e_x$ is closed and, hence, so is

$$\mathcal{D}_K = \bigcap_{x \in O \setminus K} \ker e_x \tag{1.34}$$

for each compact $K \subseteq O$, showing the $\mathcal{D}_K$ to be Fréchet as well. As in Ex. 1.43, it follows that $(X, \mathcal{T})$ is not locally bounded (and, thus, by Th. 1.41, not normable): It suffices to show that no $B_{p_i,\epsilon}(0)$, $i \in \mathbb{N}$, $\epsilon \in \mathbb{R}^+$, is bounded. Let $0 \neq f \in C_c^\infty(K_{i+1}^\circ \setminus K_i)$ (cf. [Phi17, Th. 2.54(d)]). Then $p_{i+1}(rf) = rp_{i+1}(f) > 0$ and $p_i(rf) = 0$ (i.e. $rf \in B_{p_i,\epsilon}(0)$ for each $\epsilon \in \mathbb{R}^+$) for each $r \in \mathbb{R}^+$, showing p_{i+1} is not bounded on $B_{p_i,\epsilon}(0)$, i.e. $B_{p_i,\epsilon}(0)$ is not bounded by Th. 1.40(a)(ii). Finally, that $\mathcal{T}$ does not depend on the chosen exhaustion by compact sets of O follows precisely as in Ex. 1.43. Distributions are linear functionals on the space

$$\mathcal{D} := \mathcal{D}(O, \mathbb{K}) := \bigcup_{K \subseteq O \text{ compact}} \mathcal{D}_K(O, \mathbb{K}) \tag{1.35}$$

that are continuous with respect to a suitable topology $\mathcal{S}$ on $\mathcal{D}$. For technical reasons, one chooses $\mathcal{S}$ different from the subspace topology $\mathcal{T}_{\mathcal{D}}$, where $\mathcal{S}$ is actually nonmetrizable (however, $\mathcal{T}_{\mathcal{D}_K} = \mathcal{S}_{\mathcal{D}_K}$ for each compact $K \subseteq O$, see [Rud73, Sec. 6.2–6.9] for details).

2 Main Theorems

2.1 Baire Category

Recall that a subset A of a topological space X is called *dense* if, and only if, $\overline{A} = X$. The concept of Baire category can be seen as a refinement of the concept of denseness (cf. Def. 2.1 below). The main result in this section is the so-called *Baire category theorem* (Th. 2.6). Its main applications are abstract existence proofs. As applications, we will prove the existence of continuous maps that are nowhere differentiable (Ex. 2.8) and the existence of points of continuity for pointwise limits of continuous functions as well as for derivatives (Th. 2.11).

Definition 2.1. Let $(X, \mathcal{T})$ be a topological space, $A \subseteq X$.

(a) A is called *nowhere dense* in X if, and only if, $\overline{A}$ has empty interior, i.e. if, and only if, $(\overline{A})^\circ = \emptyset$.

(b) A is said to be of the *first category* in X or *meager* if, and only if, $A = \bigcup_{k=1}^\infty A_k$ is a countable union of nowhere dense sets A_k, $k \in \mathbb{N}$.

(c) A is said to be of the *second category* in X or *nonmeager* or *fat* if, and only if, A is not of the first category in X.

Caveat: These notions of category (due to Baire) are completely different from the notion of category occurring in the more algebraic discipline called category theory.

Lemma 2.2. *Let $(X, \mathcal{T})$ be a topological space.*

(a) *$A \subseteq X$ is nowhere dense if, and only if, $(\overline{A})^{\mathrm{c}}$ is dense.*

(b) *If $A \subseteq B \subseteq X$ and B is nowhere dense (resp. of the first category), then so is A.*

(c) *Every countable union of sets of the first category in X is of the first category in X.*

(d) *If $A \subseteq X$ is closed and $A^\circ = \emptyset$, then A is nowhere dense (and, in particular, of the first category in X).*

(e) *The notions nowhere dense, of the first category, and of the second category are topological invariants, i.e. they remain invariant under homeomorphisms.*

Proof. (a): For each $B \subseteq X$, we have the disjoint union $X = B^\circ \,\dot\cup\, \partial B \,\dot\cup (B^{\mathrm{c}})^\circ$. Applying this to $\overline{A}$ yields $X = (\overline{A})^\circ \,\dot\cup\, \partial(\overline{A}) \,\dot\cup ((\overline{A})^{\mathrm{c}})^\circ$. Since also $\overline{B} = B^\circ \,\dot\cup\, \partial B$, we obtain

$$(\overline{A})^\circ = \emptyset \;\Leftrightarrow\; X = \overline{(\overline{A})^{\mathrm{c}}}$$

as claimed.

(b): The part regarding "nowhere dense" holds, since, for each $C \subseteq D \subseteq X$, one has $\overline{C} \subseteq \overline{D}$ and $C^\circ \subseteq D^\circ$. The part regarding "of the first category" holds, since $B = \bigcup_{k=1}^\infty B_k$ and $A \subseteq B$ implies $A = \bigcup_{k=1}^\infty (A \cap B_k)$.

(c) is due to the fact that countable unions of countable sets are countable.

(d): Since A is closed, we have $(\overline{A})^\circ = A^\circ = \emptyset$.

(e): Let $(Y, \mathcal{S})$ be another topological space and $\phi : X \longrightarrow Y$ a homeomorphism. If $A \subseteq X$, then

$$(\overline{A})^\circ = \emptyset \;\Leftrightarrow\; (\overline{\phi(A)})^\circ = \phi\Big((\overline{A})^\circ\Big) = \emptyset,$$
$$A = \bigcup_{k=1}^\infty A_k \;\Leftrightarrow\; \phi(A) = \bigcup_{k=1}^\infty \phi(A_k),$$

proving (e). ∎

The idea is that sets of the first category are somehow "small" and sets of the second category are somehow "large" (e.g. in the sense that complements of sets of the first category must be dense in many spaces as we will see in Th. 2.6(c)). However, one has to use care, as sets of the first category can still be “large” in other ways: For example, they can themselves be dense (see Ex. 2.3(a)) or of full measure in a measure space (see Ex. 2.3(c),(d)).

Example 2.3. **(a)** $\mathbb{Q}$ is a countable dense subset of $\mathbb{R}$. Thus $\mathbb{Q}$ is both dense and of the first category in $\mathbb{R}$. More generally, in spaces where point sets $\{x\}$ are closed, but not open, countable sets are always of the first category, but many spaces (such as $\mathbb{K}^n$) have countable dense sets (we have previously called such spaces *separable* in Analysis II/III).

(b) Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{K}$ and let Y be a vector subspace. If $Y \neq X$, then Y has empty interior (in particular, Y is either dense or nowhere dense): Let $x \in X \setminus Y$ and $U \in \mathcal{U}(0)$. Then there exists $n \in \mathbb{N}$ such that $x \in nU$, i.e. $n^{-1}x \in U \setminus Y$, showing 0 not to be an interior point of Y. Since translations are homeomorphisms in X, no $y \in Y$ can be an interior point of Y. Now, if Y is not dense, then, by Prop. 1.10(c), $\overline{Y}$ is a proper subspace of X. Then, as we have just shown, $\overline{Y}$ has empty interior, and Y is nowhere dense.

(c) Let $n \in \mathbb{N}$. It is an exercise to show every Lebesgue-measurable set $M \subseteq \mathbb{R}^n$ (i.e. each $M \in \mathcal{L}^n$) can be written as the disjoint union $M = N \,\dot\cup\, A$, where N is a λ^n-null set and A is of the first category in $\mathbb{R}^n$ (use that $\mathbb{Q}^n$ is dense in $\mathbb{R}^n$ together with a geometric series). Moreover, if M is any nontrivial interval in $\mathbb{R}^n$ (i.e. $M^\circ \neq \emptyset$), then A is of the first category in M (and, then, Th. 2.6(c) below implies N to be of the second category in M).

(d) Another way to obtain sets that are of full measure, but of the first category, is to adapt the Cantor set construction of [Phi17, Sec. 1.5.3]: In [Phi17, Sec. 1.5.3], the Cantor set was what was left from $[0,1]$ after successively removing the (open) middle third from $[0,1]$, then the (open) middle thirds of the remaining intervals etc. Now, consider what occurs if, instead of removing 2^{n-1} intervals of length $(\frac{1}{3})^n$ in step n, we remove 2^{n-1} intervals of length $\epsilon(\frac{1}{4})^n$, $0 < \epsilon \leq 1$, in step n: Then the total length of intervals removed is

$$L_\epsilon = \epsilon \sum_{n=1}^{\infty} \frac{2^{n-1}}{2^{2n}} = \epsilon \left(\frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \dots \right) = \frac{\epsilon}{2}.$$

Thus, the resulting Cantor set C_ϵ has measure $\lambda^1(C_\epsilon) = 1 - \frac{\epsilon}{2}$. As in [Phi17, Prop. 1.63], one still finds each C_ϵ to be compact with empty interior (i.e. nowhere dense). Now, if $O_\epsilon := C_\epsilon^{\mathrm{c}} = [0,1] \setminus C_\epsilon$, $B := \bigcap_{n\in\mathbb{N}} O_{1/n}$, $D := B^{\mathrm{c}} = \bigcup_{n\in\mathbb{N}} C_{1/n}$, then $\lambda^1(B) = 0$, $\lambda^1(D) = 1$, and D is of the first category in $[0,1]$.

(e) We will see in Ex. 2.23 below that, for $p, q \in [1, \infty]$ with $p < q \leq \infty$, $L^q([0,1], \mathcal{L}^1, \lambda^1)$ is of the first category in $L^p([0,1], \mathcal{L}^1, \lambda^1)$.

—

The following Baire Category Th. 2.6 holds for complete pseudometric spaces as well as for locally compact Hausdorff spaces. In preparation for the proof of the variant for locally compact Hausdorff spaces, we provide the following two propositions:

Proposition 2.4. *Let the topological space $(X, \mathcal{T})$ be T_2.*

(a) *If $(K_i)_{i \in I}$, $I \neq \emptyset$, is a family of compact subsets of X such that $\bigcap_{i \in I} K_i = \emptyset$, then there exist $i_1, \dots, i_N \in I$, $N \in \mathbb{N}$, such that $\bigcap_{k=1}^{N} K_{i_k} = \emptyset$.*

(b) *One can separate points from compact sets: Let $x \in X$, $K \subseteq X$, K compact. If $x \notin K$, then there exist open sets $O_x \subseteq X$ and $O_K \subseteq X$ such that*

$$x \in O_x \ \wedge \ K \subseteq O_K \ \wedge \ O_x \cap O_K = \emptyset.$$

Proof. (a): Since $(X, \mathcal{T})$ is T_2, each K_i, $i \in I$, is closed by [Phi16b, Prop. 3.14(b)]. Then (a) follows, since each compact set has the finite intersection property by [Phi16b, Th. 3.13(ii)].

(b): Let $K \subseteq X$ be compact and $x \in X \setminus K$. Since $(X, \mathcal{T})$ is T_2, for each $a \in K$, there exists open $O_a \in \mathcal{U}(x)$ and open $U_a \in \mathcal{U}(a)$ such that $O_a \cap U_a = \emptyset$. Now $(U_a)_{a \in K}$ is an open cover of K. Since K is compact, there exists a finite set $M \subseteq K$ such that $(U_a)_{a \in M}$ still covers K: $K \subseteq O_K := \bigcup_{a \in M} U_a \in \mathcal{T}$. On the other hand $x \in O_x := \bigcap_{a \in M} O_a \in \mathcal{T}$. Since $O_x \cap O_K = \emptyset$, this proves (b). ■

Proposition 2.5. *Let the topological space $(X, \mathcal{T})$ be locally compact and T_2. Then the following holds:*

(a) *If $O, K \subseteq X$ such that O is open, K is compact, and $K \subseteq O$, then there exists an open set $V \subseteq X$ such that $\overline{V}$ is compact and*

$$K \subseteq V \subseteq \overline{V} \subseteq O. \tag{2.1}$$

(b) *$(X, \mathcal{T})$ is T_3 (and, thus, regular).*

Proof. Exercise. ■

Theorem 2.6 (Baire Category Theorem)**.** *Let $(X, \mathcal{T})$ be a topological space, $X \neq \emptyset$, and suppose at least one of the following two hypotheses holds:*

(i) $\mathcal{T}$ *is induced by a complete pseudometric on* X.

(ii) $(X, \mathcal{T})$ *is a locally compact Hausdorff space.*

Then, the following conclusions hold as well:

(a) *If* $(O_k)_{k\in\mathbb{N}}$ *is a sequence of dense open subsets of* X *and*

$$B := \bigcap_{k=1}^{\infty} O_k, \tag{2.2}$$

then B *is dense in* X *as well.*

(b) *If* $(A_k)_{k\in\mathbb{N}}$ *is a sequence of closed subsets of* X *with empty interior, then*

$$C := \bigcup_{k=1}^{\infty} A_k \tag{2.3}$$

has empty interior as well.

(c) *If* $A \subseteq X$ *is of the first category in* X*, then* A^c *is dense in* X*. In particular,* X *is of the second category in itself.*

Proof. (a): Let $V_0 \subseteq X$ be open and nonempty. According to [Phi16b, Prop. 1.35(e)], we need to show $V_0 \cap B \neq \emptyset$. To this end, we construct a sequence $(V_k)_{k\in\mathbb{N}_0}$ of nonempty open subsets of X such that

$$\underset{k\in\mathbb{N}}{\forall} \quad \overline{V}_k \subseteq V_{k-1} \cap O_k. \tag{2.4}$$

Inductively, assume that, for $l \in \mathbb{N}_0$, $l < k$, V_l have already been constructed in accordance with (2.4). The set $V_{k-1} \cap O_k$ is open and nonempty (as O_k is dense) and we choose $x_k \in V_{k-1} \cap O_k$. In Case (i), we now choose $\epsilon_k \in \mathbb{R}^+$ such that $\epsilon_k \leq \frac{1}{k}$ and $\overline{B}_{\epsilon_k}(x_k) \subseteq V_{k-1} \cap O_k$, letting $V_k := B_{\epsilon_k}(x_k)$. Then (2.4) is satisfied. In Case (ii), we use Prop. 2.5(a) to choose an open V_k such that $\{x_k\} \subseteq V_k \subseteq \overline{V}_k \subseteq V_{k-1} \cap O_k$ with $\overline{V}_k$ compact. Then, once again, (2.4) is satisfied. We claim

$$D := \bigcap_{k\in\mathbb{N}} \overline{V}_k \neq \emptyset : \tag{2.5}$$

In Case (i), $l \geq k$ implies $x_l \in B_{\epsilon_k}(x_k)$ for each $k, l \in \mathbb{N}$, such that $(x_k)_{k\in\mathbb{N}}$ constitutes a Cauchy sequence due to $\lim_{k\to\infty} \epsilon_k = 0$. The assumed completeness of X provides a limit $x = \lim_{k\to\infty} x_k \in X$. The nested form (2.4) of the V_k implies $x \in \overline{V}_k$ for each

$k \in \mathbb{N}$, proving (2.5). In Case (ii), (2.5) holds, since the compact set $\overline{V}_1$ has the finite intersection property (cf. [Phi16b, Th. 3.13(ii)]) and

$$\underset{k\in\mathbb{N}}{\forall} \quad \emptyset \neq \overline{V}_k = \bigcap_{l\leq k} \overline{V}_l.$$

In consequence of (2.4), we have $D \subseteq B \cap V_0$, showing $V_0 \cap B \neq \emptyset$ as needed.

(b) follows from (a) by taking complements: We rewrite (2.3) as

$$B := C^{\mathrm{c}} = \left(\bigcup_{k=1}^{\infty} A_k\right)^{\mathrm{c}} = \bigcap_{k=1}^{\infty} A_k^{\mathrm{c}}.$$

The $O_k := A_k^{\mathrm{c}}$ are open, since the A_k are closed. Since the A_k have empty interior, the O_k are dense. Then B is dense by (a) and C has empty interior.

(c): If A is of the first category in X, then $A = \bigcup_{k\in\mathbb{N}} A_k$, with each $A_k \subseteq X$ being nowhere dense. Then $A \subseteq C = \bigcup_{k\in\mathbb{N}} \overline{A}_k$, where each $\overline{A}_k$ is a closed set with empty interior. By (b), C (and, thus, A) must have empty interior as well. Thus, A^{c} is dense. ■

Remark 2.7. Typical applications of the Baire Th. 2.6 are (nonconstructive) existence proofs of the following from: To show a space X contains elements having the property P, one shows X to satisfy the hypotheses of Th. 2.6 and one shows the set A of elements in X *not* having the property P to be of the first category in X. Then A^{c} must be nonempty (and even dense). The following Ex. 2.8, provides an illustration of this method.

Example 2.8. In [Phi16a, Sec. J.1], one can find the construction of functions $f : \mathbb{R} \longrightarrow \mathbb{R}$ that are continuous, but nowhere differentiable. Using the Baire Th. 2.6 together with the Weierstrass approximation theorem (provided as Th. C.1 in the Appendix), we can now show that in $C[a,b]$, where $a, b \in \mathbb{R}$ with $a < b$, the set B of nowhere differentiable functions is even dense (and the complement $D := C[a,b] \setminus B$ is of the first category): If we equip $C[a,b]$ with the max-norm $\|\cdot\|_\infty$, then we know it to be a Banach space (a closed subspace of the Banach space $L^\infty([a,b], \mathcal{L}^1, \lambda^1)$). Define

$$\underset{n\in\mathbb{N}}{\forall} \quad O_n := \left\{ f \in C[a,b] : \underset{x\in[a,b]}{\forall} \ \sup\left\{ \left| \frac{f(x+h) - f(x)}{h} \right| : 0 < |h| \leq \frac{1}{n} \right\} > n \right\} \tag{2.6}$$

(to make sure $f(x+h)$ in (2.6) is always well-defined, we extend $f \in C[a,b]$ constantly by $f(a)$ to the left and constantly by $f(b)$ to the right). Let $B_0 := \bigcap_{n\in\mathbb{N}} O_n$. Clearly, $B_0 \subseteq B$, i.e. each $f \in B_0$ is nowhere differentiable in $[a,b]$. If we can show each O_n to be open and dense in $C[a,b]$, then, by Th. 2.6(a), B_0 (and, thus, B) must be dense as well.

Taking complements, we obtain $D_0 := (B_0)^c$ (and, thus, D) to be of the first category in $C[a,b]$. It remains to show each O_n is open and dense. We first show O_n to be open: Let $f \in O_n$. Then

$$\underset{x\in[a,b]}{\forall} \quad \underset{\delta_x\in\mathbb{R}^+}{\exists} \quad \sup\left\{\left|\frac{f(x+h)-f(x)}{h}\right| : 0<|h|\leq\frac{1}{n}\right\} > n+\delta_x.$$

Thus, there exists $h_x \in \mathbb{R}$ with $0 < |h_x| \leq \frac{1}{n}$ and $\left|\frac{f(x+h_x)-f(x)}{h_x}\right| > n+\delta_x$. Now the continuity of f (at x) implies

$$\underset{U_x\in\mathcal{U}(x)}{\exists} \quad \underset{t\in U_x}{\forall} \quad \left|\frac{f(t+h_x)-f(t)}{h_x}\right| > n+\delta_x.$$

Since $[a,b]$ is compact, there exist finitely many $x_1,\dots,x_N \in [a,b]$, $N \in \mathbb{N}$, such that

$$[a,b] \subseteq \bigcup_{i=1}^{N} U_{x_i}.$$

We set $\delta := \min\{\delta_{x_1},\dots,\delta_{x_N}\}$, $h := \min\{|h_{x_1}|,\dots,|h_{x_N}|\}$, $\epsilon := \frac{1}{2}h\delta > 0$, and show $g \in O_n$ for each $g \in C[a,b]$ with $\|g-f\|_\infty < \epsilon$: If $g \in C[a,b]$ with $\|g-f\|_\infty < \epsilon$ and $x \in [a,b]$, then there exists $i \in \{1,\dots,N\}$ with $x \in U_{x_i}$ and

$$|f(x+h_{x_i})-f(x)| \leq |f(x+h_{x_i})-g(x+h_{x_i})| + |g(x+h_{x_i})-g(x)| + |g(x)-f(x)|.$$

Thus,

$$\left|\frac{g(x+h_{x_i})-g(x)}{h_{x_i}}\right| \geq \left|\frac{f(x+h_{x_i})-f(x)}{h_{x_i}}\right| - 2\,\frac{\|f-g\|_\infty}{|h_{x_i}|} > n+\delta-2\,\frac{\epsilon}{h} = n,$$

showing $g \in O_n$ and O_n open. It remains to show O_n is also dense. From the Weierstrass approximation Th. C.1, we know the set of polynomials from $\mathbb{R}$ to $\mathbb{R}$ to be dense in $(C[a,b], \|\cdot\|_\infty)$. Thus, if $\emptyset \neq O \subseteq C[a,b]$, O open, then there exist a polynomial $p : \mathbb{R} \longrightarrow \mathbb{R}$ and $\epsilon \in \mathbb{R}^+$ such that

$$\underset{f\in C[a,b]}{\forall} \quad \Big(\|f - p{\restriction_{[a,b]}}\|_\infty \leq \epsilon \;\Rightarrow\; f \in O\Big).$$

For each $n \in \mathbb{N}$, define the (continuous and piecewise affine) triangle wave function

$$\phi_n : \mathbb{R} \longrightarrow [0,\epsilon], \quad \phi_n(x) := \begin{cases} nx - k2\epsilon & \text{for } x \in [\frac{k2\epsilon}{n}, \frac{k2\epsilon+\epsilon}{n}],\ k \in \mathbb{Z}, \\ -nx + (k+1)2\epsilon & \text{for } x \in [\frac{k2\epsilon+\epsilon}{n}, \frac{(k+1)2\epsilon}{n}],\ k \in \mathbb{Z} \end{cases}$$

(the slope of ϕ_n alternates between n and $-n$ on intervals of length $\frac{\epsilon}{n}$) and also

$$g_n := \phi_n\restriction_{[a,b]} \in C[a,b].$$

Then, clearly, $\|g_n\|_\infty \leq \epsilon$ (i.e. $f_n := p + g_n \in O$) and

$$\underset{x\in[a,b]}{\forall} \quad \lim_{h\to 0}\left|\frac{g_n(x+h)-g_n(x)}{h}\right| = n. \tag{2.7}$$

One also has, for each $x \in [a,b]$ and each $h \in \mathbb{R}$,

$$|g_n(x+h)-g_n(x)| \leq |f_n(x+h)-f_n(x)| + |p(x+h)-p(x)|,$$

implying, for $h \neq 0$,

$$\begin{aligned}\left|\frac{f_n(x+h)-f_n(x)}{h}\right| &\geq \left|\frac{g_n(x+h)-g_n(x)}{h}\right| - \left|\frac{p(x+h)-p(x)}{h}\right| \\ &\geq \left|\frac{g_n(x+h)-g_n(x)}{h}\right| - \|p'\restriction_{[a,b]}\|_\infty, \end{aligned} \tag{2.8}$$

where the mean value theorem was used for the last estimate. For each $m \in \mathbb{N}$ such that $m > n + \|p'\restriction_{[a,b]}\|_\infty$, combining (2.8) with (2.7) yields

$$\sup\left\{\left|\frac{f_m(x+h)-f_m(x)}{h}\right| : 0 < |h| \leq \frac{1}{n}\right\} \geq m - \|p'\restriction_{[a,b]}\|_\infty > n,$$

showing $f_m \in O_n$, i.e. $f_m \in O \cap O_n \neq \emptyset$. Thus, O_n is dense as desired.

—

In [Phi16a, Ex. 8.3(b)], we saw an example of a sequence of continuous functions on $[0,1]$ converging pointwise to a discontinuous function; in [Phi16a, Ex. 9.14(c)], we saw a differentiable function on $\mathbb{R}$ with discontinuous derivative. So one might ask, whether pointwise limits of continuous functions can be *everywhere* discontinuous and where derivatives can be *everywhere* discontinuous. We can use the Baire category Th. 2.6 to show that the set of points of discontinuity of such a limit as well as of such a derivative must be of the first category (see Th. 2.11 below). In preparation, we introduce the *oscillation* of a real-valued function:

Definition and Remark 2.9. Let $M \subseteq \mathbb{R}$, $f : M \longrightarrow \mathbb{R}$. For each nonempty $A \subseteq M$, define the *oscillation* of f in A by

$$\omega(A) := \sup\{f(x) : x \in A\} - \inf\{f(x) : x \in A\} \in [0,\infty]. \tag{2.9}$$

For each $\xi \in M$, the function

$$\alpha_\xi : \mathbb{R}^+ \longrightarrow [0,\infty], \quad \alpha_\xi(h) := \omega\big(M\cap]\xi-h,\xi+h[\big),$$

is decreasing with lower bound 0, such that

$$\omega(\xi) := \lim_{h\to 0} \alpha_\xi(h) = \inf\{\alpha_\xi(h) : h \in \mathbb{R}^+\} \in [0,\infty] \tag{2.10}$$

is well-defined. We call $\omega(\xi)$ the *oscillation* of f at ξ.

Proposition 2.10. *Let $M \subseteq \mathbb{R}$, $f : M \longrightarrow \mathbb{R}$, $\xi \in M$.*

(a) *f is continuous at ξ if, and only if, $\omega(\xi) = 0$.*

(b) *For each $\epsilon \in \mathbb{R}^+$, $O_\epsilon := \{x \in M : \omega(x) < \epsilon\}$ is open (in M); $A_\epsilon := O_\epsilon^{\mathrm{c}} = \{x \in M : \omega(x) \geq \epsilon\}$ is closed (in M).*

Proof. (a): Let f be continuous at ξ and $\epsilon > 0$. Then there exists $h > 0$ such that $|f(\xi) - f(x)| < \epsilon$ for each $x \in M\cap]\xi-h,\xi+h[$, showing $\alpha_\xi(h) \leq 2\epsilon$ and $\omega(\xi) = 0$. Conversely, assume f is not continuous at ξ. Then there exists $\epsilon_0 > 0$ such that, for each $h > 0$, there exists $x \in M\cap]\xi-h,\xi+h[$ with $|f(\xi)-f(x)| \geq \epsilon_0$, showing $\alpha_\xi(h) \geq \epsilon_0$ and $\omega(\xi) \geq \epsilon_0 > 0$.

(b): Let $\epsilon \in \mathbb{R}^+$ and $\xi \in O_\epsilon$. Then $\omega(\xi) < \epsilon$ and there exists $h \in \mathbb{R}^+$ such that

$$\omega(\xi) \leq \alpha_\xi(h) = \omega\big(M\cap]\xi-h,\xi+h[\big) < \epsilon.$$

Thus, $\omega(x) < \epsilon$ for each $x \in M\cap]\xi-h,\xi+h[$, showing O_ϵ to be open; A_ϵ to be closed. ∎

Theorem 2.11. *Let $I \subseteq \mathbb{R}$ be a nontrivial interval (i.e. $I^\circ \neq \emptyset$), let $f : I \longrightarrow \mathbb{R}$, and let $D \subseteq I$ of points, where f is not continuous.*

(a) *Assume $(f_n)_{n\in\mathbb{N}}$ to be a sequence of continuous functions $f_n : I \longrightarrow \mathbb{R}$ such that $f_n \to f$ pointwise on I.*

(b) *Assume $f = g' : I \longrightarrow \mathbb{R}$ to be the derivative of some differentiable $g : I \longrightarrow \mathbb{R}$.*

In each case, (a) *or* (b)*, D is of the first category in I (then, by Th.* 2.6(c)*, $I \setminus D$ must be of the second category in I and also dense).*

Proof. Exercise. Hints: To show (a) first show, for each $\epsilon \in \mathbb{R}^+$, $A_\epsilon := \{x \in I : \omega(x) \geq \epsilon\}$, where $\omega(x)$ denotes the oscillation of f at x, to be nowhere dense. Then use Prop. 2.10. Show that (b) follows from (a). ∎

2.2 Uniform Boundedness Principle, Banach-Steinhaus Theorem

In the present section, we combine the Baire category concept of the previous section with concepts of continuity, uniformity, and boundedness.

Definition 2.12. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological vector spaces over $\mathbb{K}$ and let $\mathcal{F}$ be a set of functions from X into Y. Then the set $\mathcal{F}$ (or the functions in $\mathcal{F}$) are said to be *uniformly equicontinuous* if, and only if,

$$\forall_{U \in \mathcal{U}(0) \subseteq \mathcal{P}(Y)} \quad \exists_{V \in \mathcal{U}(0) \subseteq \mathcal{P}(X)} \quad \forall_{x,y \in X} \quad \Big(y - x \in V \;\Rightarrow\; \forall_{f \in \mathcal{F}} f(y) - f(x) \in U \Big) \tag{2.11}$$

(then, for $\mathcal{F} = \{f\}$, $\mathcal{F}$ is uniformly equicontinuous if, and only if, f is uniformly continuous (cf. Th. 1.13(iii)).

Proposition 2.13. *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological vector spaces over $\mathbb{K}$ and let $\mathcal{F}$ be a set of* $\mathbb{K}$-linear *functions from X into Y.*

(a) *$\mathcal{F}$ is uniformly equicontinuous if, and only if,*

$$\forall_{U \in \mathcal{U}(0) \subseteq \mathcal{P}(Y)} \quad \exists_{V \in \mathcal{U}(0) \subseteq \mathcal{P}(X)} \quad \forall_{A \in \mathcal{F}} \quad A(V) \subseteq U. \tag{2.12}$$

(b) *If $\mathcal{F}$ is uniformly equicontinuous, then $\mathcal{F}$ is uniformly bounded in the following sense: For each bounded $E \subseteq X$, there exists a bounded set $F \subseteq Y$ such that*

$$\forall_{A \in \mathcal{F}} \quad A(E) \subseteq F.$$

Proof. (a): If (2.11) holds with f replaced by A, then setting $x := 0$ proves (2.12). Conversely, if (2.12) holds, $A \in \mathcal{F}$, and $y - x \in V$, then $A(y) - A(x) = A(y - x) \in U$, proving (2.11).

(b): Let $\mathcal{F}$ be uniformly equicontinuous, let $E \subseteq X$ be bounded, and set $F := \bigcup_{A \in \mathcal{F}} A(E)$. If $U \subseteq Y$, $U \in \mathcal{U}(0)$, then, as $\mathcal{F}$ is uniformly equicontinuous, there exists $V \subseteq X$, $V \in \mathcal{U}(0)$, such that $A(V) \subseteq U$ for each $A \in \mathcal{F}$. Moreover, since E is bounded, there exists $s \in \mathbb{R}^+$ such that $E \subseteq sV$. Then

$$\forall_{A \in \mathcal{F}} \quad A(E) \subseteq A(sV) = sA(V) \subseteq sU,$$

showing $F \subseteq sU$, i.e. F is bounded. ■

Theorem 2.14 (Uniform Boundedness Principle)**.** *Let $(X, \mathcal{T}_X)$, $(Y, \mathcal{T}_Y)$ be nonemtpy topological spaces, where we assume to have a notion of boundedness on Y. Consider a set of continuous functions $\mathcal{F} \subseteq C(X, Y)$ and define*

$$\underset{x \in X}{\forall} \quad M_x := \{f(x) : f \in \mathcal{F}\} \tag{2.13}$$

as well as

$$B := \{x \in X : M_x \text{ is bounded}\}. \tag{2.14}$$

Assume B to be of the second category in X.

(a) *If $(Y, \mathcal{T}_Y)$ is pseudometrizable (e.g., if $(Y, \mathcal{T}_Y)$ is a seminormed space), then there exists a nonempty open set $O \subseteq X$ such that $\mathcal{F}$ is uniformly bounded on O in the sense that there exists a (pseudometric-)bounded set $F \subseteq Y$ such that*

$$\underset{x \in O}{\forall} \quad M_x \subseteq F.$$

(b) *If $(X, \mathcal{T}_X)$, $(Y, \mathcal{T}_Y)$ are topological vector spaces over $\mathbb{K}$ and the elements of $\mathcal{F}$ are both continuous and linear, then $B = X$ and $\mathcal{F}$ is uniformly equicontinuous.*

Proof. If $\mathcal{F} = \emptyset$, then $M_x = \emptyset$ for each $x \in X$, $\mathcal{F}$ is (in case (b)), trivially, uniformly equicontinuous, and there is nothing to prove. Thus, let $\mathcal{F} \neq \emptyset$.

(a): Assume $\mathcal{T}_Y$ to be induced by the pseudometric d on Y, and fix some $y_0 \in Y$. For each $f \in \mathcal{F}$, $k \in \mathbb{N}$, as both f and the $y \mapsto d(y, y_0)$ are continuous, the set

$$A_{k,f} := \big\{x \in X : d(f(x), y_0) \leq k\big\}$$

is an continuous inverse image of the closed set $[0, k]$ and, hence, closed. Since arbitrary intersections of closed sets are closed, so is

$$A_k := \bigcap_{f \in \mathcal{F}} A_{k,f} = \big\{x \in X : d(f(x), y_0) \leq k \text{ for each } f \in \mathcal{F}\big\}.$$

If $x \in B$, then M_x is bounded, i.e. there exists $k \in \mathbb{N}$ with $d(f(x), y_0) \leq k$ for each $f \in \mathcal{F}$, implying

$$B = \bigcup_{k=1}^{\infty} A_k.$$

Since B is of the second category, there exists $k_0 \in \mathbb{N}$ such that A_{k_0} has nonempty interior O. In consequence,

$$\underset{x \in O}{\forall} \quad M_x \subseteq F := B_{d,k_0}(y_0),$$

proving (a).

(b): First, we show $\mathcal{F}$ to be uniformly equicontinuous: Let $W \subseteq Y$, $W \in \mathcal{U}(0)$. Choose some closed and balanced $U \in \mathcal{U}(0)$ such that $U + U \subseteq W$. Then, since each $A \in \mathcal{F}$ is continuous,

$$E := \bigcap_{A \in \mathcal{F}} A^{-1}(U)$$

is a closed subset of X. For each $x \in B$, since M_x is bounded, there exists $k \in \mathbb{N}$ such that $M_x \subseteq kU$ and $x \in kE$, implying

$$B \subseteq \bigcup_{k \in \mathbb{N}} (kE).$$

Since B is of the second category in X, there exists k such that kE is of the second category in X. Then, since $x \mapsto kx$ is a homeomorphism, E itself must be of the second category in X. In particular, since E is also closed, $E^\circ \neq \emptyset$. Let $x \in E^\circ$. Then

$$V := x - E \in \mathcal{U}(0),$$

implying

$$\forall_{A \in \mathcal{F}} \quad A(V) = A(x) - A(E) \subseteq U - U \subseteq W,$$

showing $\mathcal{F}$ to be uniformly equicontinuous. Thus, $\mathcal{F}$ is uniformly bounded by Prop. 2.13(b) and, since each $\{x\}$, $x \in X$, is bounded in X, each M_x must be bounded in Y, proving $B = X$. ■

Using the Baire category Th. 2.6, we now obtain the following corollaries:

Corollary 2.15. *As in Th. 2.14, let $(X, \mathcal{T}_X)$, $(Y, \mathcal{T}_Y)$ be nonemtpy topological spaces, $\mathcal{F} \subseteq C(X, Y)$, $M_x := \{f(x) : f \in \mathcal{F}\}$ for each $x \in X$, $B := \{x \in X : M_x$ is bounded$\}$.*

(a) *If X is a nonempty complete pseudometric space and Y is a seminormed vector space with*

$$\sigma_x := \sup\{\|y\| : y \in M_x\} < \infty \quad \textit{for each } x \in X, \tag{2.15}$$

then there exists $x_0 \in X$ and $\epsilon_0 > 0$ such that

$$\sup\left\{\sigma_x : x \in \overline{B}_{\epsilon_0}(x_0)\right\} < \infty. \tag{2.16}$$

In other words, if a collection of continuous functions from X into Y is bounded pointwise in X, then it is uniformly *bounded on an entire ball.*

(b) *If $(X, \mathcal{T}_X)$, $(Y, \mathcal{T}_Y)$ are topological vector spaces over $\mathbb{K}$, where X is an F-space, $B = X$ (i.e. all M_x are bounded), and the elements of $\mathcal{F}$ are both continuous and linear, then $\mathcal{F}$ is uniformly equicontinuous.*

(c) Banach-Steinhaus Theorem: *If X, Y are normed spaces and X is a Banach space (i.e. complete), $B = X$ (i.e. $\mathcal{F}$ is pointwise bounded), then*

$$\underset{M\in\mathbb{R}_0^+}{\exists} \quad \underset{A\in\mathcal{F}}{\forall} \quad \underset{\substack{x\in X,\\ \|x\|\leq 1}}{\forall} \quad \|Ax\| \leq M, \tag{2.17}$$

i.e., for each $A \in \mathcal{F}$, $\|A\| := \sup\{\|Ax\| : x \in X, \|x\| \leq 1\} \leq M$. We will see later that $\|A\|$ (the so-called operator norm *of A) does actually constitute a norm on $\mathcal{L}(X, Y)$, the space of continuous linear functions from X into Y. Thus, we can restate the Banach-Steinhaus theorem by saying that, if $\mathcal{F}$ is pointwise bounded, than it is a bounded subset of $\mathcal{L}(X, Y)$.*

Proof. (a): According to (2.15), we have that M_x is bounded by σ_x for each $x \in X$, i.e. $B = X$. Since X is a complete pseudometric space, by the Baire category Th. 2.6(c), X is of the second category in itself. Thus, Th. 2.14(a) implies $\mathcal{F}$ to be uniformly bounded on a nonempty open set O, such that (2.16) holds with O instead of $\overline{B}_{\epsilon_0}(x_0)$. But then (2.16) also follows, as O contains some closed ball $\overline{B}_{\epsilon_0}(x_0)$.

(b): As in (a), we have $B = X$ and the Baire category Th. 2.6(c) yields X to be of the second category in itself (since an F-space is a complete metric space). Now Th. 2.14(b) implies $\mathcal{F}$ to be uniformly equicontinuous.

(c) follows by combining (b) with Th. 2.13(b). ■

Under suitable hypotheses, the uniform boundedness principle allows to establish the continuity of pointwise limits of continuous linear maps. First, we provide a linearity result:

Proposition 2.16. *Let X be a vector space over $\mathbb{K}$ and $(Y, \mathcal{T}_Y)$ a topological vector space over $\mathbb{K}$. Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of linear maps $A_n : X \longrightarrow Y$ and define*

$$C := \{x \in X : (A_n(x))_{n\in\mathbb{N}} \text{ Cauchy in } Y\},$$
$$L := \{x \in X : (A_n(x))_{n\in\mathbb{N}} \text{ converges in } Y\}.$$

Then C and L are vector subspaces of X. If $L = X$ and $(Y, \mathcal{T}_Y)$ is T_1, then the pointwise limit

$$A : X \longrightarrow Y, \quad A(x) := \lim_{n\to\infty} A_n(x),$$

is linear (if $(Y, \mathcal{T}_Y)$ is T_1, then it is T_2 by Prop. 1.5(d)*, i.e. limits in Y are unique and A is well-defined).*

Proof. Exercise. ■

Theorem 2.17. *Let $(X, \mathcal{T}_X)$, $(Y, \mathcal{T}_Y)$ be topological vector spaces over $\mathbb{K}$, and let $(A_n)_{n\in\mathbb{N}}$ be a sequence of linear continuous maps $A_n : X \longrightarrow Y$, $\mathcal{F} := \{A_n : n \in \mathbb{N}\}$. Define C and L as in Prop.* 2.16 *above.*

(a) *If C is of the second category in X, then $\mathcal{F}$ is uniformly equicontinuous and $X = C$.*

(b) *If $\mathcal{F}$ is uniformly equicontinuous, $X = L$, and $(Y, \mathcal{T}_Y)$ is T_1 (i.e. the A_n converge pointwise to a unique limit), then the pointwise limit*

$$A : X \longrightarrow Y, \quad A(x) := \lim_{n\to\infty} A_n(x),$$

is linear and continuous.

(c) *If Y is complete (e.g. and F-space) and L is of the second category in X, then $X = L$, and $\mathcal{F}$ is uniformly equicontinuous. If Y is also T_1, then the pointwise limit A as above is linear and continuous.*

(d) *If X is an F-space (e.g. a Banach space) and $L = X$, then $\mathcal{F}$ is uniformly equicontinuous. If, in addition, Y is also T_1, then the pointwise limit A as above is linear and continuous.*

Proof. (a): Since Cauchy sequences are bounded by Prop. 1.28(c), $\mathcal{F}$ is pointwise bounded on C. Thus, Th. 2.14(b) applies, showing $\mathcal{F}$ to be uniformly equicontinuous. Moreover, from Prop. 2.16, we know C to be a vector subspace of X. Since C is of the second category, it can not be nowhere dense. Thus, by Ex. 2.3(b), it must be dense. Let $x \in X$, $W \subseteq Y$, $W \in \mathcal{U}(0)$, and choose $U \in \mathcal{U}(0)$ such that $U + U + U \subseteq W$. Since $\mathcal{F}$ is uniformly equicontinuous, there exists $V \subseteq X$, $V \in \mathcal{U}(0)$ balanced, such that $A_n(V) \subseteq U$ for each $n \in \mathbb{N}$. As C is dense, there exists $z \in C \cap (x + V)$. Choose $N \in \mathbb{N}$ such that

$$\underset{m,n>N}{\forall} \quad A_n(z) - A_m(z) \in U.$$

Then,

$$\underset{m,n>N}{\forall} \quad A_n(x) - A_m(x) = A_n(x - z) + A_n(z) - A_m(z) + A_m(z - x) \in U + U + U \subseteq W,$$

showing $x \in C$, i.e. $C = X$.

(b): A is linear by Prop. 2.16. To see the continuity of A, let $W \subseteq Y$, $W \in \mathcal{U}(0)$, and choose $U \in \mathcal{U}(0)$ such that $\overline{U} \subseteq W$ (which is possible by Prop. 1.5(c)). Since $\mathcal{F}$ is uniformly equicontinuous, there exists $V \subseteq X$, $V \in \mathcal{U}(0)$, such that $A_n(V) \subseteq U$ for each $n \in \mathbb{N}$. Thus, $A(V) \subseteq \overline{U} \subseteq W$, proving A to be continuous.

(c): Since $L \subseteq C$ by Prop. 1.28(b), C is of the second category in X. Thus, by (a), $\mathcal{F}$ is uniformly equicontinuous and $X = C$. Since Y is complete, we also have $C \subseteq L$ and, thus, $X = L$. If Y is also T_1, then (b) applies and we obtain A to be linear and continuous.

(d): By the Baire category Th. 2.6(c), we know the F-space $L = X$ to be of the second category in itself. Now $C = X$ (since covergent sequences are Cauchy), i.e. (a) implies $\mathcal{F}$ to be uniformly equicontinuous. Thus, if, in addition, Y is T_1, then A is linear and continuous by (b). ■

Example 2.18. Consider

$$X := \{f \in C^1(\mathbb{R}, \mathbb{K}) : f \text{ and } f' \text{ are bounded}\}.$$

Clearly, $(X, \|\cdot\|_\infty)$ is a normed vector space over $\mathbb{K}$ (e.g. a subspace of $L^\infty(\mathbb{R}, \mathcal{L}^1, \lambda^1)$). Define the continuous linear functionals

$$\underset{n\in\mathbb{N}}{\forall} \quad A_n : X \longrightarrow \mathbb{K}, \quad A_n(f) := \frac{f(n^{-1}) - f(0)}{n^{-1}}.$$

Each A_n is a difference quotient map (at 0) and a linear combination of two (linear) evaluation functionals $A_n = n(\pi_{n^{-1}} + \pi_0)$, where the evaluation functionals are precisely the projections

$$\underset{x\in\mathbb{R}}{\forall} \quad \pi_x : \mathbb{K}^{\mathbb{R}} \longrightarrow \mathbb{K}, \quad \pi_x(f) := f(x).$$

Since each $f \in X$ is differentiable, we have the pointwise convergence $A_n \to A$, where A is the linear functional

$$A : X \longrightarrow \mathbb{K}, \quad A(f) := f'(0).$$

However, A is *not* continuous: For each $k \in \mathbb{N}$, let $f_k : \mathbb{R} \longrightarrow \mathbb{K}$, $f_k(x) := (1/k)\sin(kx)$, $f_k'(x) = k(1/k)\cos(kx) = \cos(kx)$. Clearly, $f_k \in X$ for each $k \in \mathbb{N}$ with $\lim_{k\to\infty} \|f_k\|_\infty = 0$, but

$$\lim_{k\to\infty} A(f_k) = \cos(0) = 1 \neq 0 = A(0).$$

Now Th. 2.17(d) implies that $(X, \|\cdot\|_\infty)$ is not a Banach space (i.e. not complete). However, we can make X into a Banach space by modifying the norm: Let

$$\|\cdot\| : X \longrightarrow \mathbb{R}_0^+, \quad \|f\| := \|f\|_\infty + \|f'\|_\infty. \tag{2.18}$$

Suppose $(f_k)_{k\in\mathbb{N}}$ is Cauchy in $(X, \|\cdot\|)$. Then we know the continuous maps f_k to converge uniformly to some continuous $f : \mathbb{R} \longrightarrow \mathbb{K}$. Moreover, the continuous maps f_k' also converge uniformly to some continuous $g : \mathbb{R} \longrightarrow \mathbb{K}$. Now Th. B.1 implies f to be continuously differentiable with $f' = g$, i.e. $f_k \to f$ in $(X, \|\cdot\|)$ and $(X, \|\cdot\|)$ is a Banach space. Now Th. 2.17(d) implies the map A from above (evaluating the derivative at 0) to be continuous (on $(X, \|\cdot\|)$ – of course, here it is also easy to see that without Th. 2.17(d); also note that, for the f_k from above, $f_k \not\to 0$ in $(X, \|\cdot\|)$).

2.3 Open Mapping Theorem

Definition 2.19. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces, $\xi \in X$. A map $f : X \longrightarrow Y$ is called *open* at ξ if, and only if,

$$\underset{V \in \mathcal{U}(\xi)}{\forall} \quad \underset{U \in \mathcal{U}(f(\xi))}{\exists} \quad U \subseteq f(V). \tag{2.19}$$

Moreover, f is called *open* if, and only if, f maps open set to open sets (i.e. $f(O) \in \mathcal{T}_Y$ for each $O \in \mathcal{T}_X$).

Caveat 2.20. While we know that a map is continuous if, and only if, preimages of open sets are open and also if, and only if, preimages of closed sets are closed, there exist open maps that do not map closed sets to closed sets: Projections from a product to the factors are always open maps according to [Phi16b, Ex. 2.12(b)(ii)]. In particular, $\pi_1 : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $\pi_1(s,t) := s$, is open. However, $A := \{(s,t) \in \mathbb{R}^2 : s \geq 0, st \geq 1\}$ is closed, but $\pi_1(A) =]0, \infty[$ is not closed.

Proposition 2.21. *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces, $f : X \longrightarrow Y$.*

(a) *f is open if, and only if, f is open at every $\xi \in X$.*

(b) *If $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ are topological vector spaces over $\mathbb{K}$ and f is linear, then f is open if, and only if, f is open at 0. If $f : X \longrightarrow \mathbb{K}$ is linear, then $f \equiv 0$ or f is open.*

(c) *If f is bijective, then f^{-1} is continuous if, and only if, f is open. In particular, if f is bijective and continuous, then f is a homeomorphism if, and only if, f is open.*

Proof. Exercise. ■

Theorem 2.22 (Open Mapping Theorem)**.** *Let $(X, \mathcal{T}_X)$, $(Y, \mathcal{T}_Y)$ be T_1 topological vector spaces over $\mathbb{K}$ and assume $(X, \mathcal{T}_X)$ to be an F-space. Let $A : X \longrightarrow Y$ be continuous and linear and assume $A(X)$ to be of the second category in Y. Then the following assertions hold true:*

(a) *$A(X) = Y$.*

(b) *A is an open mapping.*

(c) *$(Y, \mathcal{T}_Y)$ is an F-space.*

Proof. (a) follows from (b), since proper vector subspaces can never be open (see Ex. 2.3(b)).

(b): We have to show A is open at 0. To this end, let $V \subseteq X$, $V \in \mathcal{U}(0)$. We have to find $U \subseteq Y$ such that $U \in \mathcal{U}(0)$ and $U \subseteq A(V)$. Since $(X, \mathcal{T}_X)$ is an F-space, $\mathcal{T}_X$ is induced by some complete, translation-invariant metric d on X. Choose $r \in \mathbb{R}^+$ such that $V_0 := B_{d,r}(0) \subseteq V$ and define

$$\underset{n\in\mathbb{N}_0}{\forall} \quad V_n := B_{d,2^{-n}r}(0) \in \mathcal{U}(0).$$

Let $m, n \in \mathbb{N}_0$, $m \leq n$. Then $V_n \subseteq V_m \subseteq V_0 \subseteq V$ and also, for $m < n$, $V_n - V_n \subseteq V_m$. Since $y \mapsto -y$ is a homeomorphism, we can apply Prop. 1.10(b) to conclude

$$\underset{\substack{m,n\in\mathbb{N}_0,\\ m<n}}{\forall} \quad \overline{A(V_n)} - \overline{A(V_n)} = \overline{A(V_n)} + \overline{-A(V_n)} \subseteq \overline{A(V_n) - A(V_n)} \subseteq \overline{A(V_m)}. \tag{2.20}$$

Next, by Prop. 1.12(f), we have

$$\underset{n\in\mathbb{N}_0}{\forall} \quad X = \bigcup_{k\in\mathbb{N}} (kV_n) \quad \Rightarrow \quad A(X) = \bigcup_{k\in\mathbb{N}} \big(kA(V_n)\big).$$

Since, by hypothesis, $A(X)$ is of the second category in Y, at least one $kA(V_n)$ must be of the second category in Y. Thus, since $y \mapsto ky$ is a homeomorphism, each $A(V_n)$ must be of the second category in Y, implying $\overline{A(V_n)}$ to have nonempty interior. In other words, for each $n \in \mathbb{N}_0$, there exists a nonempty open set $W_n \subseteq \overline{A(V_n)}$. Then $U_n := W_n - W_n$ is still nonempty and open, $U_n \in \mathcal{U}(0)$ and, for $n \geq 1$, $m < n$, $U_n \subseteq \overline{A(V_m)}$ by (2.20). We will still prove

$$\overline{A(V_1)} \subseteq A(V), \tag{2.21}$$

where we note that (2.21) shows (b), as we can let $U := U_2$. Given $y \in \overline{A(V_1)}$, we inductively construct a sequence $(y_n)_{n\in\mathbb{N}}$ in Y such that each $y_n \in \overline{A(V_n)}$ as follows: Start with setting $y_1 := y$. Then, assuming $y_1, \dots, y_n$ to be constructed, $y_n \in \overline{A(V_n)}$ means that every neighborhood of y_n has nonempty intersection with $A(V_n)$. Since $\overline{A(V_{n+1})} \in \mathcal{U}(0)$, we have $y_n - \overline{A(V_{n+1})} \in \mathcal{U}(y_n)$ and

$$\Big(y_n - \overline{A(V_{n+1})}\Big) \cap A(V_n) \neq \emptyset.$$

In other words,

$$\underset{x_n\in V_n}{\exists} \quad A(x_n) \in y_n - \overline{A(V_{n+1})}.$$

Thus, if we let $y_{n+1} := y_n - A(x_n)$, then we have $y_{n+1} \in \overline{A(V_{n+1})}$ as desired. Then the continuity of A implies

$$\lim_{n\to\infty} y_n = 0 : \tag{2.22}$$

Indeed, let $W \subseteq Y$, $W \in \mathcal{U}(0)$. The continuity of A at 0 yields $n_0 \in \mathbb{N}$ with $A(V_{n_0}) \subseteq W$ (since the V_n, $n \in \mathbb{N}$, form a local base for $\mathcal{T}_X$ at 0). Since the V_n are decreasing, $A(V_{n_0}) \subseteq W$ proves (2.22). We now define, for each $n \in \mathbb{N}$, $z_n := \sum_{i=1}^n x_i \in X$. Then, using translation-invariance of d,

$$\underset{m<n\in\mathbb{N}}{\forall} \quad d(z_n, z_m) \leq \sum_{i=m+1}^{n} d(z_i, z_{i-1}) = \sum_{i=m+1}^{n} d(z_i - z_{i-1}, 0) = \sum_{i=m+1}^{n} d(x_i, 0) < r \sum_{i=m+1}^{n} 2^{-i},$$

showing $(z_n)_{n\in\mathbb{N}}$ to be a Cauchy sequence in X (since the partial sums of the geometric series form a Cauchy sequence in $\mathbb{R}$). As we assume X to be complete, there exists a limit $z := \lim_{n\to\infty} z_n \in X$. Moreover,

$$d(z, 0) = \lim_{n\to\infty} d(z_n, 0) \leq \sum_{i=1}^{\infty} d(x_i, 0) < r \sum_{i=1}^{\infty} 2^{-i} = r,$$

showing $z \in B_{d,r}(0) = V_0 \subseteq V$. Using the continuity of A once again, we obtain

$$A(z) = \lim_{n\to\infty} A(z_n) = \lim_{n\to\infty} \sum_{i=1}^{n} A(x_i) = \lim_{n\to\infty} \sum_{i=1}^{n} (y_i - y_{i+1}) = \lim_{n\to\infty} (y_1 - y_{n+1}) \overset{(2.22)}{=} y_1 = y$$

(here we also used that limits in Y are unique), showing $y \in A(V)$, (2.21), and (b).

(c): One uses that X being an F-space implies the factor space $X/\ker A$ to be an F-space as well, and shows $X/\ker A$ to be homeomorphic to Y (we refer to [Rud73, Th. 2.11(iii)] for the details). ■

Example 2.23. Let $p, q \in [1, \infty]$ with $p < q \leq \infty$. From [Phi17, Th. 2.42], we know $L^q([0,1], \mathcal{L}^1, \lambda^1) \subseteq L^p([0,1], \mathcal{L}^1, \lambda^1)$. Since

$$f :]0,1] \longrightarrow \mathbb{K}, \quad f(t) := t^{-\frac{1}{q}}, \tag{2.23}$$

is in $L^p([0,1], \mathcal{L}^1, \lambda^1)$ (by [Phi16a, Ex. 10.35(a)], since $\frac{p}{q} < 1$), but not in $L^q([0,1], \mathcal{L}^1, \lambda^1)$ (by [Phi16a, Ex. 10.35(b)]), we also know $L^q([0,1], \mathcal{L}^1, \lambda^1) \subsetneq L^p([0,1], \mathcal{L}^1, \lambda^1)$. As an application of Th. 2.22, we can now show that $X := L^q([0,1], \mathcal{L}^1, \lambda^1)$ is of the first category in $Y := L^p([0,1], \mathcal{L}^1, \lambda^1)$: Let $A := \mathrm{Id} : X \longrightarrow Y$. Then A is linear. According to [Phi17, Th. 2.42], we have

$$\underset{f\in L^q}{\forall} \quad \|f\|_p \leq \|f\|_q,$$

showing A to be bounded. Then, by Th. 1.32, A is continuous. Since X is an F-space (even a Banach space) and $A(X) \neq Y$, Th. 2.22 yields that X must be of the first category in Y.

Proposition 2.24. *Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be normed vector spaces over $\mathbb{K}$ and let $f : X \longrightarrow Y$. If*

$$\underset{\alpha\in\mathbb{R}^+}{\exists} \quad \underset{x\in X}{\forall} \quad \|f(x)\|_Y \leq \alpha\|x\|_X, \tag{2.24}$$

then f maps bounded sets into bounded sets. If f is linear, then the converse also holds (i.e. (2.24) is equivalent to the linear map being bounded).

Proof. Exercise. ■

Corollary 2.25. *Let $(X, \mathcal{S})$, $(Y, \mathcal{T})$ be F-spaces over $\mathbb{K}$ (e.g. Banach spaces) and assume $A : X \longrightarrow Y$ to be continuous and linear. Then the following assertions hold true:*

(a) *If A is surjective, then A is open.*

(b) *If A is bijective, then A^{-1} is continuous (i.e. A is then a homeomorphism).*

(c) *If $X = Y$ and $\mathcal{S} \subseteq \mathcal{T}$, then $\mathcal{S} = \mathcal{T}$.*

(d) *If $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are Banach spaces, where $\|\cdot\|_X$ induces $\mathcal{S}$, $\|\cdot\|_Y$ induces $\mathcal{T}$, and A is bijective, then*

$$\underset{\alpha,\beta\in\mathbb{R}^+}{\exists} \quad \underset{x\in X}{\forall} \quad \alpha\|x\|_X \leq \|Ax\|_Y \leq \beta\|x\|_X. \tag{2.25}$$

(e) *If $X = Y$, $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ are Banach spaces such that $\|\cdot\|_1$ induces $\mathcal{S}$, $\|\cdot\|_2$ induces $\mathcal{T}$, and*

$$\underset{\alpha\in\mathbb{R}^+}{\exists} \quad \underset{x\in X}{\forall} \quad \|x\|_1 \leq \alpha\|x\|_2, \tag{2.26}$$

then $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent (i.e. $\mathcal{S} = \mathcal{T}$).

Proof. (a): As A is sujective, we have $A(X) = Y$. Since Y is an F-space, the Baire category Th. 2.6(c) implies $A(X) = Y$ to be of the second category in itself. Thus, A is open by Th. 2.22(b).

(b) follows from (a) combined with Prop. 2.21(c).

(c): If $\mathcal{S} \subseteq \mathcal{T}$, then $\mathrm{Id} : (X, \mathcal{T}) \longrightarrow (X, \mathcal{S})$ is continuous. Then $\mathrm{Id} : (X, \mathcal{S}) \longrightarrow (X, \mathcal{T})$ is continuous by (b), implying $\mathcal{T} \subseteq \mathcal{S}$.

(d): A is continuous by hypothesis, A^{-1} is continuous by (b). Thus, A and A^{-1} are bounded. According to Prop. 2.24, the boundedness of A implies the second inequality of (2.25) and the boundedness of A^{-1} implies the second inequality of (2.25): Indeed,

$$\underset{\gamma>0}{\forall} \quad \underset{x\in X}{\exists} \quad \|x\|_X = \|A^{-1}(A(x))\|_X \leq \gamma\,\|A(x)\|_Y$$

(set $\alpha := \gamma^{-1}$).

(e): According to (2.26) and Prop. 2.24, $\mathrm{Id} : (X, \|\cdot\|_2) \longrightarrow (X, \|\cdot\|_1)$ is bounded and, thus, by Th. 1.32, continuous, implying $\mathcal{S} \subseteq \mathcal{T}$. Now $\mathcal{S} = \mathcal{T}$ holds according to (c). ■

Example 2.26. Consider the Banach space $(X, \|\cdot\|_1)$, where

$$X := \{f \in C^1(\mathbb{R}, \mathbb{K}) : f \text{ and } f' \text{ are bounded}\},$$
$$\|f\|_1 := \|f\|_\infty + \|f'\|_\infty$$

(cf. Ex. 2.18 and, of course, $\|\cdot\|_1$ is not the L^1-norm). Let

$$\phi : \mathbb{R} \longrightarrow \mathbb{R}, \quad \phi(x) := e^{-x^2}.$$

Then we know $\phi \in L^1(\mathbb{R}, \mathcal{L}^1, \lambda^1)$ and we obtains two other norms on X by setting

$$\|f\| := \int_{\mathbb{R}} \phi\,|f|\,\mathrm{d}\lambda^1\,,$$
$$\|f\|_2 := \|f\|_1 + \|f\|,$$

immediately implying

$$\forall_{f \in X} \quad \|f\|_1 \leq \|f\|_2. \tag{2.27}$$

Let $(f_k)_{k\in\mathbb{N}}$ be a sequence in X that is $\|\cdot\|_2$-Cauchy. Then (2.27) implies $(f_k)_{k\in\mathbb{N}}$ to be $\|\cdot\|_1$-Cauchy as well. Thus, there exists $f \in X$ with $\lim_{k\to\infty} \|f_k - f\|_1 = 0$. By the dominated convergence theorem [Phi17, Th. 2.20], we have $\lim_{k\to\infty} \|f_k - f\| = 0$ as well, implying $\lim_{k\to\infty} \|f_k - f\|_2 = 0$. In consequence, $(X, \|\cdot\|_2)$ is also a Banach space. Since (2.27) means that (2.26) is satisfied with $\alpha = 1$, Cor. 2.25(e) implies $\|\cdot\|_1$ and $\|\cdot\|_2$ to be equivalent.

2.4 Closed Graph Theorem

2.4.1 The Theorem

We start by recalling that, for each function $f : X \longrightarrow Y$, the *graph* of f is $\mathrm{graph}(f) = \{(x, y) \in X \times Y : y = f(x)\}$. In topological spaces, the graph of f is closely related to the Hausdorff separation property:

Lemma 2.27. *Let $(X, \mathcal{T})$ be a topological space. Then $(X, \mathcal{T})$ is T_2 if, and only if,* graph(Id) *is closed with respect to the product topology $\mathcal{P}$ on $X \times X$ (where* Id *denotes the identity on X).*

Proof. The set

$$\Delta := \operatorname{graph}(\mathrm{Id}) = \{(x,x) \in X \times X : x \in X\} \tag{2.28}$$

is also called the *diagonal* in $X \times X$. We have

$$\begin{aligned}
\Delta \text{ closed} \quad &\Leftrightarrow \quad \Delta^{\mathrm{c}} \text{ open} \\
&\Leftrightarrow \quad \underset{x,y\in X,\, x\neq y}{\forall} \quad \underset{O\in\mathcal{P}}{\exists} \quad (x,y) \in O \subseteq \Delta^{\mathrm{c}} \\
&\Leftrightarrow \quad \underset{x,y\in X,\, x\neq y}{\forall} \quad \underset{O_x,O_y\in\mathcal{T}}{\exists} \quad (x,y) \in O_x \times O_y \subseteq O \subseteq \Delta^{\mathrm{c}} \\
&\Leftrightarrow \quad \underset{x,y\in X,\, x\neq y}{\forall} \quad \underset{O_x,O_y\in\mathcal{T}}{\exists} \quad \Big(x \in O_x \ \wedge \ y \in O_y \ \wedge \ O_x \cap O_y = \emptyset \Big) \\
&\Leftrightarrow \quad (X, \mathcal{T}) \ T_2,
\end{aligned}$$

proving the lemma. ■

We now proceed to the situation of a continuous map $f : X \longrightarrow Y$. We know from Lem. 2.27 that, if $X = Y$, $f = \mathrm{Id}$, and Y is a T_2-space, then $\operatorname{graph}(f)$ is closed. We can now show that, if Y is a T_2-space, then $\operatorname{graph}(f)$ is closed for *every* continuous $f : X \longrightarrow Y$:

Proposition 2.28. *Let $(X, \mathcal{S})$, $(Y, \mathcal{T})$ be topological spaces and $f : X \longrightarrow Y$. If f is continuous and $(Y, \mathcal{T})$ is T_2, then* $\operatorname{graph}(f)$ *is closed with respect to the product topology $\mathcal{P}$ on $X \times Y$.*

Proof. Let $G := \operatorname{graph}(f)$. We have to show G^{c} to be open. To this end, let $(x,y) \in G^{\mathrm{c}}$. Then $y \neq z := f(x)$. Since Y is T_2, there exist $O_y, O_z \in \mathcal{T}$ with $y \in O_y$, $z \in O_z$, and $O_y \cap O_z = \emptyset$. Since f is continuous, there exists an open $U \in \mathcal{U}(x)$ such that $f(U) \subseteq O_z$. Then $U \times O_y \in \mathcal{P}$ with $(x,y) \in U \times O_y \subseteq G^{\mathrm{c}}$, showing G^{c} to be open. ■

While simple examples such as

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \quad f(x) := \begin{cases} x^{-1} & \text{for } x \neq 0, \\ 0 & \text{for } x = 0, \end{cases} \tag{2.29}$$

where $\operatorname{graph}(f)$ is closed, but f is not continuous, show that the converse of Prop. 2.28 is not true in general, according to the following Th. 2.30, the converse does hold for linear maps between F-spaces. In preparation, we provide some simple facts regarding the product of two metric spaces:

Proposition 2.29. *Let $(X, \mathcal{S})$, $(Y, \mathcal{T})$ be topological spaces and let $\mathcal{P}$ be the product topology on $X \times Y$. Moreover, let $\mathcal{S}$, $\mathcal{T}$ be induced by metrics d_X and d_Y on X and Y, respectively.*

(a) *The map*

$$d: (X\times Y)^2 \longrightarrow \mathbb{R}_0^+, \quad d\big((x_1,y_1),(x_2,y_2)\big) := d_X(x_1,x_2)+d_Y(y_1,y_2), \tag{2.30}$$

defines a metric on $X\times Y$ that induces $\mathcal{P}$. Moreover, if d_X, d_Y are complete, then so is d. If $(X,\mathcal{S})$, $(Y,\mathcal{T})$ are topological vector spaces over $\mathbb{K}$, then $(X\times Y,\mathcal{P})$ is a topological vector space over $\mathbb{K}$; if d_X, d_Y are translation-invariant, then d is translation-invariant as well.

(b) *If $f: X\longrightarrow Y$, then* graph(f) *is closed if, and only if, for each sequence $(x_k)_{k\in\mathbb{N}}$ in X such that the limits*

$$x := \lim_{k\to\infty} x_k \quad \textit{and} \quad y := \lim_{k\to\infty} f(x_k) \tag{2.31}$$

both exist, one has $y = f(x)$.

Proof. Exercise. ■

Theorem 2.30 (Closed Graph Theorem)**.** *Let $(X,\mathcal{S})$, $(Y,\mathcal{T})$ be F-spaces over $\mathbb{K}$ (e.g. Banach spaces) and let $A: X\longrightarrow Y$ be linear. Then the following statements are equivalent:*

(i) *A is continuous.*

(ii) graph(A) *is closed with respect to the product topology $\mathcal{P}$ on $X\times Y$.*

(iii) *For each sequence $(x_k)_{k\in\mathbb{N}}$ in X such that the limits*

$$x := \lim_{k\to\infty} x_k \quad \textit{and} \quad y := \lim_{k\to\infty} A(x_k) \tag{2.32}$$

both exist with $x=0$, one has $y=0$.

Proof. The equivalence of (ii) and (iii) is due to Prop. 2.29(b): (ii) implies (iii), since $A(0)=0$. If (iii) holds with $\lim_{k\to\infty} x_k = x$ and $\lim_{k\to\infty} A(x_k) = y$, then $\lim_{k\to\infty}(x_k - x) = 0$, $\lim_{k\to\infty} A(x_k - x) = y - A(x)$ and (iii) implies $y - A(x) = 0$. Thus, $y = A(x)$ and Prop. 2.29(b) yields (ii). Next, note that (i) implies (ii) according to Prop. 2.28. Thus, it only remains to show (ii) implies (i). So let $G :=$ graph(A) be closed and consider the map

$$B: X\longrightarrow X\times Y, \quad B(x) := (x, A(x)).$$

Then

$$\underset{a,b\in X}{\forall} \quad \underset{\lambda,\mu\in\mathbb{K}}{\forall} \quad \begin{aligned} B(\lambda a+\mu b) &= \big(\lambda a+\mu b, \lambda A(a)+\mu A(b)\big) \\ &= \lambda\big(a, A(a)\big) + \mu\big(b, B(b)\big) = \lambda B(a) + \mu B(b) \end{aligned}$$

shows B to be linear and $G = B(X)$ to be a vector subspace of $X \times Y$. As a closed subspace of $X \times Y$ (which is an F-space by Prop. 2.29(a)), G is itself an F-space. The projections $\pi_X : X \times Y \longrightarrow X$ and $\pi_Y : X \times Y \longrightarrow Y$ are linear, continuous, and surjective. Moreover, $\pi_X \restriction_G$ is even bijective as well. Thus, $(\pi_X \restriction_G)^{-1} : X \longrightarrow G$ is continuous by Cor. 2.25(b), showing $A = \pi_Y \circ (\pi_X \restriction_G)^{-1}$ to be continuous as well. ∎

2.4.2 Application to Sequence Spaces

As an application of Th. 2.30, we will prove a theorem due to Toeplitz (Th. 2.34 below), regarding linear operators on sequence spaces. We start by recalling/introducing some notation for various sequence spaces:

Notation 2.31. Let $p \in]0, \infty[$ and define

$$\mathbb{K}^{\mathbb{N}} := \Big\{ (x_k)_{k \in \mathbb{N}} : x_k \in \mathbb{K} \Big\}, \tag{2.33a}$$

$$l^p := \Big\{ (x_k)_{k \in \mathbb{N}} \in \mathbb{K}^{\mathbb{N}} : \sum_{k=1}^{\infty} |x_k|^p < \infty \Big\}, \tag{2.33b}$$

$$l^\infty := \Big\{ (x_k)_{k \in \mathbb{N}} \in \mathbb{K}^{\mathbb{N}} : (x_k)_{k \in \mathbb{N}} \text{ bounded} \Big\}, \tag{2.33c}$$

$$c := \Big\{ (x_k)_{k \in \mathbb{N}} \in l^\infty : (x_k)_{k \in \mathbb{N}} \text{ convergent} \Big\}, \tag{2.33d}$$

$$c_0 := \Big\{ (x_k)_{k \in \mathbb{N}} \in c : \lim_{k \to \infty} x_k = 0 \Big\}. \tag{2.33e}$$

Note that the definition of l^p in (2.33b) and (2.33c) is consistent with [Phi17, Def. and Rem. 2.41(b)]. In particular, each l^p is equipped with the norm $\| \cdot \|_p$; c and c_0 are subspaces of $(l^\infty, \| \cdot \|_\infty)$. In particular, it makes sense to define

$$(c_0)' := \Big\{ (\alpha : c_0 \longrightarrow \mathbb{K}) : \alpha \text{ linear and continuous} \Big\}. \tag{2.33f}$$

We endow $(c_0)'$ with the so-called *operator norm*

$$\begin{aligned} \|\alpha\| &:= \sup \left\{ \frac{|\alpha(x)|}{\|x\|_\infty} : x \in c_0,\ x \neq 0 \right\} \\ &= \sup \big\{ |\alpha(x)| : x \in c_0,\ \|x\|_\infty = 1 \big\}. \end{aligned} \tag{2.34}$$

It is a simple exercise to check that the operator norm is, indeed, a norm (we will also come back to operator norms in a more general context later).

Proposition 2.32. (a) *c and c_0 are Banach spaces over $\mathbb{K}$. Moreover,*

$$\lambda : c \longrightarrow \mathbb{K}, \quad \lambda\big((x_k)_{k\in\mathbb{N}}\big) := \lim_{k\to\infty} x_k, \tag{2.35}$$

defines a continuous linear functional.

(b) *We have the representation $(c_0)' \cong l^1$: More precisely, $(c_0)'$ and l^1 are isometrically isomorphic, where a linear isometry is given by*

$$\phi_1 : l^1 \longrightarrow (c_0)', \qquad \phi_1\big((a_k)_{k\in\mathbb{N}}\big)\big((x_k)_{k\in\mathbb{N}}\big) := \sum_{k=1}^{\infty} a_k x_k, \tag{2.36a}$$

$$\phi_1^{-1} : (c_0)' \longrightarrow l^1, \qquad \phi_1^{-1}(f) = \big(f(e_k)\big)_{k\in\mathbb{N}} \tag{2.36b}$$

(e_k denoting the kth standard unit vector in $\mathbb{K}^{\mathbb{N}}$).

Proof. (a): Since c and c_0 are subspaces of the Banach space l^∞, it suffices to show they are closed. Since λ is, clearly, linear, (a) is proved once we have shown, for each sequence $(x^n)_{n\in\mathbb{N}}$ in c converging to $x \in l^\infty$, that $\lambda(x^n)$ converges to some $L \in \mathbb{K}$ and

$$L := \lim_{n\to\infty} \lambda(x^n) = \lim_{k\to\infty} x_k. \tag{2.37}$$

First, we note that

$$\underset{m,n\in\mathbb{N}}{\forall} \quad |\lambda(x^n) - \lambda(x^m)| = |\lambda(x^n - x^m)| \le \|x^n - x^m\|_\infty$$

implies $(\lambda(x^n))_{n\in\mathbb{N}}$ to be Cauchy in $\mathbb{K}$, i.e. $L := \lim_{n\to\infty} \lambda(x^n) \in \mathbb{K}$ exists. To prove (2.37), let $\epsilon > 0$. Let $N \in \mathbb{N}$ be such that

$$\|x^N - x\|_\infty < \frac{\epsilon}{3} \quad \text{and} \quad |\lambda(x^N) - L| < \frac{\epsilon}{3}$$

and choose $M \in \mathbb{N}$ such that

$$\underset{n\ge M}{\forall} \quad |x^N_n - \lambda(x^N)| < \frac{\epsilon}{3}.$$

Then

$$\underset{n\ge M}{\forall} \quad |x_n - L| \le |x_n - x^N_n| + |x^N_n - \lambda(x^N)| + |\lambda(x^N) - L| < 3 \cdot \frac{\epsilon}{3} = \epsilon,$$

proving (2.37) and (a).

(b): Exercise. ■

Proposition 2.33. *Let $A = (a_{ij}) \in \mathbb{K}^{\mathbb{N}\times\mathbb{N}}$ be a double sequence (some authors think of such an A as a countable matrix). We identify A with the map*

$$A : \mathbb{K}^{\mathbb{N}} \longrightarrow (\mathbb{K}^{\mathbb{N}})^{\mathbb{N}}, \quad A\big((x_j)_{j\in\mathbb{N}}\big) := \left(\sum_{j=1}^{\infty} a_{ij} x_j\right)_{i\in\mathbb{N}}, \tag{2.38}$$

where each $\sum_{j=1}^{\infty} a_{ij} x_j$ is meant to be a sequence of partial sums (the series does not necessarily converge). Let $X \subseteq \mathbb{K}^{\mathbb{N}}$ be a vector subspace. We call A simple *on X if, and only if, $A(x)$ converges (in $\mathbb{K}$) for each $x \in X$; for A simple on X, we also write $A : X \longrightarrow \mathbb{K}^{\mathbb{N}}$, i.e. we then consider A as a map into $\mathbb{K}^{\mathbb{N}}$. A given A might or might not satisfy some or all of the following conditions* (2.39a) – (2.39c)*, where*

$$\underset{j\in\mathbb{N}}{\forall} \quad \lim_{i\to\infty} a_{ij} = 0, \tag{2.39a}$$

$$M := \sup\left\{\sum_{j=1}^{\infty} |a_{ij}| : i \in \mathbb{N}\right\} < \infty, \tag{2.39b}$$

$$\text{all} \sum_{j=1}^{\infty} a_{ij} \text{ converge with} \quad \lim_{i\to\infty} \sum_{j=1}^{\infty} a_{ij} = 1. \tag{2.39c}$$

The following statements hold true for each double sequence A:

(a) *A is linear. If A is simple on X, then $A : X \longrightarrow \mathbb{K}^{\mathbb{N}}$ is linear as well.*

(b) *If* (2.39b) *holds, then A is simple on l^∞ and $A : l^\infty \longrightarrow l^\infty$ is a continuous linear map.*

(c) *If* (2.39a) *and* (2.39b) *hold, then $A(c_0) \subseteq c_0$.*

(d) (2.39c) *is equivalent to $y := A\big((1)_{j\in\mathbb{N}}\big) \in c$ with $\lambda(y) = 1$, where λ is the functional of* (2.35)*.*

(e) *If A is simple on c_0 with $A(c_0) \subseteq c_0$, then* (2.39b) *is equivalent to $A : c_0 \longrightarrow c_0$ being continuous.*

Proof. (a): The linearity of A, in the general as well as in the simple case, is immediate from the properties of finite sums, series, and limits in $\mathbb{K}$.

(b): If $x = (x_j)_{j\in\mathbb{N}} \in l^\infty$ and (2.39b) holds, then

$$\underset{i\in\mathbb{N}}{\forall} \quad \sum_{j=1}^{\infty} |a_{ij}|\,|x_j| \leq M\,\|x\|_\infty,$$

showing A to be simple on l^∞ with $\|A(x)\|_\infty \leq M\,\|x\|_\infty$ for each $x \in l^\infty$. In particular, the linear map $A: l^\infty \longrightarrow l^\infty$ is bounded and, thus, continuous.

(c): Exercise.

(d) is immediate from $y = A\big((1)_{j\in\mathbb{N}}\big) = (\sum_{j=1}^{\infty} a_{ij})_{i\in\mathbb{N}}$.

(e): For each $i \in \mathbb{N}$, let $A_i := (a_{ij})_{j\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$. Then (2.39b) is equivalent to $\{A_i : i \in \mathbb{N}\}$ being a bounded subset of l^1. Due to Prop. 2.32(b), (2.39b) is further equivalent to

$$M = \sup\{\|\phi_1(A_i)\| : i \in \mathbb{N}\} < \infty, \tag{2.40}$$

where $\|\cdot\|$ denotes the operator norm on $(c_0)'$. However, (2.40) is equivalent to

$$\underset{M'\in\mathbb{R}_0^+}{\exists}\quad \underset{i\in\mathbb{N}}{\forall}\quad \underset{x\in c_0}{\forall}\quad \left|\sum_{j=1}^{\infty} a_{ij}x_j\right| \leq M'\|x\|_\infty,$$

which is equivalent to

$$\underset{M'\in\mathbb{R}_0^+}{\exists}\quad \underset{x\in c_0}{\forall}\quad \|A(x)\|_\infty \leq M'\|x\|_\infty,$$

i.e. to $A: c_0 \longrightarrow c_0$ being bounded, i.e. to $A: c_0 \longrightarrow c_0$ being continuous. ■

Theorem 2.34 (Toeplitz)**.** *Let $A = (a_{ij}) \in \mathbb{K}^{\mathbb{N}\times\mathbb{N}}$ and identify A with the map of* (2.38)*. Then the following statements* (i) *and* (ii) *are equivalent:*

(i) *A is simple on c (in the sense of Prop.* 2.33*), maps c into c, and satisfies*

$$\underset{x:=(x_j)_{j\in\mathbb{N}}\in c}{\forall}\quad \lambda(A(x)) = \lim_{i\to\infty}\sum_{j=1}^{\infty} a_{ij}x_j = \lim_{j\to\infty} x_j = \lambda(x), \tag{2.41}$$

where λ is the functional of (2.35)*.*

(ii) *$A = (a_{ij})$ satisfies the above conditions* (2.39a) – (2.39c)*.*

Proof. Theorem 2.30 will be applied to show (i) implies (2.39b). However, we begin with the other direction:

“(ii)⇒(i)”: Let $x \in c$ with $\xi := \lambda(x) \in \mathbb{K}$. We identify ξ with $(\xi)_{j\in\mathbb{N}} \in c$. Then $x - \xi \in c_0$ and $A(x-\xi) \in c_0$ by Prop. 2.33(c). Since

$$A(x) = A(x-\xi) + A(\xi) = \left(\sum_{j=1}^{\infty} a_{ij}(x_j - \xi)\right)_{i\in\mathbb{N}} + \left(\xi\sum_{j=1}^{\infty} a_{ij}\right)_{i\in\mathbb{N}}$$

with $\lambda(A(x-\xi)) = 0$ and, by Prop. 2.33(d), $\lambda(A(\xi)) = \xi$, we obtain $A(x) \in c$ with $\lambda(A(x)) = \lambda(A(x-\xi)) + \lambda(A(\xi)) = \xi = \lambda(x)$, as desired.

"(i)⇒(ii)": For each $j \in \mathbb{N}$, let e_j denote the standard unit vector in $\mathbb{K}^{\mathbb{N}}$. Then $A(e_j) = (a_{ij})_{i\in\mathbb{N}}$ for each $j \in \mathbb{N}$ and $0 = \lambda(e_j) = \lambda(A(e_j))$ yields (2.39a). Moreover, (2.39c) is satisfied by Prop. 2.33(d). It remains to verify the validity of (2.39b). According to Prop. 2.33(e), it suffices to show that $A : c_0 \longrightarrow c_0$ is continuous. According to the closed graph Th. 2.30, we have to consider a sequence $(x^k)_{k\in\mathbb{N}}$ in c_0 such that $\lim_{k\to\infty} x^k = 0$ and such that $y := \lim_{k\to\infty} A(x^k)$ exists in c_0, showing $y = 0$. For each $i \in \mathbb{N}$, we consider the linear functional

$$A_i : c_0 \longrightarrow \mathbb{K}, \quad A_i(x) := \sum_{j=1}^{\infty} a_{ij} x_j.$$

Moreover, for each $i, n \in \mathbb{N}$, we also consider the continuous linear functional

$$A_{in} : c_0 \longrightarrow \mathbb{K}, \quad A_{in}(x) := \sum_{j=1}^{n} a_{ij} x_j$$

(where the continuity of A_{in} can be verified directly, but is also immediate from Prop. 2.32(b)). Then

$$\underset{i\in\mathbb{N}}{\forall} \quad \underset{x\in c_0}{\forall} \quad A_i(x) = \lim_{n\to\infty} A_{in}(x),$$

showing the pointwise convergence of A_{in} to A_i for $n \to \infty$. Thus, each A_i is continuous (i.e. $A_i \in (c_0)'$) by Th. 2.17(d) and each $(a_{ij})_{j\in\mathbb{N}} \in l^1$ by Prop. 2.32(b). Thus,

$$\underset{i\in\mathbb{N}}{\forall} \quad |A_i(x^k)| \leq \sum_{j=1}^{\infty} |a_{ij}|\,|x_j^k| \leq \left\|(a_{ij})_{j\in\mathbb{N}}\right\|_1 \|x^k\|_\infty \quad \to 0 \text{ for } k \to \infty,$$

showing $A(x^k)$ to converge to 0 in each component (i.e. pointwise). Thus, if $y = \lim_{k\to\infty} A(x^k)$ exists in c_0 (i.e. $A(x^k)$ converges with respect to $\|\cdot\|_\infty$, i.e. uniformly), then $y = 0$ as desired, showing A to be continuous and completing the proof. ■

3 Convexity

3.1 Hahn-Banach Theorems

In the literature, several extension and separation theorems in regard to linear functionals are associated with the name *Hahn-Banach*. We study such theorems in the present section. In this spirit, there will not be *the* Hahn-Banach theorem, but rather a number of theorems, each of which might subsequently be referred to as a Hahn-Banach theorem. It is often not easy to establish the existence of nontrivial continuous linear functionals (or, more generally, of linear functionals with desired prescribed properties). The Hahn-Banach theorems are designed to be applied in precisely this kind of situation. Theorem 3.3 is a dominated extension theorem that does not involve any topology, but has immediate applications to seminormed spaces in Cor. 3.4. Theorem 3.5 and Cor. 3.7 are separation-type Hahn-Banach theorems holding on topological vector spaces. We begin with some notation and some simple observations regarding the relation between $\mathbb{R}$-linear and $\mathbb{C}$-linear functionals:

Definition 3.1. Let $(X, \mathcal{T}_X)$, $(Y, \mathcal{T}_Y)$ be topological vector spaces over $\mathbb{K}$. We introduce the following notation:

$$\mathcal{L}(X,Y) := \{(A : X \longrightarrow Y) : A \text{ linear and continuous}\}, \tag{3.1a}$$

$$X' := \mathcal{L}(X, \mathbb{K}). \tag{3.1b}$$

The space X' is called the *dual space* (or just the *dual*) of X (in the literature, one often also finds the notation X^* instead of X').

—

Clearly,

$$\forall_{z \in \mathbb{C}} \quad z = \operatorname{Re} z + i \operatorname{Im} z = \operatorname{Re} z - i \operatorname{Re}(iz) \tag{3.2}$$

and every vector space over $\mathbb{C}$ is also a vector space over $\mathbb{R}$.

Lemma 3.2. *Let X be a vector space over $\mathbb{C}$.*

(a) *If $\alpha : X \longrightarrow \mathbb{C}$ is $\mathbb{C}$-linear, then $\operatorname{Re} \alpha : X \longrightarrow \mathbb{R}$ is $\mathbb{R}$-linear with*

$$\forall_{x \in X} \quad \alpha(x) = \operatorname{Re} \alpha(x) - i \operatorname{Re} \alpha\big(ix\big).$$

(b) *If $f : X \longrightarrow \mathbb{R}$ is $\mathbb{R}$-linear, then*

$$\alpha : X \longrightarrow \mathbb{C}, \quad \alpha(x) := f(x) - i\, f(ix),$$

is $\mathbb{C}$-linear.

(c) *Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{C}$. Let $\alpha : X \longrightarrow \mathbb{C}$ be $\mathbb{C}$-linear. Then α is continuous (i.e. $\alpha \in X'$) if, and only if, $\operatorname{Re} \alpha$ is continuous.*

(d) *Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{C}$. Let $f : X \longrightarrow \mathbb{R}$ be $\mathbb{R}$-linear and continuous. Then there exists a unique $\alpha \in X'$ such that $f = \operatorname{Re} \alpha$.*

Proof. Exercise (use (3.2) in the proof of (a)). ∎

Theorem 3.3. *Let X be a vector space over $\mathbb{R}$ and $V \subseteq X$ a vector subspace. Moreover, consider a function $p: X \longrightarrow \mathbb{R}$, satisfying*

$$\forall_{x,y\in X} \quad p(x+y) \leq p(x) + p(y), \tag{3.3a}$$

$$\forall_{x\in X} \quad \forall_{t\in\mathbb{R}_0^+} \quad p(tx) = t\,p(x). \tag{3.3b}$$

(e.g., each seminorm on X satisfies the above conditions), and a linear functional $\alpha: V \longrightarrow \mathbb{R}$ that is dominated by p, i.e. such that

$$\forall_{v\in V} \quad \alpha(v) \leq p(v). \tag{3.3c}$$

Then α can be linearly extended to X such that the extension is dominated by p on all of X: More precisely, there exists a linear functional $\beta: X \longrightarrow \mathbb{R}$ such that $\beta\restriction_V = \alpha$ and

$$\forall_{x\in X} \quad -p(-x) \leq \beta(x) \leq p(x). \tag{3.4}$$

Proof. First note that domination and linearity imply (3.4): If $\alpha(v) \leq p(v)$ and $\alpha(-v) \leq p(-v)$, then $-p(-v) \leq -\alpha(-v) = \alpha(v) \leq p(v)$. Thus, if $V = X$, then we can set $\beta := \alpha$. It remains to consider $V \neq X$. We let $x_1 \in X \setminus V$ and first show how to extend α to

$$V_1 := \operatorname{span}(V \cup \{x_1\}) = \{v + \lambda x_1 : v \in V, \lambda \in \mathbb{R}\}:$$

From

$$\forall_{u,v\in V} \quad \alpha(u) + \alpha(v) = \alpha(u+v) \leq p(u+v) \leq p(u - x_1) + p(x_1 + v)$$

we obtain

$$\forall_{u,v\in V} \quad \alpha(u) - p(u - x_1) \leq p(x_1 + v) - \alpha(v)$$

and, hence,

$$\sigma := \sup\big\{\alpha(u) - p(u - x_1) : u \in V\big\} \leq p(x_1) < \infty.$$

Then

$$\forall_{u,v\in V} \quad \alpha(u) - p(u - x_1) \leq \sigma \leq p(x_1 + v) - \alpha(v),$$

implying

$$\forall_{u\in V} \quad \alpha(u) - \sigma \leq p(u - x_1), \tag{3.5a}$$

$$\forall_{v\in V} \quad \alpha(v) + \sigma \leq p(x_1 + v). \tag{3.5b}$$

Define

$$\alpha_1 : V_1 \longrightarrow \mathbb{R}, \quad \alpha_1(v + \lambda x_1) := \alpha(v) + \lambda\sigma.$$

Then, clearly, α_1 is linear with $\alpha_1\restriction_V = \alpha$. To obtain $\alpha_1 \leq p$, for each $\lambda \in \mathbb{R}^+$, replace u by $\lambda^{-1}u$ in (3.5a), replace v by $\lambda^{-1}v$ in (3.5b), and multiply the respective result by t to obtain

$$\underset{u\in V}{\forall} \quad \alpha(u) - \lambda\sigma \leq p(u - \lambda x_1),$$
$$\underset{v\in V}{\forall} \quad \alpha(v) + \lambda\sigma \leq p(\lambda x_1 + v),$$

indeed, implying $\alpha_1 \leq p$. We can now apply Zorn's lemma to finish the proof: Define a partial order on the set

$$\mathcal{S} := \{(W, \gamma) : V \subseteq W \subseteq X,\, W \text{ vector space},\, \gamma : W \longrightarrow \mathbb{R} \text{ linear},\, \gamma\restriction_V = \alpha,\, \gamma \leq p\}$$

by letting

$$(W, \gamma) \leq (W', \gamma') \quad :\Leftrightarrow \quad W \subseteq W', \quad \gamma'\restriction_W = \gamma.$$

Every chain $\mathcal{C}$, i.e. every totally ordered subset of $\mathcal{S}$, has an upper bound, namely $(W_\mathcal{C}, \gamma_\mathcal{C})$ with $W_\mathcal{C} := \bigcup_{(W,\gamma)\in\mathcal{C}} W$ and $\gamma_\mathcal{C}(x) := \gamma(x)$, where $(W, \gamma) \in \mathcal{C}$ is chosen such that $x \in W$ (since $\mathcal{C}$ is a chain, the value of $\gamma_\mathcal{C}(x)$ does not actually depend on the choice of $(W, \gamma) \in \mathcal{C}$ and is, thus, well-defined). Clearly, $W_\mathcal{C}$ is a vector subspace of X, $V \subseteq W_\mathcal{C}$, and $\gamma_\mathcal{C}$ extends α, and $\gamma_\mathcal{C} \leq p$, i.e. $(W_\mathcal{C}, \gamma_\mathcal{C}) \in \mathcal{S}$ (that $(W_\mathcal{C}, \gamma_\mathcal{C})$ is an upper bound for $\mathcal{C}$ is then immediate). Thus, all hypotheses of Zorn's lemma have been verified and we obtain the existence of a maximal element $(W_{\max}, \gamma_{\max}) \in \mathcal{S}$. But then $W_{\max} = X$, since, otherwise, we could extend $\gamma_{\max}$ to $W_{\max} + \operatorname{span}\{x_0\}$, $x_0 \notin W_{\max}$, as in the first step of the proof, in contradiction to the maximality of $(W_{\max}, \gamma_{\max})$. Thus, we can set $\beta := \gamma_{\max}$ to complete the proof. ■

Corollary 3.4. *Let $(X, \|\cdot\|)$ be a seminormed vector space over $\mathbb{K}$.*

(a) *Let $V \subseteq X$ be a vector subspace of X and let $\alpha : V \longrightarrow \mathbb{K}$ be a linear functional such that*

$$\underset{v\in V}{\forall} \quad |\alpha(v)| \leq \|v\|. \tag{3.6}$$

Then there exists $\beta \in X'$ such that $\beta\restriction_V = \alpha$ and

$$\underset{x\in X}{\forall} \quad |\beta(x)| \leq \|x\|. \tag{3.7}$$

(b) *For each $x_0 \in X$, there exists $\beta \in X'$ such that* (3.7) *holds as well as*

$$\beta(x_0) = \|x_0\|. \tag{3.8}$$

Proof. First note that (3.7) implies β to be continuous (e.g., by Th. 1.14(iv), as $\beta(B_1(0))$ is bounded).

(a): If $\mathbb{K} = \mathbb{R}$, then we merely apply Th. 3.3 with $p := \|\cdot\|$, where (3.7) follows from (3.4), since $p(-x) = \|-x\| = \|x\| = p(x)$. If $\mathbb{K} = \mathbb{C}$, then $f := \operatorname{Re}\alpha$ is $\mathbb{R}$-linear on V and we can extend it to some $\mathbb{R}$-linear map $f : X \longrightarrow \mathbb{R}$, satisfying (3.7) (with β replaced by f). According to Lem. 3.2(a),(b), there exists a unique $\mathbb{C}$-linear $\beta : X \longrightarrow \mathbb{C}$ with $\operatorname{Re}\beta = f$. Using Lem. 3.2(a) again, then yields $\beta{\restriction_V} = \alpha$. Since (3.7), clearly, holds for $\beta(x) = 0$, it only remains to check it holds for each $x \in X$ with $\beta(x) \neq 0$. However, if $\beta(x) \neq 0$, then

$$|\beta(x)| = \frac{|\beta(x)|}{\beta(x)} \cdot \beta(x) = \beta\left(\frac{|\beta(x)|}{\beta(x)} \cdot x\right) = f\left(\frac{|\beta(x)|}{\beta(x)} \cdot x\right) \leq \left\|\frac{|\beta(x)|}{\beta(x)} \cdot x\right\| = \|x\|,$$

completing the proof of (a).

(b): If $\|x_0\| = 0$, then $\beta \equiv 0$ works. If $\|x_0\| \neq 0$, then, to apply (a), let $V := \operatorname{span}\{x_0\}$ and define

$$\alpha : V \longrightarrow \mathbb{K}, \quad \alpha(\lambda x_0) := \lambda\,\|x_0\|.$$

Then (3.6) holds and we obtain a linear extension β, satisfying (3.7) by (a). ■

Theorem 3.5. *Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{K}$. Assume $A, B \subseteq X$ such that A, B are nonempty and convex with $A \cap B = \emptyset$.*

(a) *If A is open, then there exist $\alpha \in X'$ and $s \in \mathbb{R}$ such that*

$$\underset{a\in A}{\forall}\ \underset{b\in B}{\forall}\quad \operatorname{Re}\alpha(a) < s \leq \operatorname{Re}\alpha(b). \tag{3.9}$$

(b) *If A is compact, B is closed, and $(X, \mathcal{T})$ is locally convex, then there exist $\alpha \in X'$ and $s_1, s_2 \in \mathbb{R}$ such that*

$$\underset{a\in A}{\forall}\ \underset{b\in B}{\forall}\quad \operatorname{Re}\alpha(a) < s_1 < s_2 < \operatorname{Re}\alpha(b). \tag{3.10}$$

Proof. Suppose, we have proved the theorem for $\mathbb{K} = \mathbb{R}$. Then it also holds for $\mathbb{K} = \mathbb{C}$: We first obtain a continuous $\mathbb{R}$-linear function $f : X \longrightarrow \mathbb{R}$, satisfying (3.9) or (3.10), respectively (with α replaced by f). Then, by Lem. 3.2(d), there exists a unique $\alpha \in X'$ with $\operatorname{Re}\alpha = f$, proving the case $\mathbb{K} = \mathbb{C}$. It remains to prove (a),(b) for $\mathbb{K} = \mathbb{R}$.

(a): Fix $a_0 \in A$, $b_0 \in B$, and set $x_0 := b_0 - a_0$. If $C := A - B + x_0$, then $C \in \mathcal{U}(0)$ (C is open by Prop. 1.10(d)). Moreover, C is convex due to Prop. 1.8(a). Let $p := \mu_C$ be the Minkowski functional of C as defined in Def. and Rem. 1.35(b). Then

$$A \cap B = \emptyset \quad \Rightarrow \quad x_0 \notin C \quad \Rightarrow \quad p(x_0) \geq 1.$$

The idea is now to apply Th. 3.3, where we note that p satisfies (3.3a) and (3.3b) due to Th. 1.37(b),(e). Define

$$V := \operatorname{span}\{x_0\}, \quad \alpha : V \longrightarrow \mathbb{R}, \quad \alpha(\lambda x_0) := \lambda,$$

and note

$$\left.\begin{array}{l} \lambda \geq 0 \;\Rightarrow \alpha(\lambda x_0) = \lambda \leq \lambda p(x_0) = p(\lambda x_0), \\ \lambda < 0 \;\Rightarrow \alpha(\lambda x_0) = \lambda < 0 \leq p(\lambda x_0) \end{array}\right\} \;\Rightarrow\; \alpha \leq p \text{ on } V,$$

showing p to satisfy (3.3c) as well. Thus, by Th. 3.3, α has a linear extension defined on X such that $\alpha \leq p$ holds on all of X. In particular, $\alpha \leq 1$ on C and $\alpha \geq -1$ on $-C$, implying $|\alpha| \leq 1$ on $U := C \cap (-C) \in \mathcal{U}(0)$. Thus, $\alpha \in X'$ by Th. 1.14(iv). Next, we observe

$$\underset{a\in A}{\forall} \quad \underset{b\in B}{\forall} \quad \alpha(a) - \alpha(b) + 1 = \alpha(a - b + x_0) \leq p(a - b + x_0) < 1,$$

where the last inequality is due to $a - b + x_0 \in C$ open and (1.25). Thus, $\alpha(a) < \alpha(b)$, showing $\alpha(A)$ and $\alpha(B)$ to be disjoint. Since A, B are convex and α is linear, $\alpha(A)$ and $\alpha(B)$ are also convex, i.e. they are intervals with $\sup \alpha(A) \leq \inf \alpha(B)$. Moreover, α is open by Prop. 2.21(b), i.e. $\alpha(A)$ is open (as A is open by hypothesis). Thus, $\sup \alpha(A) \notin \alpha(A)$ and we can set $s := \sup \alpha(A)$ to finish the proof of (a).

(b): If A is compact and B is closed, then, by (1.5b), there exists a convex open neighborhood V of 0 in X such that $(A + V) \cap B = \emptyset$. According to (a), there exists $\alpha \in X'$ such that $\alpha(A+V)$ and $\alpha(B)$ are disjoint intervals with $\sup \alpha(A+V) \leq \inf \alpha(B)$. Now $\alpha(A+V)$ is open with $\alpha(A)$ as a compact subinterval, proving (3.10) and (b). ■

Definition 3.6. Let I, X, X_i be sets, $i \in I \neq \emptyset$. Given a family of functions $\mathcal{F} := (f_i)_{i\in I}$, $f_i : X \longrightarrow X_i$, we say that $\mathcal{F}$ *separates points* on X if, and only if,

$$\underset{x_1,x_2\in X}{\forall} \quad \left(x_1 \neq x_2 \quad \Rightarrow \quad \underset{i\in I}{\exists} \; f_i(x_1) \neq f_i(x_2) \right). \tag{3.11}$$

If $Y = X_i$ for each $i \in I$ and $\mathcal{M} \subseteq Y^X$, then we say $\mathcal{M}$ *separates points* on X if, and only if, $(f)_{f\in\mathcal{M}}$ separates points on X.

Corollary 3.7. *Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{K}$ that is locally convex (e.g. a normed space).*

(a) *If X is T_1, then X' separates points on X.*

(b) *Let $V \subseteq X$ be a vector subspace and $x_0 \in X$ such that $x_0 \notin \overline{V}$. Then*

$$\underset{\alpha\in X'}{\exists} \quad \Big(\alpha\restriction_V = 0 \;\wedge\; \alpha(x_0) = 1 \Big).$$

(c) *If $V \subseteq X$ is a vector subspace and $\alpha \in V'$, then α has a continuous linear extension to X, i.e.*

$$\underset{\beta \in X'}{\exists} \quad \beta\restriction_V = \alpha.$$

(d) *Let $B \subseteq X$ be convex, balanced, and closed. If $x_0 \in X \setminus B$, then*

$$\underset{\alpha \in X'}{\exists} \quad \Big(|\alpha\restriction_B| \leq 1 \ \wedge \ \alpha(x_0) \in]1, \infty[\Big).$$

Proof. Exercise. ∎

Example 3.8. In an application of Hahn-Banach to sequence spaces, we show that not every $\alpha \in (l^\infty)'$ can be represented by some $x \in l^1$: In other words, we will show that the map

$$\phi : l^1 \longrightarrow (l^\infty)', \quad \phi\big((a_k)_{k\in\mathbb{N}}\big)\big((x_k)_{k\in\mathbb{N}}\big) := \sum_{k=1}^{\infty} a_k x_k, \tag{3.12}$$

is not surjective. First, we note that ϕ is defined by the same formula as $\phi_1 : l^1 \longrightarrow (c_0)'$ in (2.36a) and that the same estimate used in the proof of Prop. 2.32(b), namely

$$|\phi(a)(x)| \leq \sum_{k=1}^{\infty} |a_k x_k| \leq \|x\|_\infty \, \|a\|_1,$$

shows $(a_k x_k)_{k\in\mathbb{N}} \in l^1$ and $\phi(a)$ to be bounded (and, thus, continuous, i.e. $\phi(a) \in (l^\infty)'$). From Prop. 2.32(a), we know the functional

$$\lambda : c \longrightarrow \mathbb{K}, \quad \lambda\big((x_k)_{k\in\mathbb{N}}\big) := \lim_{k\to\infty} x_k,$$

to be linear and continuous. Thus, we can apply Hahn-Banach in the form of Cor. 3.7(c) with $X := l^\infty$ and $V := c$ to obtain $\beta \in (l^\infty)'$ such that $\beta\restriction_c = \lambda$. Suppose there were $a \in l^1$ such that $\beta = \phi(a)$. Letting e_k denote the kth standard unit vector in $\mathbb{K}^\mathbb{N}$, we obtain

$$\underset{k\in\mathbb{N}}{\forall} \quad a_k = \beta(e_k) = \lambda(e_k) = 0,$$

implying $a = 0$ and $\beta = 0$. This contradiction shows $\beta \notin \phi(l^1)$, i.e. ϕ is not surjective.

3.2 Weak Topology, Weak Convergence

Definition 3.9. Let X be a set and let $\mathcal{T}_1, \mathcal{T}_2$ be topologies on X. We call $\mathcal{T}_1$ *weaker* or *smaller* or *coarser* than $\mathcal{T}_2$, and we call $\mathcal{T}_2$ *stronger* or *bigger* or *finer* than $\mathcal{T}_1$, if and only if, $\mathcal{T}_1 \subseteq \mathcal{T}_2$.

—

Given a topological vector space $(X, \mathcal{T})$, we know that (by definition) each $\alpha \in X'$ is continuous. It turns out to be useful to also consider another topology $\mathcal{T}_{\mathrm{w}}$ on X, the so-called *weak topology*, consisting of the weakest (i.e. the smallest) topology on X making each $\alpha \in X'$ continuous, cf. Def. 3.14. We will see that, in many interesting cases, $\mathcal{T}_{\mathrm{w}}$ is strictly weaker than $\mathcal{T}$. The weak topology $\mathcal{T}_{\mathrm{w}}$ is the so-called *initial topology* on X with respect to X'. It is a special case of the contruction defined in the following Def. 3.10, which is also of general interest (another initial topology we will encounter shortly is the so-called weak*-topology on X', cf. Rem. and Def. 3.21)[1]:

Definition 3.10. Let X be a set and let $((X_i, \mathcal{T}_i))_{i\in I}$ be a family of topological spaces, $I \neq \emptyset$. Given a family of functions $\mathcal{F} := (f_i)_{i\in I}$, $f_i : X \longrightarrow X_i$, the *initial* or *weak topology* on X with respect to the family $(f_i)_{i\in I}$ (also called the $\mathcal{F}$-topology on X) is the coarsest topology $\mathcal{T}$ on X that makes all f_i continuous (i.e. $\mathcal{T}$ is the intersection of all topologies that make all f_i continuous – this intersection is well-defined, since the discrete topology on X always makes all f_i continuous). The name initial topology stems from the f_i being *initially* in X. If $Y = X_i$ for each $i \in I$ and $\mathcal{M} \subseteq Y^X$, then the $\mathcal{M}$-topology on X is the $\mathcal{F}$-topology on X, where $\mathcal{F} := (f)_{f\in\mathcal{M}}$.

Lemma 3.11. *Let X be a set and let $((X_i, \mathcal{T}_i))_{i\in I}$ be a family of topological spaces, $I \neq \emptyset$. Given a family of functions $(f_i)_{i\in I}$, $f_i : X \longrightarrow X_i$, the set*

$$\mathcal{S} := \left\{ f_i^{-1}(O_i) : O_i \in \mathcal{T}_i,\ i \in I \right\} \tag{3.13}$$

is a subbase of the initial topology $\mathcal{T}$ on X with respect to the family $(f_i)_{i\in I}$.

Proof. Let $\tau(\mathcal{S})$ be the topology on X generated by $\mathcal{S}$, and let $\mathcal{T}'$ be an arbitrary topology on X that makes all f_i continuous. Then, clearly, $\mathcal{S} \subseteq \mathcal{T}'$, also implying $\tau(\mathcal{S}) \subseteq \mathcal{T}'$. Thus, $\tau(\mathcal{S}) \subseteq \mathcal{T}$. On the other hand, by the definition of $\mathcal{S}$, $\tau(\mathcal{S})$ also has the property of making every f_i continuous, proving $\tau(\mathcal{S}) = \mathcal{T}$. ■

Proposition 3.12. *Let X be a set and let $((X_i, \mathcal{T}_i))_{i\in I}$ be a family of topological spaces, $I \neq \emptyset$.*

(a) *Given a family of functions $(f_i)_{i\in I}$, $f_i : X \longrightarrow X_i$, let $\mathcal{T}$ denote the initial topology on X with respect to the family $(f_i)_{i\in I}$. Then $\mathcal{T}$ has the property that each map $g : Z \longrightarrow X$ from a topological space $(Z, \mathcal{T}_Z)$ into X is continuous if, and only if, each map $(f_i \circ g) : Z \longrightarrow X_i$ is continuous. Moreover, $\mathcal{T}$ is the* only *topology on X with this property.*

[1]Some readers might be familiar with the present treatment of initial topologies from [Phi16b, Sec. D.2].

(b) *Let $(x_k)_{k\in\mathbb{N}}$ be a sequence in X, $x \in X$. Then*

$$\lim_{k\to\infty} x_k = x \quad \Leftrightarrow \quad \forall_{i\in I} \quad \lim_{k\to\infty} f_i(x_k) = f_i(x). \tag{3.14}$$

Proof. (a): If g is continuous, then each composition $f_i \circ g$, $i \in I$, is also continuous. For the converse, assume that, for each $i \in I$, $f_i \circ g$ is continuous. If $O \in \mathcal{S}$, where $\mathcal{S}$ is the subbase from (3.13), then there exist $i \in I$ and $O_i \in \mathcal{T}_i$ such that $O = f_i^{-1}(O_i)$. Since $f_i \circ g$ is continuous, we have

$$g^{-1}(O) = g^{-1}\big(f_i^{-1}(O_i)\big) = (f_i \circ g)^{-1}(O_i) \in \mathcal{T}_Z,$$

proving the continuity of g. Now let $\mathcal{A}$ be an arbitrary topology on X with the property stated in the hypothesis. Letting $(Z, \mathcal{T}_Z) := (X, \mathcal{A})$ and $g := \mathrm{Id}_X$, we see that each f_i is continuous with respect to $\mathcal{A}$, implying $\mathcal{T} \subseteq \mathcal{A}$. Now let $\mathcal{T}'$ be an arbitrary topology on X that makes all f_i continuous. Letting $(Z, \mathcal{T}_Z) := (X, \mathcal{T}')$, we see that $g := \mathrm{Id}_X$ is $\mathcal{T}'$-$\mathcal{A}$ continuous (since each $f_i = \mathrm{Id}_X \circ f_i$ is $\mathcal{T}'$-$\mathcal{T}_i$ continuous) i.e., for each $O \in \mathcal{A}$, we have $g^{-1}(O) = O \in \mathcal{T}'$, showing $\mathcal{A} \subseteq \mathcal{T}'$ and $\mathcal{A} \subseteq \mathcal{T}$, also completing the proof of $\mathcal{A} = \mathcal{T}$.

(b): If $\lim_{k\to\infty} x_k = x$ and $i \in I$, then f_i is continuous, implying $\lim_{k\to\infty} f_i(x_k) = f_i(x)$. Conversely, assume $\lim_{k\to\infty} f_i(x_k) = f_i(x)$ holds for each $i \in I$. Let $O \in \mathcal{S}$ with $x \in O$, where $\mathcal{S}$ is the subbase from (3.13). Then there exist $i \in I$ and $O_i \in \mathcal{T}_i$ such that $O = f_i^{-1}(O_i)$. Then there exists $N \in \mathbb{N}$ such that

$$\forall_{k>N} \quad \Big(f_i(x_k) \in O_i, \quad \text{i.e.} \quad x_k \in O = f_i^{-1}(O_i)\Big),$$

showing $\lim_{k\to\infty} x_k = x$ by [Phi16b, Cor. 1.50(a)]. ∎

Example 3.13. **(a)** The product topology on $X = \prod_{i\in I} X_i$ is the initial topology with respect to the projections $(\pi_i)_{i\in I}$, $\pi_i : X \longrightarrow X_i$ (as is clear from Lem. 3.11).

(b) The subspace topology on $M \subseteq X$, where $(X, \mathcal{T})$ is a topological space is the initial topology with respect to the identity inclusion map $\iota : M \longrightarrow X$, $\iota(x) := x$: This is also clear from Lem. 3.11, since

$$\mathcal{T}_M = \big\{O \cap M : O \in \mathcal{T}\big\} = \big\{\iota^{-1}(O) : O \in \mathcal{T}\big\}.$$

Definition 3.14. Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{K}$. We call the X'-topology on X in the sense of Def. 3.10, the *weak* topology on X, denoted by $\mathcal{T}_{\mathrm{w}}$. We will use terms such as *weakly convergent* (and write $x_n \rightharpoonup x$ to denote such weak convergence), *weakly closed*, *weakly compact*, *weakly bounded* etc. to refer to notions in the space $(X, \mathcal{T}_{\mathrm{w}})$; in contrast, $\mathcal{T}$ is called the *strong* or *original* topology on X and the corresponding notions in $(X, \mathcal{T})$ are sometimes called *strongly* or *originally* convergent, closed, etc.

—

The following Lem. 3.15 will be used in the proof of Th. 3.16(d).

Lemma 3.15. *Let X be a vector space over $\mathbb{K}$, let $\alpha, \alpha_1, \dots, \alpha_n : X \longrightarrow \mathbb{K}$ be linear functionals, $n \in \mathbb{N}$, and let*

$$N := \bigcap_{i=1}^{n} \ker \alpha_i.$$

Then (i) – (iii) *are equivalent, where*

(i) *α is a linear combination of $\alpha_1, \dots, \alpha_n$.*

(ii) *There exists $M \in \mathbb{R}_0^+$ such that*

$$\forall_{x \in X} \quad |\alpha(x)| \leq M \max\big\{|\alpha_i(x)| : i \in \{1, \dots, n\}\big\}. \tag{3.15}$$

(iii) *$N \subseteq \ker \alpha$.*

Proof. "(i) $\Rightarrow$ (ii)": If (i) holds, then there exist $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ such that $\alpha = \sum_{i=1}^n \lambda_i \alpha_i$, implying,

$$\forall_{x \in X} \quad |\alpha(x)| \leq \sum_{i=1}^{n} |\lambda_i \alpha_i(x)| \leq \left(\sum_{i=1}^{n} |\lambda_i|\right) \max\big\{|\alpha_i(x)| : i \in \{1, \dots, n\}\big\},$$

i.e. (3.15) holds with $M := \sum_{i=1}^n |\lambda_i|$.

"(ii) $\Rightarrow$ (iii)" is immediate.

"(iii) $\Rightarrow$ (i)": Define

$$A : X \longrightarrow \mathbb{K}^n, \quad A(x) := \big(\alpha_1(x), \dots, \alpha_n(x)\big).$$

Clearly, A is linear. Moreover, for $x, y \in X$ with $A(x) = A(y)$, we have $\alpha_1(x - y) = \dots = \alpha_n(x - y) = 0$, i.e. $x - y \in N$, implying $\alpha(x) - \alpha(y) = \alpha(x - y) = 0$ by (iii). Thus, $A(x) = A(y)$ implies $\alpha(x) = \alpha(y)$ and

$$\beta : A(X) \longrightarrow \mathbb{K}, \quad \beta\big(A(x)\big) := \alpha(x),$$

well-defines a functional on $A(X)$. Moreover,

$$\begin{aligned} \forall_{\lambda, \mu \in \mathbb{K}} \quad \forall_{x, y \in X} \qquad \beta\big(\lambda A(x) + \mu A(y)\big) &= \beta\big(A(\lambda x + \mu y)\big) = \alpha(\lambda x + \mu y) \\ &= \lambda \alpha(x) + \mu \alpha(y) = \lambda \beta(A(x)) + \mu \beta(A(y)), \end{aligned}$$

showing β to be linear. Thus, β can be extended to a linear functional $\beta : \mathbb{K}^n \longrightarrow \mathbb{K}$. Then there exist $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ such that

$$\forall_{(z_1,\dots,z_n)\in\mathbb{K}^n} \quad \beta(z_1,\dots,z_n) = \sum_{i=1}^n \lambda_i z_i,$$

implying

$$\forall_{x\in X} \quad \alpha(x) = \beta(A(x)) = \sum_{i=1}^n \lambda_i \alpha_i(x),$$

proving (i). ■

Theorem 3.16. *Let X be a vector space over $\mathbb{K}$ with*

$$V \subseteq \{(A : X \longrightarrow \mathbb{K}) : A \text{ linear}\}$$

also being a vector space over $\mathbb{K}$. Let $\mathcal{T}_V$ be the V-topology on X. With each $\alpha \in V$, we associate the seminorm

$$p_\alpha : X \longrightarrow \mathbb{R}_0^+, \quad p_\alpha(x) := |\alpha(x)|. \tag{3.16}$$

(a) *$(X, \mathcal{T}_V)$ is a locally convex topological vector space, where $\mathcal{T}_V$ is the topology induced by the family of seminorms $\mathcal{F} := (p_\alpha)_{\alpha\in V}$ according to Th. 1.40(a).*

(b) *$(X, \mathcal{T}_V)$ is T_1 if, and only if, V separates points on X.*

(c) *$E \subseteq X$ is $\mathcal{T}_V$-bounded if, and only if, each $\alpha \in V$ is bounded on E.*

(d) *The dual of $(X, \mathcal{T}_V)$ is V.*

Proof. (a): Let $\mathcal{T}_0$ be the topology induced by $\mathcal{F}$. If $\mathcal{T}_0 = \mathcal{T}_V$, then $(X, \mathcal{T}_V)$ is a locally convex topological vector space by Th. 1.40(a). It remains to show $\mathcal{T}_0 = \mathcal{T}_V$. Since each p_α is $\mathcal{T}_0$-continuous by Th. 1.40(a)(i), we already know $\mathcal{T}_V \subseteq \mathcal{T}_0$. For the remaining inclusion, it suffices to show $\mathcal{B} \subseteq \mathcal{T}_V$, where $\mathcal{B}$ is the base of $\mathcal{T}_0$ given by Th. 1.40(a). Thus, let $\alpha \in V$, $n \in \mathbb{N}$, $y \in X$. Then

$$\begin{aligned} B_{p_\alpha,n^{-1}}(y) &= \{x \in X : p_\alpha(x-y) < n^{-1}\} = \big\{x \in X : |\alpha(x) - \alpha(y)| < n^{-1}\big\} \\ &= \alpha^{-1}\big(B_{n^{-1}}(\alpha(y))\big) \overset{\text{Lem. 3.11}}{\in} \mathcal{T}_V, \end{aligned}$$

completing the proof of (a).

(b) and (c) are now a direct consequence of (a) and Th. 1.40(a).

(d): Let X'_V denote the dual of $(X, \mathcal{T}_V)$. Each $\alpha \in V$ is $\mathcal{T}_V$-continuous, showing $V \subseteq X'_V$. It remains to show $X'_V \subseteq V$. Let $\alpha \in X'_V$. Then $O := \alpha^{-1}(B_1(0)) \in \mathcal{T}_V$. Thus, according to Th. 1.40(a), there exist $\epsilon \in \mathbb{R}^+$ and $\alpha_1, \dots, \alpha_n \in V$, $n \in \mathbb{N}$, such that

$$U := \bigcap_{i=1}^{n} \{x \in X : |\alpha_i(x)| < \epsilon\} \subseteq O.$$

As in Lem. 3.15, let $N := \bigcap_{i=1}^{n} \ker \alpha_i$. Suppose, there is $x \in N$ such that $s := |\alpha(x)| > 0$. Then $s^{-1}x \in N \subseteq U$, but $|\alpha(s^{-1}x)| = s^{-1}s = 1$, in contradiction to $U \subseteq O$. Thus, $N \subseteq \ker \alpha$, and the equivalence between (iii) and (i) of Lem. 3.15 yields $\alpha \in V$, i.e. $X'_V \subseteq V$. ∎

Corollary 3.17. *Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{K}$, $\mathcal{T}_{\mathrm{w}}$ denoting the corresponding weak topology on X.*

(a) *$(X, \mathcal{T}_{\mathrm{w}})$ is a locally convex topological vector space, where $\mathcal{T}_{\mathrm{w}}$ is the topology induced by the family of seminorms $\mathcal{F} := (p_\alpha)_{\alpha \in X'}$ according to Th.* 1.40(a) *(p_α defined by* (3.16)*).*

(b) *$(X, \mathcal{T}_{\mathrm{w}})$ is T_1 if, and only if, X' separates points on X (($X, \mathcal{T}$) being T_1 and locally convex is sufficient by Cor.* 3.7(a)*). If X' separates points on X, then weak limits are unique.*

(c) *$E \subseteq X$ is weakly bounded if, and only if, each $\alpha \in X'$ is bounded on E.*

(d) *Let $(x_k)_{k \in \mathbb{N}}$ be a sequence in X, $x \in X$. Then*

$$x_k \rightharpoonup x \quad \Leftrightarrow \quad \forall_{\alpha \in X'} \quad \lim_{k \to \infty} \alpha(x_k) = \alpha(x). \tag{3.17}$$

(e) *If $V \subseteq X'$, then $\mathcal{T}_V \subseteq \mathcal{T}_{\mathrm{w}} \subseteq \mathcal{T}$ (i.e. the weak topology is, indeed, weaker than the strong topology; in many, but not all, cases, it is strictly weaker, cf. Cor.* 3.17(g) *and Ex.* 3.18(a),(b) *below). In consequence, strong convergence implies convergence with respect to $\mathcal{T}_V$ (in particular, weak convergence) and, if $(Y, \mathcal{T}_Y)$ is a topological space, then $f : X \longrightarrow Y$ being $\mathcal{T}_V$-$\mathcal{T}_Y$-continuous (e.g. $\mathcal{T}_{\mathrm{w}}$-$\mathcal{T}_Y$-continuous) implies f to be $\mathcal{T}$-$\mathcal{T}_Y$-continuous.*

(f) *The dual of $(X, \mathcal{T}_{\mathrm{w}})$ is still X'.*

(g) *If $\mathcal{T}_{\mathrm{ww}}$ denotes the weak topology corresponding to $(X, \mathcal{T}_{\mathrm{w}})$, then $\mathcal{T}_{\mathrm{w}} = \mathcal{T}_{\mathrm{ww}}$.*

(h) *If $E \subseteq X$, then the strong closure of E is always contained in the weak closure of E, i.e. $\overline{E} \subseteq \mathrm{cl}_{\mathrm{w}}(E)$.*

Proof. (a) – (c) are just Th. 3.16(a)–(c), respectively, applied with $V := X'$ (where, for (b), we once again use that, for topological vector spaces, T_1 implies T_2 and, thus, uniqueness of limits).

(d) is immediate from Prop. 3.12(b).

(e): $\mathcal{T}_V \subseteq \mathcal{T}$ holds since each $\alpha \in V \subseteq X'$ is $\mathcal{T}$-continuous and $\mathcal{T}_V$ is the weakest topology making each $\alpha \in V$ continuous. Then strong convergence implies $\mathcal{T}_V$-convergence, as every $\mathcal{T}_V$-neighborhood is a $\mathcal{T}$-neighborhood. If $f : X \longrightarrow Y$ is $\mathcal{T}_V$-$\mathcal{T}_Y$-continuous and $O \in \mathcal{T}_Y$, then $f^{-1}(O) \in \mathcal{T}_V \subseteq \mathcal{T}$, showing f to be $\mathcal{T}$-$\mathcal{T}_Y$-continuous.

(f) is immediate from Th. 3.16(d), but here the proof is actually easier: Let X'_{w} denote the dual of $(X, \mathcal{T}_{\mathrm{w}})$. Each $\alpha \in X'$ is $\mathcal{T}_{\mathrm{w}}$-continuous, showing $X' \subseteq X'_{\mathrm{w}}$. Conversely, each $\alpha \in X'_{\mathrm{w}}$ is $\mathcal{T}$-continuous by (e), showing $X'_{\mathrm{w}} \subseteq X'$.

(g): Due to Th. 3.16(d), T_{w} and T_{ww} both are the X'-topology on X.

(h): From (e), we know $\mathcal{T}_{\mathrm{w}} \subseteq \mathcal{T}$. Thus, if $E \subseteq X$ is weakly closed, then it is strongly closed. In particular, $\mathrm{cl}_{\mathrm{w}}(E)$ is strongly closed, showing $\overline{E} \subseteq \mathrm{cl}_{\mathrm{w}}(E)$. ■

Example 3.18. (a) If $(X, \mathcal{T})$ is a finite-dimensional T_1 topological vector space over $\mathbb{K}$, $\dim X = n \in \mathbb{N}$, then we know from Th. 1.16(a) that $(X, \mathcal{T})$ is linearly homeomorphic to $\mathbb{K}^n$ with the norm topology. Then every linear functional $\alpha : X \longrightarrow \mathbb{K}$ is continuous, i.e. $X' \cong \mathbb{K}^n$ as well. As $(X, \mathcal{T})$ is a normed space, we also know X' to separate points on X. Thus, by Th. 3.16(b), $(X, \mathcal{T}_{\mathrm{w}})$ is also a T_1 topological vector space. By Cor. 1.17, $\mathcal{T}_{\mathrm{w}}$ must also be the norm topology on X, i.e. $\mathcal{T} = \mathcal{T}_{\mathrm{w}}$.

(b) Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{K}$, $\dim X = \infty$. Let $O \subseteq X$ be a weak neighborhood of 0. According to Th. 3.16(a) and Th. 1.40(a), there exist $\epsilon \in \mathbb{R}^+$ and $\alpha_1, \dots, \alpha_n \in X'$, $n \in \mathbb{N}$, such that

$$U := \bigcap_{i=1}^{n} \{x \in X : |\alpha_i(x)| < \epsilon\} \subseteq O.$$

Since $A : X \longrightarrow \mathbb{K}^n$, $A(x) := (\alpha_1(x), \dots, \alpha_n(x))$, is linear with $N := \ker A = \bigcap_{i=1}^n \ker \alpha_i$, we have $\dim N = \infty$ (otherwise, $\dim X \leq n + \dim N < \infty$). Since $N \subseteq U \subseteq O$, each weak neighborhood of 0 contains an infinite-dimensional subspace N. Thus, if $\mathcal{T}_{\mathrm{w}}$ is T_1 (i.e. if X' separates points on X), then N is not weakly bounded by Rem. 1.26 and, thus, $(X, \mathcal{T}_{\mathrm{w}})$ is not locally bounded. In particular, if $(X, \mathcal{T})$ is a normed space, then it is T_1, locally convex, and locally bounded, whereas $(X, \mathcal{T}_{\mathrm{w}})$ is T_1, locally convex, but not locally bounded, showing $\mathcal{T}_{\mathrm{w}}$ to be strictly weaker than $\mathcal{T}$ on each infinite-dimensional normed space.

(c) Consider the space $(c_0, \|\cdot\|_\infty)$. We will show that for the sequence $(e_k)_{k \in \mathbb{N}}$ in c_0, where, as before, e_k is the kth standard unit vector in $\mathbb{K}^{\mathbb{N}}$, one has $e_k \rightharpoonup 0$ for

$k \to \infty$, but $e_k \not\to 0$ strongly for $k \to \infty$: Since $\|e_k\|_\infty = 1$ for each k, $e_k \not\to 0$ strongly is already clear. To show $e_k \rightharpoonup 0$, we use the representation $(c_0)' \cong l^1$ of Prop. 2.32(b): If $\alpha \in (c_0)'$, then there exists $(a_i)_{i\in\mathbb{N}} \in l^1$ such that, for each $k \in \mathbb{N}$, $\alpha(e_k) = a_k$. Thus, $\lim_{k\to\infty} \alpha(e_k) = \lim_{k\to\infty} a_k = 0$, proving $e_k \rightharpoonup 0$.

Remark 3.19. Weak convergence can often be of use when solving minimization problems: In a first step, it is often easier to show the existence of a sequence that converges *weakly* to a potential solution of the problem.

Theorem 3.20. *Let $(X, \mathcal{T})$ be a locally convex topological vector space over $\mathbb{K}$, $\mathcal{T}_\mathrm{w}$ denoting the corresponding weak topology on X. Let C be a convex subset of X.*

(a) *The weak and the strong closure of C are the same:* $\mathrm{cl}_\mathrm{w}(C) = \overline{C}$.

(b) *C is weakly closed if, and only if, C is strongly closed.*

(c) *Let $A \subseteq X$, $C \subseteq A$. Then C is weakly dense in A if, and only if, C is strongly dense in A.*

Proof. (a): We know that $\overline{C} \subseteq \mathrm{cl}_\mathrm{w}(C)$ always holds according to Cor. 3.17(h). For the remaining inclusion, we use the Hahn-Banach separation Th. 3.5(b): Let $x_0 \in X$, $x_0 \notin \overline{C}$. We apply Th. 3.5(b) with $A := \{x_0\}$ and $B := \overline{C}$ to obtain $\alpha \in X'$ and $s \in \mathbb{R}$ with

$$\underset{x\in\overline{C}}{\forall} \quad \mathrm{Re}\,\alpha(x_0) < s < \mathrm{Re}\,\alpha(x).$$

Thus, $W := \{x \in X : \mathrm{Re}\,\alpha(x) < s\}$ is a weak neighborhood of x_0 such that $W \cap C = \emptyset$, showing $x_0 \notin \mathrm{cl}_\mathrm{w}(C)$, proving (a).

(b): If C is weakly closed, then it is strongly closed, as $\mathcal{T}_\mathrm{w} \subseteq \mathcal{T}$. If C is strongly closed, then, by (a), $C = \mathrm{cl}_\mathrm{w}(C) = \overline{C}$, showing C to be weakly closed.

(c) holds, as

$$C \text{ is weakly dense in } A \quad \Leftrightarrow \quad \mathrm{cl}_\mathrm{w}(C) = A \quad \overset{\text{(a)}}{\Leftrightarrow} \quad \overline{C} = A \quad C \text{ is strongly dense in } A,$$

completing the proof of the theorem. ∎

Remark and Definition 3.21. Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{K}$ with dual space X'. The following construction yields a topology on X' that turns out to be quite useful: First, define

$$\underset{x\in X}{\forall} \quad f_x : X' \longrightarrow \mathbb{K}, \quad f_x(\alpha) := \alpha(x). \tag{3.18}$$

Then each f_x is linear, since

$$\underset{\lambda,\mu\in\mathbb{K}}{\forall} \quad \underset{\alpha,\beta\in X'}{\forall} \quad f_x(\lambda\alpha+\mu\beta) = (\lambda\alpha+\mu\beta)(x) = \lambda\alpha(x)+\mu\beta(x) = \lambda f_x(\alpha)+\mu f_x(\beta).$$

The map

$$\Phi : X \longrightarrow \{(f : X' \longrightarrow \mathbb{K}) : f \text{ linear}\}, \quad \Phi(x) := f_x, \tag{3.19}$$

is also linear, since

$$\underset{\lambda,\mu\in\mathbb{K}}{\forall} \quad \underset{x,y\in X}{\forall} \quad \underset{\alpha\in X'}{\forall} \quad \begin{aligned} \Phi(\lambda x+\mu y)(\alpha) &= f_{\lambda x+\mu y}(\alpha) = \alpha(\lambda x+\mu y) = \lambda\alpha(x)+\mu\alpha(y) \\ &= \lambda f_x(\alpha)+\mu f_y(\alpha) = \big(\lambda\Phi(x)+\mu\Phi(y)\big)(\alpha). \end{aligned}$$

We call the $\Phi(X)$-topology on X' in the sense of Def. 3.10, the *weak star* (write: *weak**) topology on X', denoted by $\mathcal{T}_{w*}$. We will use terms such as *weak*-convergent* (and write $\alpha_n \overset{*}{\rightharpoonup} \alpha$ to denote such weak*-convergence), *weak*-closed*, *weak*-compact*, etc. The usefulness of the weak*-topology is mainly due to compactness results for the following Sec. 3.3.

Remark 3.22. Let $(X,\mathcal{T})$ be a topological vector space over $\mathbb{K}$ with dual space X', where the map Φ is defined as in (3.19), and $\mathcal{T}_{w*}$ denotes the weak*-topology on X'.

(a) $\Phi(X)$ separates points on X': Indeed, proceeding by contraposition, if $\alpha,\beta\in X'$ are such that

$$\underset{x\in X}{\forall} \quad \alpha(x) = \Phi(x)(\alpha) = \Phi(x)(\beta) = \beta(x),$$

then $\alpha=\beta$.

(b) According to (a) and Th. 3.16(a),(b), $(X',\mathcal{T}_{w*})$ is a T_1 locally convex topological vector space over $\mathbb{K}$, where $\mathcal{T}_{w*}$ is the topology induced by the family of seminorms $\mathcal{F} := (p_x)_{x\in X}$, where

$$\underset{x\in X}{\forall} \quad p_x : X' \longrightarrow \mathbb{R}_0^+, \quad p_x(\alpha) := |\alpha(x)|.$$

According to Th. 3.16(d), the dual of $(X',\mathcal{T}_{w*})$ is $\Phi(X)$.

(c) In consequence of (b), weak*-limits are unique. Let $(\alpha_k)_{k\in\mathbb{N}}$ be a sequence in X', $\alpha\in X'$. Then, due to Prop. 3.12(b),

$$\alpha_k \overset{*}{\rightharpoonup} \alpha \quad \Leftrightarrow \quad \underset{x\in X}{\forall} \quad \lim_{k\to\infty} \alpha_k(x) = \alpha(x). \tag{3.20}$$

(d) In general, one can not expect Φ to be injective: For example, we know from Ex. 1.11(b), that $X := L^p([0,1],\mathcal{L}^1,\lambda^1)$ with $0<p<1$ has $X'=\{0\}$, implying

$\Phi(X) = \{0\}$ and Φ is not injective. However, Φ *is* injective if X' separates points on X (by Cor. 3.7(a), $(X, \mathcal{T})$ being T_1 and locally convex is sufficient): If $x, y \in X$ with $x \neq y$, then let $\alpha \in X'$ such that $\alpha(x) \neq \alpha(y)$. Then

$$\Phi(x)(\alpha) = f_x(\alpha) = \alpha(x) \neq \alpha(y) = f_y(\alpha) = \Phi(y)(\alpha),$$

showing $\Phi(x) \neq \Phi(y)$. In cases, where Φ is injective, one often identifies $\Phi(X)$ with X.

(e) In general, one can not expect Φ to be surjective: Let W be a vector space over $\mathbb{K}$. The set

$$W'_{\text{lin}} := \{(A : W \longrightarrow \mathbb{K}) : A \text{ linear}\}$$

is known as the *linear dual* of W. It is a general result of Linear Algebra that always $\dim W \leq \dim W'_{\text{lin}}$, and $\dim W = \dim W'_{\text{lin}}$ if, and only if, the dimension of W is finite[2]. If X is a vector space over $\mathbb{K}$, then, using $V := X'_{\text{lin}}$ in Th. 3.16, we know from Th. 3.16(a),(d) that $(X, \mathcal{T}_{X'_{\text{lin}}})$ is a topological vector space with dual $X' = X'_{\text{lin}}$. Thus, in this case, if X is infinite-dimensional, then $\dim X < \dim X' < \dim(X')'_{\text{lin}}$ and $\Phi : X \longrightarrow (X')'_{\text{lin}}$ can not be surjective.

3.3 Banach-Alaoglu

The proof of the Banach-Alaoglu Th. 3.26 is based on Tychonoff Th. 3.25. A proof of Th. 3.25 was already provided in [Phi16b, Sec. E.3], using nets. Here we provide a different proof, based on the Alexander subbase Th. 3.24, that does not require the use of nets. We start by showing that Zorn's lemma implies *Hausdorff's Maximality Principle* (both are actually equivalent, see [Phi16a, Th. A.52]):

Theorem 3.23 (Hausdorff's Maximality Principle). *Every nonempty partially ordered set $(X, \leq)$ contains a maximal chain, where we recall that a chain is a totally ordered subset.*

Proof. To apply Zorn's lemma, let

$$\mathcal{P} := \{C \subseteq X : C \text{ is a chain}\}$$

and note that $\mathcal{P}$ is partially ordered by set inclusion $\subseteq$. Now every chain $\mathcal{C}$ in $\mathcal{P}$ has an upper bound, namely $W_{\mathcal{C}} := \bigcup_{C \in \mathcal{C}} C$ ($W_{\mathcal{C}} \in \mathcal{P}$ follows from $\mathcal{C}$ being a chain in $\mathcal{P}$). Thus, Zorn's lemma yields a maximal element of $\mathcal{P}$, i.e. a maximal chain in X as desired. ■

[2] If the dimension of W is infinite, then $\dim W < \dim W'_{\text{lin}}$ in the sense that for each basis B of W and each basis B' of W'_{lin}, one has $\#B < \#B'$ (i.e. there does not exist a surjective map $f : B \longrightarrow B'$).

Theorem 3.24 (Alexander Subbase Theorem)**.** *Let $(X, \mathcal{T})$ be a topological space and assume $\mathcal{S}$ to be a subbase for $\mathcal{T}$. Let $C \subseteq X$. Then C is compact if, and only if, every open cover of C with sets from $\mathcal{S}$ has a finite subcover, i.e. if $(O_i)_{i \in I}$ is a family in $\mathcal{S}$ with $C \subseteq \bigcup_{i \in I} O_i$, then there exist $i_1, \dots, i_N \in I$, $N \in \mathbb{N}$, such that $C \subseteq \bigcup_{k=1}^{N} O_{i_k}$.*

Proof. Since $\mathcal{S} \subseteq \mathcal{T}$, one only has to show that the subbase condition implies compactness. We proceed by contraposition and assume C is not compact. Using Hausdorff's Maximality Principle, we will produce a cover of C with sets from $\mathcal{S}$ that does not have a finite subcover: Let

$$\mathcal{P} := \{\mathcal{M} \subseteq \mathcal{T} : \mathcal{M} \text{ is cover of } C \text{ without finite subcover}\}$$

with the partial order given by set inclusion (note $X \notin \mathcal{M}$ for each $\mathcal{M} \in \mathcal{P}$, as $\{X\}$ would always constitute a finite subcover). As C is not compact, $\mathcal{P} \neq \emptyset$ and, by Hausdorff's Maximality Principle of Th. 3.23, we let $\Omega \subseteq \mathcal{P}$ be a maximal chain and set

$$\mathcal{M}_\Omega := \bigcup_{\mathcal{M} \in \Omega} \mathcal{M}.$$

Then, clearly, $\mathcal{M}_\Omega = \max \Omega$: $\mathcal{M}_\Omega \in \mathcal{P}$, i.e. $\mathcal{M}_\Omega$ is a cover of C without a finite subcover, but if we add any new $O \in \mathcal{T} \setminus \mathcal{M}_\Omega$, then $\mathcal{M}_\Omega \,\dot\cup \{O\}$ does have a finite subcover. Next, let

$$\mathcal{M}_\mathcal{S} := \mathcal{M}_\Omega \cap \mathcal{S}.$$

By definition, $\mathcal{M}_\mathcal{S} \subseteq \mathcal{M}_\Omega$, i.e. no finite subset of $\mathcal{M}_\mathcal{S}$ can cover C. So far, we have not excluded $\mathcal{M}_\mathcal{S} = \emptyset$, but, in the next step, we will even show that $\mathcal{M}_\mathcal{S}$ must cover C: Seeking a contradiction, assume there is $x \in C$, x not covered by $\mathcal{M}_\mathcal{S}$. Since $\mathcal{M}_\Omega$ covers C, there exists $O \in \mathcal{M}_\Omega$ such that $x \in O$. As O is open and $\mathcal{S}$ is a subbase, there exist $S_1, \dots, S_n \in \mathcal{S}$, $n \in \mathbb{N}$, such that

$$x \in \bigcap_{i=1}^{n} S_i \subseteq O.$$

As x is not covered by $\mathcal{M}_\mathcal{S}$, $S_i \notin \mathcal{M}_\mathcal{S}$, implying $S_i \notin \mathcal{M}_\Omega$ for each $i \in \{1, \dots, n\}$. Thus, each $\mathcal{M}_\Omega \,\dot\cup \{S_i\}$ must have a finite subcover of C. In other words, for each $i \in \{1, \dots, n\}$, there is an open set U_i, being a finite union of sets in $\mathcal{M}_\Omega$, such that $C \subseteq S_i \cup U_i$, implying

$$C \subseteq U_1 \cup \dots \cup U_n \cup \bigcap_{i=1}^{n} S_i \subseteq U_1 \cup \dots \cup U_n \cup O,$$

i.e. $\mathcal{M}_\Omega$ has a finite subcover, providing the desired contradiction. Thus, we have shown $\mathcal{M}_\mathcal{S}$ to be a cover of C with sets from $\mathcal{S}$, not having a finite subcover, thereby proving the theorem. ■

Theorem 3.25 (Tychonoff)**.** *Let $(X_i, \mathcal{T}_i)$ be topological spaces, $i \in I$. If $X = \prod_{i\in I} X_i$ is endowed with the product topology $\mathcal{T}$ and each X_i is compact, then X is compact.*

Proof. For each $i \in I$, we have the projection $\pi_i : X \longrightarrow X_i$ and we define

$$\mathcal{S}_i := \left\{ \pi_i^{-1}(O) : O \in \mathcal{T}_i \right\}.$$

We know from [Phi16b, Ex. 1.53(a)], that

$$\mathcal{S} = \left\{ \pi_i^{-1}(O_i) : i \in I,\, O_i \in \mathcal{T}_i \right\} = \bigcup_{i\in I} \mathcal{S}_i$$

constitutes a subbase of $\mathcal{T}$. Let $\mathcal{M} \subseteq \mathcal{S}$ be a cover of X. By Th. 3.24, it suffices to show $\mathcal{M}$ has a finite subcover. Define

$$\forall_{i\in I} \quad \mathcal{M}_i := \mathcal{M} \cap \mathcal{S}_i.$$

Seeking a contradiction, assume no $\mathcal{M}_i$ covers X. Then, for each $i \in I$, there is $x^i \in X$ not covered by $\mathcal{M}_i$. Let $x_i^i := \pi_i(x^i) \in X_i$. Then *no* $x \in \pi_i^{-1}\{x_i^i\}$ can be covered by $\mathcal{M}_i$, since if there exists $O \in \mathcal{T}_i$ such that $x \in U := \pi_i^{-1}(O) \in \mathcal{M}_i$ with $x \in \pi_i^{-1}\{x_i^i\}$, then $x^i \in U$ as well. Now let $z := (x_i^i)_{i\in I} \in X$. Then z is not covered by $\mathcal{M}$, the desired contradiction. Thus, we have shown there exists some $i_0 \in I$ such that $\mathcal{M}_{i_0}$ covers X. There exists $\mathcal{C} \subseteq \mathcal{T}_{i_0}$ such that $\mathcal{M}_{i_0} = \{\pi_{i_0}^{-1}(O) : O \in \mathcal{C}\}$, i.e. $\mathcal{C}$ must be a cover of X_{i_0}. As X_{i_0} is compact, there must be a finite subcover $O_1, \dots, O_N \in \mathcal{C}$, $N \in \mathbb{N}$. Then

$$\pi_{i_0}^{-1}(O_1), \dots, \pi_{i_0}^{-1}(O_N) \in \mathcal{M}_{i_0} \subseteq \mathcal{M}$$

cover X and form a finite subcover of $\mathcal{M}$, completing the proof. ■

Theorem 3.26 (Banach-Alaoglu)**.** *Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{K}$. If $U \in \mathcal{U}(0)$, then*

$$K := K(U) := \left\{ \alpha \in X' : \forall_{x\in U} \ |\alpha(x)| \leq 1 \right\} \tag{3.21}$$

is weak-compact.*

Proof. Since $U \in \mathcal{U}(0)$, we know U to be absorbing. Thus,

$$\forall_{x\in X} \quad \exists_{s(x)\in\mathbb{R}^+} \quad x \in s(x)\, U,$$

implying

$$\forall_{x\in X} \quad \forall_{\alpha\in K} \quad |\alpha(x)| \leq s(x).$$

Now define

$$P := \prod_{x\in X} \overline{B}_{s(x)}(0) \subseteq \mathbb{K}^X,$$

letting $\mathcal{P}$ denote the product topology on P. As each $\overline{B}_{s(x)}(0) \subseteq \mathbb{K}$ is compact, $(P,\mathcal{P})$ is compact by Tychonoff's Th. 3.25. By definition, P consists of all functions $f : X \longrightarrow \mathbb{K}$ (not necessarily linear) such that $|f(x)| \leq s(x)$ for each $x \in X$. In particular, we have $K \subseteq X' \cap P$, i.e. we obtain two relative topologies on K, namely $\mathcal{P}_K$ and $(\mathcal{T}_{w*})_K$. We will show:

(1) The relative topologies $\mathcal{P}_K$ and $(\mathcal{T}_{w*})_K$ are identical, i.e. $\mathcal{P}_K = (\mathcal{T}_{w*})_K$.

(2) K is $\mathcal{P}$-closed.

Then K is a closed subset of the compact set P, showing K to be $\mathcal{P}$-compact by [Phi16b, Prop. 3.14(a)]. Then K is $\mathcal{T}_{w*}$-compact by (1), proving the theorem. It remains to prove (1), (2).

(1): We show $\mathcal{P}_K$ and $(\mathcal{T}_{w*})_K$ to have identical local bases at each $\alpha \in K$. Let $\alpha \in K$. A local base for $\mathcal{P}$ at α is given by

$$\mathcal{B}_{\mathcal{P}} := \left\{ P \cap \bigcap_{x\in J} \pi_x^{-1}\big(B_\epsilon(\alpha(x))\big) :\ J \subseteq X,\ 0 < \#J < \infty,\ \epsilon \in \mathbb{R}^+ \right\}.$$

A local base for $\mathcal{T}_{w*}$ at α is given by

$$\mathcal{B}_{w*} := \left\{ X' \cap \bigcap_{x\in J} \pi_x^{-1}\big(B_\epsilon(\alpha(x))\big) :\ J \subseteq X,\ 0 < \#J < \infty,\ \epsilon \in \mathbb{R}^+ \right\}.$$

A local base for $\mathcal{P}_K$ at α is $\mathcal{K}_{\mathcal{P}} := \{B \cap K :\ B \in \mathcal{B}_{\mathcal{P}}\}$, and a local base for $(\mathcal{T}_{w*})_K$ at α is $\mathcal{K}_{w*} := \{B \cap K :\ B \in \mathcal{B}_{w*}\}$. Thus, since $K \subseteq X' \cap P$, we see that $\mathcal{K}_{\mathcal{P}} = \mathcal{K}_{w*}$, proving (1).

(2): Let $f_0 \in P$ be in the $\mathcal{P}$-closure of K. We wish to show f_0 is linear. To this end, let $x, y \in X$ and $\lambda, \mu \in \mathbb{K}$ and $\epsilon \in \mathbb{R}^+$. Then

$$\begin{aligned} O := \{f \in P :\ |f(x) - f_0(x)| < \epsilon\} \cap \{f \in P :\ |f(y) - f_0(y)| < \epsilon\} \\ \cap \{f \in P :\ |f(\lambda x + \mu y) - f_0(\lambda x + \mu y)| < \epsilon\} \end{aligned}$$

is an open $\mathcal{P}$-neighborhood of f_0, implying $O \cap K \neq \emptyset$. Let $f \in O \cap K$. As f is linear, we obtain

$$\begin{aligned} &\big|f_0(\lambda x + \mu y) - \lambda f_0(x) - \mu f_0(y)\big| \\ &\leq \big|f_0(\lambda x + \mu y) - f(\lambda x + \mu y)\big| + \big|\lambda f(x) + \mu f(y) - \lambda f_0(x) - \mu f_0(y)\big| \\ &< \epsilon + |\lambda|\,\epsilon + |\mu|\,\epsilon, \end{aligned}$$

showing f_0 to be linear, since $\epsilon > 0$ was arbitrary. Now, for $x \in U$, $\epsilon > 0$ and f as above, we have

$$|f_0(x)| \leq |f_0(x) - f(x)| + |f(x)| < \epsilon + 1,$$

since $|f(x)| \leq 1$ for $x \in U$ and $f \in K$. Thus, $|f_0(x)| \leq 1$, showing $f_0 \in K$, proving (2) and the theorem. ■

In preparation for the proof of Th. 3.28 below, we provide the following proposition:

Proposition 3.27. **(a)** *Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be topologies on a set X. If $\mathcal{T}_1 \subseteq \mathcal{T}_2$, $\mathcal{T}_1$ is Hausdorff (i.e. T_2), and $\mathcal{T}_2$ is compact, then $\mathcal{T}_1 = \mathcal{T}_2$ (in particular, if $(X, \mathcal{T})$ is a compact T_2 space, then each strictly weaker topology on X is not T_2 and each strictly stronger topology on X is not compact).*

(b) *Let $(X, \mathcal{T})$ be a compact topological space. If there exists a sequence $\mathcal{F} := (f_n)_{n\in\mathbb{N}}$ of continuous functions $f_n : X \longrightarrow \mathbb{R}$ such that $\mathcal{F}$ separates points on X, then $(X, \mathcal{T})$ is metrizable.*

Proof. Exercise. ■

Theorem 3.28. *If $(X, \mathcal{T})$ is a separable topological vector space over $\mathbb{K}$ and $K \subseteq X'$ is weak*-compact, then K is metrizable in its weak*-topology.*

Proof. Let the sequence $(x_n)_{n\in\mathbb{N}}$ in X be dense. Then each

$$f_n : X' \longrightarrow \mathbb{K}, \quad f_n(\alpha) := \alpha(x_n),$$

is weak*-continuous by the definition of the weak*-topology. Let $\alpha, \beta \in X'$. If

$$\underset{n\in\mathbb{N}}{\forall} \quad f_n(\alpha) = f_n(\beta),$$

then the continuous maps α, β agree on the dense set $\{x_n : n \in \mathbb{N}\}$, implying $\alpha = \beta$, showing $(f_n)_{n\in\mathbb{N}}$ to separate points on X' and, in particular, on K. An application of Prop. 3.27(b) proves K to be metrizable in its weak*-topology. ■

Caveat 3.29. Theorem 3.28 does *not* claim that the dual of a separable topological vector space is itself metrizable in its weak*-topology. Indeed, in many cases, it is not: For example, if X' separates points on X, then it is an exercise to show $(X', \mathcal{T}_{w*})$ is metrizable if, and only if, the dimension of X is finite or countable.

Corollary 3.30. *If $(X, \mathcal{T})$ is a separable topological vector space over $\mathbb{K}$, $U \in \mathcal{U}(0)$ and K as in Th. 3.26, i.e.*

$$K := K(U) := \left\{ \alpha \in X' : \underset{x\in U}{\forall} \ |\alpha(x)| \leq 1 \right\},$$

then K is weak-sequentially compact, i.e. for each sequence $(\alpha_n)_{n\in\mathbb{N}}$ in K, there exists a subsequence $(\alpha_{n_k})_{k\in\mathbb{N}}$ and $\alpha \in X'$ such that*

$$\underset{x\in X}{\forall} \quad \lim_{k\to\infty} \alpha_{n_k}(x) = \alpha(x).$$

Proof. One merely combines Th. 3.26 with Th. 3.28 and recalls that, by [Phi16b, Th. 3.20], compactness and sequential compactness are the same in metric spaces. ■

Remark 3.31. For normed spaces $(X, \|\cdot\|)$, we can express the Banach-Alaoglu theorem by stating that balls in X' are weak*-compact (weak*-sequentially compact if X is separable): Since translations and non-zero scalings are homeomorphisms, it suffices to consider the unit ball

$$B' := \{\alpha \in X' : \|\alpha\| \leq 1\},$$

where $\|\alpha\| := \sup\{|\alpha(x)| : x \in X, \|x\| \leq 1\}$ is the operator norm (cf. Cor. 2.15(c) and Sec. 4.1 below). If $U := \{x \in X : \|x\| \leq 1\}$, then $B' = K(U)$, which is weak*-compact by the Banach-Alaoglu Th. 3.26 (weak*-sequentially compact for X separable by Cor. 3.30).

—

In preparation for the proof of Th. 3.33 below, we provide the following proposition, which constitutes another variant of the uniform boundedness principle:

Proposition 3.32. *Let $(X, \mathcal{T})$ and $(Y, \mathcal{S})$ be topological vector spaces over $\mathbb{K}$, assume $(X, \mathcal{T})$ to be T_1 and $K \subseteq X$ to be compact and convex. If $\mathcal{F} \subseteq \mathcal{L}(X, Y)$ is a collection of continuous linear maps such that*

$$\underset{x\in K}{\forall} \quad M_x := \{A(x) : A \in \mathcal{F}\} \text{ bounded in } Y, \tag{3.22}$$

then $\mathcal{F}$ is uniformly bounded in the sense that

$$\underset{B\subseteq Y \text{ bounded}}{\exists} \quad \underset{A\in\mathcal{F}}{\forall} \quad A(K) \subseteq B. \tag{3.23}$$

Proof. Exercise. ■

Theorem 3.33. *Let $(X, \mathcal{T})$ be a locally convex topological vector space over $\mathbb{K}$ and $E \subseteq X$. Then E is weakly bounded if, and only if, E is strongly bounded.*

Proof. Since $\mathcal{T}_w \subseteq \mathcal{T}$, every weak neighborhood of 0 is a strong neighborhood of 0, i.e. it always holds that a strongly bounded set is also weakly bounded. It remains to prove the converse: Assume E to be weakly bounded and let $U \subseteq X$ be a strong neighborhood of 0. We have to show there is $s \in \mathbb{R}^+$ such that $E \subseteq sU$. As $(X, \mathcal{T})$ is locally convex, we can choose a strong neighborhood V of 0 such that V is convex, balanced, and closed, and such that $V \subseteq U$. As in the Banach-Alaoglu Th. 3.26, let

$$K := K(V) := \left\{ \alpha \in X' : \underset{x \in V}{\forall} |\alpha(x)| \leq 1 \right\}.$$

We show that in the present situation

$$V = \tilde{V} := \left\{ x \in X : \underset{\alpha \in K}{\forall} |\alpha(x)| \leq 1 \right\} : \tag{3.24}$$

We have $V \subseteq \tilde{V}$ directly from the definition of K. Now suppose $x_0 \in X \setminus V$. Applying Cor. 3.7(d) (with $B := V$), we obtain

$$\underset{\alpha \in X'}{\exists} \quad \Big(|\alpha{\restriction}_V| \leq 1 \ \wedge \ \alpha(x_0) \in]1, \infty[\Big),$$

i.e. $x_0 \notin \tilde{V}$, proving $\tilde{V} \subseteq V$ and (3.24). In the next step, we intend to apply Prop. 3.32, using $(X', \mathcal{T}_{w*})$ for $(X, \mathcal{T})$, $\mathbb{K}$ for Y, K for K, and $\mathcal{F} := \Phi(E)$, where Φ is as in (3.19), i.e.

$$\underset{x \in X}{\forall} \quad \Phi(x) := f_x, \quad f_x : X' \longrightarrow \mathbb{K}, \quad f_x(\alpha) := \alpha(x).$$

Note that $(X', \mathcal{T}_{w*})$ is a T_1 topological vector space by Rem. 3.22(b), K is, clearly, convex, and K is $\mathcal{T}_{w*}$-compact by the Banach-Alaoglu Th. 3.26. Moreover, each f_x is $\mathcal{T}_{w*}$-continuous by the very definition of $\mathcal{T}_{w*}$. Since we assume E to be weakly bounded, by Cor. 3.17(c), each $\alpha \in X'$ is bounded on E, i.e.

$$\underset{\alpha \in X'}{\forall} \quad \underset{\mu(\alpha) \in \mathbb{R}^+}{\exists} \quad \underset{x \in E}{\forall} \quad |f_x(\alpha)| = |\alpha(x)| \leq \mu(\alpha),$$

implying

$$\underset{\alpha \in K}{\forall} \quad M_\alpha := \{\alpha(x) : x \in E\} \text{ bounded in } \mathbb{K}.$$

Thus, we have verified all hypotheses of Prop. 3.32 and conclude

$$\underset{M \in \mathbb{R}^+}{\exists} \quad \underset{x \in E}{\forall} \quad \underset{\alpha \in K}{\forall} \quad |\alpha(x)| \leq M,$$

which, together with (3.24), yields

$$\underset{x \in E}{\forall} \quad M^{-1}x \in V \subseteq U.$$

In other words,

$$E \subseteq M\,V \subseteq M\,U,$$

showing E to be strongly bounded. ∎

Corollary 3.34. *Let $(X, \|\cdot\|)$ be a normed vector space over $\mathbb{K}$, $E \subseteq X$. If*

$$\forall_{\alpha \in X'} \quad \sup\{|\alpha(x)| : x \in E\} < \infty, \tag{3.25}$$

then

$$\exists_{M \in \mathbb{R}^+} \quad \sup\{\|x\| : x \in E\} \leq M. \tag{3.26}$$

Proof. As $(X, \|\cdot\|)$ is locally convex, we can apply Th. 3.33. As (3.25) means that E is weakly bounded, E must be strongly bounded, which is (3.26). ■

3.4 Extreme Points, Krein-Milman

Definition 3.35. Let X be a vector space over $\mathbb{K}$.

(a) Let $C \subseteq X$ be convex. Then $E \subseteq C$ is called an *extreme set* of C if, and only if, $p = s\,x + (1-s)\,y \in E$ with $x, y \in C$ and $0 < s < 1$ implies $x, y \in E$ (i.e. extreme sets of C are precisely those subsets E of C that do not contain interior points of line segments in C with endpoints outside of E). Moreover, $p \in C$ is called an *extreme point* of C if, and only if, $\{p\}$ is an extreme set of C, i.e. if, and only if, $p = s\,x + (1-s)\,y$ with $x, y \in C$ and $0 < s < 1$ implies $p = x = y$. The set of all extreme points of C is denoted by $\mathrm{ex}(C)$.

(b) Let $A \subseteq X$ be arbitrary. The *convex hull* of A, denoted $\mathrm{conv}(A)$, is the intersection of all convex subsets of X containing A (by Prop. 1.8(a), $\mathrm{conv}(A)$ is convex, i.e. it is the smallest convex set containing A).

(c) Let $A \subseteq X$ be arbitrary and let $\mathcal{T}$ be a topology on X. The *closed convex hull* of A, denoted $\overline{\mathrm{conv}(A)}$, is the $\mathcal{T}$-closure of $\mathrm{conv}(A)$.

Lemma 3.36. *Let X be a vector space over $\mathbb{K}$, and let $C \subseteq X$ be convex.*

(a) *If $(E_i)_{i \in I}$, $I \neq \emptyset$, is a family of extreme subsets of C, then $E := \bigcap_{i \in I} E_i$ is an extreme set of C as well.*

(b) *For $p \in C$, the following statements are equivalent:*

(i) $p \in \mathrm{ex}(C)$.

(ii) $p = \frac{1}{2}(x + y)$ *with* $x, y \in C$ *implies* $p = x = y$.

(iii) $p \pm x \in C$ *with* $x \in X$ *implies* $x = 0$.

Proof. Exercise. ■

Example 3.37. (a) Clearly, for real intervals with $a, b \in \mathbb{R}$, $a < b$, one has $\operatorname{ex}([a,b]) = \{a, b\}$ and $\operatorname{ex}(]a,b[) = \emptyset$.

(b) Let $(X, \|\cdot\|)$ be a normed vector space over $\mathbb{K}$, $\dim X > 0$, $C := \overline{B}_1(0)$. Then $\operatorname{ex}(C) \subseteq S_1(0) = \{x \in X : \|x\| = 1\}$: If $x = 0$, and $\|y\| = 1$, then $x \pm y \in C$, showing $x \notin \operatorname{ex}(C)$. If $x \in X$ with $0 < \|x\| < 1$, then

$$x = \|x\| \frac{x}{\|x\|} + \big(1 - \|x\|\big)\, 0,$$

showing $x \notin \operatorname{ex}(C)$.

(c) Let $\big(X, \langle\cdot,\cdot\rangle\big)$ be an inner product space over $\mathbb{K}$ with induced norm $\|x\| := \sqrt{\langle x, x\rangle}$ (cf. [Phi16b, Def. 1.66]). Then $\operatorname{ex}(\overline{B}_1(0)) = S_1(0)$: According to (b), we only have to show $S_1(0) \subseteq \operatorname{ex}(\overline{B}_1(0))$. Thus, let $p \in S_1(0)$ and $x \in X$ with $p \pm x \in \overline{B}_1(0)$: Then

$$1 \geq \|p \pm x\|^2 = \|p\|^2 \pm 2\operatorname{Re}\langle p, x\rangle + \|x\|^2,$$

i.e. $\|x\|^2 \leq \mp 2\operatorname{Re}\langle p, x\rangle$, i.e. $x = 0$, showing $p \in \operatorname{ex}(\overline{B}_1(0))$.

(d) Consider $(c_0, \|\cdot\|_\infty)$. Then $\operatorname{ex}(\overline{B}_1(0)) = \emptyset$: Let $x = (x_k)_{k\in\mathbb{N}} \in S_1(0)$. According to (b), it suffices to show $x \notin \operatorname{ex}(\overline{B}_1(0))$. Since $\lim_{k\to\infty} x_k = 0$, there exists $k_0 \in \mathbb{N}$ such that $|x_{k_0}| < \frac{1}{2}$. Then $x \pm \frac{1}{2} e_{k_0} \in \overline{B}_1(0)$, where $e_{k_0} \in c_0$ is the standard unit vector corresponding to k_0. Thus, $x \notin \operatorname{ex}(\overline{B}_1(0))$.

(e) Consider $X := L^1([0,1], \mathcal{L}^1, \lambda^1)$. Then $\operatorname{ex}(\overline{B}_1(0)) = \emptyset$: Let $f \in S_1(0)$. According to (b), it suffices to show $f \notin \operatorname{ex}(\overline{B}_1(0))$. Since $\|f\|_1 = \int_0^1 |f| \,\mathrm{d}\lambda^1 = 1$ and

$$F : [0,1] \longrightarrow \mathbb{R}, \quad F(s) := \int_0^s |f| \,\mathrm{d}\lambda^1,$$

is continuous, with $F(0) = 0$ and $F(1) = 1$, there exists $s_0 \in]0,1[$ such that $F(s_0) = \frac{1}{2}$. Letting $f_1 := 2\, f\, \chi_{[0,s_0]}$, $f_2 := 2\, f\, \chi_{[s_0,1]}$, we have $\|f_1\|_1 = \|f_2\|_1 = 1$, $f_1 \neq f$, $f_2 \neq f$, but $f = \frac{1}{2}(f_1 + f_2)$, showing $f \notin \operatorname{ex}(\overline{B}_1(0))$.

—

The following variant of the Hahn-Banach separation Th. 3.5(b) will be used in the proof of the Krein-Milman Th. 3.39 below:

Proposition 3.38. *Let $(X, \mathcal{T})$ be a topological vector space over $\mathbb{K}$ with the property that X' separates points on X. Assume $A, B \subseteq X$ such that A, B are nonempty, convex, and compact with $A \cap B = \emptyset$. Then there exists $\alpha \in X'$, satisfying*

$$\sup(\operatorname{Re}\alpha)(A) < \inf(\operatorname{Re}\alpha)(B). \tag{3.27}$$

Proof. Let $\mathcal{T}_{\mathrm{w}}$ denote the weak topology on X. Since A, B are $\mathcal{T}$-compact and $\mathcal{T}_{\mathrm{w}} \subseteq \mathcal{T}$, A, B are also $\mathcal{T}_{\mathrm{w}}$-compact. By assumption, X' separates points on X, i.e. $(X, \mathcal{T}_{\mathrm{w}})$ is T_2 as a consequence of Cor. 3.17(b). Thus, A, B are $\mathcal{T}_{\mathrm{w}}$-closed as well. As $(X, \mathcal{T}_{\mathrm{w}})$ is locally convex, we can apply Th. 3.5(b) to obtain $\alpha \in X'_{\mathrm{w}}$, satisfying (3.27), where X'_{w} denotes the dual of $(X, \mathcal{T}_{\mathrm{w}})$. Since $X' = X'_{\mathrm{w}}$ by Cor. 3.17(f), the proof is complete. ∎

Theorem 3.39 (Krein-Milman)**.** *Let $(X, \mathcal{T})$ be a T_1 topological vector space over $\mathbb{K}$ with the property that X' separates points on X. If $K \subseteq X$ is compact and convex, then K is the closed convex hull of its extreme points, i.e.*

$$K = \overline{\operatorname{conv}(\operatorname{ex}(K))}. \tag{3.28}$$

Proof. The main ingredients are Prop. 3.38 and another application of Hausdorff's Maximality Principle of Th. 3.23. If $K = \emptyset$, then there is nothing to prove. Thus, assume $K \neq \emptyset$. Define

$$\mathcal{P} := \{E \subseteq K : E \text{ nonempty, compact, extreme set of } K\}$$

and note $K \in \mathcal{P}$. We will use the following two properties of $\mathcal{P}$:

(1) For each $\emptyset \neq \mathcal{M} \subseteq \mathcal{P}$, we have $\bigcap_{E \in \mathcal{M}} E = \emptyset$ or $\bigcap_{E \in \mathcal{M}} E \in \mathcal{P}$.

(2) If $E \in \mathcal{P}$, $\alpha \in X'$, and $\mu := \max\{\operatorname{Re}\alpha(x) : x \in E\}$, then

$$E_\alpha := \{x \in E : \operatorname{Re}\alpha(x) = \mu\} \in \mathcal{P}.$$

(1): Let $C := \bigcap_{E \in \mathcal{M}} E$. Since $(X, \mathcal{T})$ is T_2, C is compact by [Phi16b, Prop. 3.17(b)]; C is an extreme set of K by Lem. 3.36.

(2): Let $E \in \mathcal{P}$ and $\alpha \in X'$. Then $\mu = \max\{\operatorname{Re}\alpha(x) : x \in E\}$ is well-defined, since $\emptyset \neq E$ is compact and $\operatorname{Re}\alpha$ is continuous. As a closed subset of the compact set E, E_α is compact. Suppose $p = s\,x + (1-s)\,y \in E_\alpha$ with $x, y \in K$ and $0 < s < 1$. Then $p \in E$ implies $x, y \in E$, i.e. $\operatorname{Re}\alpha(x) \leq \mu$ and $\operatorname{Re}\alpha(y) \leq \mu$. Since $\operatorname{Re}\alpha(p) = \mu$ and $\operatorname{Re}\alpha$ is linear, this implies $\operatorname{Re}\alpha(x) = \mu = \operatorname{Re}\alpha(y)$, i.e. $x, y \in E_\alpha$. Thus, E_α is an extreme set of K, proving (2).

Fix $S \in \mathcal{P}$ and define

$$\mathcal{P}_S := \{E \in \mathcal{P} : E \subseteq S\},$$

partially ordered by set inclusion. Note $\mathcal{P}_S \neq \emptyset$, as $S \in \mathcal{P}_S$. By Hausdorff's Maximality Principle of Th. 3.23, $\mathcal{P}_S$ contains a maximal chain $\mathcal{C}$. Set

$$M := \bigcap_{E \in \mathcal{C}} E.$$

Since $\emptyset \notin \mathcal{C}$ and $\mathcal{C}$ is totally ordered, the intersection over finitely many elements of $\mathcal{C}$ is always nonempty. Morever, each $E \in \mathcal{C}$ is compact and, thus, closed (having used that $(X, \mathcal{T})$ is T_2 once more). Since the compact set S has the finite intersection property, we obtain $M \neq \emptyset$ and (1) implies $M \in \mathcal{P}_S$ and $M = \min \mathcal{P}_S$. As $\mathcal{C}$ is maximal, no proper subset of M can be an element of $\mathcal{P}$. Now (2) implies that each $\alpha \in X'$ must be constant on M. Since, by assmumption, X' separates points on X, M contains precisely one point, showing $M \subseteq \operatorname{ex}(K)$. Thus, we have proved

$$\underset{S \in \mathcal{P}}{\forall} \quad S \cap \operatorname{ex}(K) \neq \emptyset. \tag{3.29}$$

Next, we have $\operatorname{conv}(\operatorname{ex}(K)) \subseteq K$, since K is convex, implying $\overline{\operatorname{conv}(\operatorname{ex}(K))} \subseteq K$, since K is also closed. As a closed subset of the compact set K, $\overline{\operatorname{conv}(\operatorname{ex}(K))}$ must itself be compact. Seeking a contradiction, assume there exists $x_0 \in K \setminus \overline{\operatorname{conv}(\operatorname{ex}(K))}$. Then Prop. 3.38 furnishes $\alpha \in X'$ such that

$$\max\left\{ \operatorname{Re} \alpha(x) : x \in \overline{\operatorname{conv}(\operatorname{ex}(K))} \right\} < \alpha(x_0). \tag{3.30}$$

Let $\mu := \max\{\operatorname{Re} \alpha(x) : x \in K\}$. According to (2),

$$K_\alpha := \{x \in K : \operatorname{Re} \alpha(x) = \mu\} \in \mathcal{P}.$$

However, (3.30) implies $K_\alpha \cap \operatorname{ex}(K) = \emptyset$ in contradiction to (3.29). Thus, $K = \overline{\operatorname{conv}(\operatorname{ex}(K))}$, completing the proof. ■

Example 3.40. (a) We apply the Krein-Milman Th. 3.39 to show, for each normed space X, that the closed unit ball B' of the dual X' always has extreme points: From Rem. 3.31, we know

$$B' = \{\alpha \in X' : \|\alpha\| \leq 1\},$$

is weak*-compact. We apply the Krein-Milman Th. 3.39 with $(X, \mathcal{T})$ replaced by $(X', \mathcal{T}_{w*})$ and $K := B'$: According to Rem. 3.22(b), the dual of $(X', \mathcal{T}_{w*})$ is $\Phi(X)$, which separates points on X' by Rem. 3.22(a). As $(X', \mathcal{T}_{w*})$ is also T_1 and K is convex and weak*-compact, Th. 3.39 implies

$$B' = \operatorname{cl}_{\mathcal{T}_{w*}}(\operatorname{conv}(\operatorname{ex}(B'))),$$

in particular, $\operatorname{ex}(B') \neq \emptyset$.

(b) The spaces c_0 and $L^1([0,1], \mathcal{L}^1, \lambda^1)$ can not be isometrically isomorphic to the dual of any normed space: According to Ex. 3.37(d),(e), the closed unit balls of c_0 and of $L^1([0,1], \mathcal{L}^1, \lambda^1)$, respectively, do not have any extreme points. On the other hand, we know from (a), that the closed unit ball of every dual of a normed space always has extreme points.

4 Duality, Representation Theorems

4.1 General Normed Space, Adjoint Operators

While the weak*-topology $\mathcal{T}_{w*}$ is defined on the dual X' for *every* topological vector space (cf. Rem. and Def. 3.21, Rem. 3.22), for normed spaces $(X, \|\cdot\|)$, there turns out to be an even more natural and useful topology on X', namely the topology induced by the so-called *operator norm* on X' (in special cases, we have encountered this norm and topology before).

Definition 4.1. Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed vector spaces over $\mathbb{K}$ and let $A : X \longrightarrow Y$ be linear (in this context, the map A is often called a (linear) *operator*). The number

$$\begin{aligned}\|A\| &:= \sup\left\{\frac{\|Ax\|_Y}{\|x\|_X} : x \in X,\ x \neq 0\right\}\\ &= \sup\left\{\|Ax\|_Y : x \in X,\ \|x\|_X \leq 1\right\}\\ &= \sup\left\{\|Ax\|_Y : x \in X,\ \|x\|_X = 1\right\} \in [0,\infty] \qquad (4.1)\end{aligned}$$

is called the *operator norm* of A induced by $\|\cdot\|_X$ and $\|\cdot\|_Y$ (strictly speaking, the term operator norm is only justified if the value is finite, but it is often convenient to use the term in the generalized way defined here).

—

From now on, the space index of a norm will usually be suppressed, i.e. we write just $\|\cdot\|$ instead of both $\|\cdot\|_X$ and $\|\cdot\|_Y$, and also use the same symbol for the operator norm.

Theorem 4.2. *Let $(X, \|\cdot\|)$ and $(Y, \|\cdot\|)$ be normed vector spaces over $\mathbb{K}$.*

(a) *The operator norm does, indeed, constitute a norm on the set of bounded linear maps $\mathcal{L}(X, Y)$.*

(b) *If $A \in \mathcal{L}(X, Y)$, then $\|A\|$ is the smallest Lipschitz constant for A, i.e. $\|A\|$ is a Lipschitz constant for A and $\|Ax - Ay\| \leq L\,\|x - y\|$ for each $x, y \in X$ implies $\|A\| \leq L$.*

(c) *If $(Y, \|\cdot\|)$ is a Banach space, then $(\mathcal{L}(X, Y), \|\cdot\|)$ is a Banach space as well (this holds, indeed, even if $(X, \|\cdot\|)$ is not a Banach space).*

Proof. We leave the proofs of (a) and (b) as an exercise.

(c): Assuming Y to be complete, we show $(\mathcal{L}(X,Y), \|\cdot\|)$ to be complete: Let $(A_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\mathcal{L}(X,Y)$. Fix $x \in X$. According to (b),

$$\underset{m,n\in\mathbb{N}}{\forall} \quad \|A_m(x) - A_n(x)\| \leq \|A_m - A_n\| \, \|x\|, \tag{4.2}$$

showing $(A_n(x))_{n\in\mathbb{N}}$ to be Cauchy in Y. As Y is complete, there exists $A(x) \in Y$ such that $A(x) = \lim_{n\to\infty} A_n(x)$, defining a map

$$A : X \longrightarrow Y, \quad A(x) := \lim_{n\to\infty} A_n(x).$$

As a pointwise limit of linear maps, A is linear by Prop. 2.16. Let $\epsilon \in \mathbb{R}^+$. From (4.2), we obtain

$$\underset{N\in\mathbb{N}}{\exists} \quad \underset{m,n>N}{\forall} \quad \underset{x\in X,\, \|x\|=1}{\forall} \quad \|A_m(x) - A_n(x)\| < \epsilon,$$

implying, as $\lim_{m\to\infty} A_m(x) = A(x)$,

$$\underset{n>N}{\forall} \quad \underset{x\in X,\, \|x\|=1}{\forall} \quad \|A(x) - A_n(x)\| \leq \epsilon. \tag{4.3}$$

Thus,

$$\underset{n>N}{\forall} \quad \underset{x\in X,\, \|x\|=1}{\forall} \quad \|A(x)\| \leq \|A(x) - A_n(x)\| + \|A_n(x)\| \overset{\|x\|=1}{\leq} \epsilon + \|A_n\|,$$

showing $A \in \mathcal{L}(X,Y)$. From (4.3), we obtain $\lim_{n\to\infty} \|A - A_n\| = 0$, completing the proof. ∎

Corollary 4.3. *For a linear map $A : X \longrightarrow Y$ between two normed vector spaces $(X, \|\cdot\|)$ and $(Y, \|\cdot\|)$ over $\mathbb{K}$, the following statements are equivalent:*

(a) *A is bounded.*

(b) *$\|A\| < \infty$.*

(c) *A is Lipschitz continuous.*

(d) *A is continuous.*

(e) *There is $x_0 \in X$ such that A is continuous at x_0.*

Proof. The equivalence of (a), (c), (d), (e) is due to Th. 1.32 and [Phi16b, Th. 2.22].

"(b) $\Rightarrow$ (a)": Let $\|A\| < \infty$ and let $M \subseteq X$ be bounded. Then there is $r > 0$ such that $M \subseteq B_r(0)$. Moreover,

$$\underset{x \in M}{\forall} \quad \|Ax\| \leq \|A\|\|x\| \leq r\|A\|,$$

showing $A(M) \subseteq \overline{B}_{r\|A\|}(0)$. Thus, $A(M)$ is bounded, thereby establishing the case.

"(a) $\Rightarrow$ (b)": Since A is bounded, it maps the bounded set $\overline{B}_1(0) \subseteq X$ into some bounded subset of Y. Thus, there is $r > 0$ such that $A(\overline{B}_1(0)) \subseteq B_r(0) \subseteq Y$. In particular, $\|Ax\| < r$ for each $x \in X$ satisfying $\|x\| = 1$, showing $\|A\| \leq r < \infty$. ■

The most important special case of the above considerations is the case $Y = \mathbb{K}$, where $\mathcal{L}(X, Y) = X'$, which we treat in the following corollary:

Corollary 4.4. *Let $(X, \|\cdot\|)$ be a normed vector space over $\mathbb{K}$ and consider $(X', \|\cdot\|)$, i.e. the dual with the operator norm according to Def.* 4.1 *($(X', \|\cdot\|)$ is called the* normed dual *of the normed space X).*

(a) *$(X', \|\cdot\|)$ is a Banach space (even if X is not).*

(b) *$B' := \{\alpha \in X' : \|\alpha\| \leq 1\}$ is weak*-compact.*

Proof. (a) is due to Th. 4.2(c), as $\mathbb{K}$ is a Banach space; (b) is due to Rem. 3.31. ■

Let $(X, \|\cdot\|)$ be a normed vector space over $\mathbb{K}$. Recall the map Φ on X from Rem. and Def. 3.21, $x \mapsto \Phi(x) = f_x$, where

$$f_x : X' \longrightarrow \mathbb{K}, \quad f_x(\alpha) := \alpha(x).$$

As X is a normed space, X' separates points on X, Φ is injective, and we may identify x with $\Phi(x)$. Thus, if $\alpha \in X'$, then α acts on x and x acts on α, the result being $\alpha(x)$ is both cases. This symmetry (or *duality*) gives rise to the following notation:

Definition 4.5. Let $(X, \|\cdot\|)$ be a normed vector space over $\mathbb{K}$.

(a) The map

$$\langle \cdot, \cdot \rangle : X \times X' \longrightarrow \mathbb{K}, \quad \langle x, \alpha \rangle := \alpha(x), \tag{4.4}$$

is called the *dual pairing* corresponding to X.

(b) The normed dual of $(X', \|\cdot\|)$ is called the *bidual* or the *second dual* of X. One writes $X'' := (X')'$.

(c) The map $\Phi : X \longrightarrow X''$ is called the *canonical embedding* of X into X'' (cf. Th. 4.6 below). The space X is called *reflexive* if, and only if, the map Φ is surjective, i.e. if, and only if, Φ constitutes an isometric isomorphism between X and its bidual.

Theorem 4.6. *Let $(X, \|\cdot\|_X)$ be a normed vector space over $\mathbb{K}$. The canonical embedding $\Phi : X \longrightarrow X''$ does, indeed, map into X''. It constitutes an isometric isomorphism between X and a subspace $\Phi(X)$ of X''. Moreover, if X is a Banach space, then $\Phi(X)$ is closed (i.e. a Banach space).*

Proof. By the definition of the operator norm,

$$\underset{x\in X}{\forall}\quad \underset{\alpha\in X'}{\forall}\quad |\Phi(x)(\alpha)| = |\alpha(x)| \leq \|\alpha\|_{X'}\,\|x\|_X,$$

showing $\Phi(x)$ to be a bounded linear functional on X' (i.e. $\Phi(x) \in X''$) and $\|\Phi(x)\|_{X''} \leq \|x\|_X$. On the other hand, given $x \in X$, by the Hahn-Banach Cor. 3.4(b), there exists $\alpha \in X'$ such that

$$|\Phi(x)(\alpha)| = |\alpha(x)| = \|x\|_X,$$

implying $\|\Phi(x)\|_{X''} = \|x\|_X$. Thus, Φ is an isometric isomorphism onto $\Phi(X)$. In particular, if X is complete, then so is $\Phi(X)$, showing $\Phi(X)$ to be Banach as well as a closed subspace of X''. ■

Remark 4.7. Let $(X, \|\cdot\|)$ be a normed vector space over $\mathbb{K}$ with normed dual $(X', \|\cdot\|)$.

(a) Clearly, the dual pairing, as defined in (4.4), is bilinear.

(b) Consider the following three topologies on X': The (operator) norm topology $\mathcal{T}'$ (also called the *strong* topology on X'), the weak*-topology $\mathcal{T}_{\mathrm{w}*}$, and the weak topology (i.e. the X''-topology) $\mathcal{T}'_{\mathrm{w}}$. Since $\Phi(X) \subseteq X''$, we have

$$\mathcal{T}_{\mathrm{w}*} \subseteq \mathcal{T}'_{\mathrm{w}} \subseteq \mathcal{T}',$$

that means, the strong topology is, indeed, the strongest of the three, whereas the weak*-topology is the weakest. If $\dim X = n \in \mathbb{N}$, then $X'' \cong X' \cong \mathbb{K}^n$ and all three topologies are the same (cf. Ex. 3.18(a)). If $\dim X = \infty$, then $\dim X' = \infty$ and we know that $\mathcal{T}'_{\mathrm{w}}$ is strictly weaker than $\mathcal{T}'$ according to Ex. 3.18(b), implying $\mathcal{T}_{\mathrm{w}*}$ to be strictly weaker than $\mathcal{T}'$ as well. If X is reflexive, then $\mathcal{T}'_{\mathrm{w}} = \mathcal{T}_{\mathrm{w}*}$ (since $\Phi(X) = X''$). In general, $\mathcal{T}_{\mathrm{w}*}$ can be strictly weaker than $\mathcal{T}'_{\mathrm{w}}$, but $\mathcal{T}'_{\mathrm{w}}$ does not appear to be of particular use in such cases (for an example, where $\mathcal{T}_{\mathrm{w}*}$ is strictly weaker, see Ex. 4.8(e) below).

Example 4.8. (a) If $(X, \|\cdot\|)$ is a normed vector space over $\mathbb{K}$, $\dim X = n \in \mathbb{N}$, then $X'' \cong X' \cong \mathbb{K}^n$ (cf. Rem. 4.7(b) above). In particular, $(X, \|\cdot\|)$ is reflexive.

(b) If $(X, \|\cdot\|)$ is a normed vector space over $\mathbb{K}$ that is not a Banach space, then it can never be reflexive, since X'' is always a Banach space.

(c) In Sec. 4.2 below, we will see that every Hilbert space is reflexive. In Sec. 4.4 below, we will see that, for each measure space $(\Omega, \mathcal{A}, \mu)$ and each $1 < p < \infty$, $L^p(\mu)$ is reflexive.

(d) As a caveat, we note that it can occur that a Banach space $(X, \|\cdot\|)$ is isometrically isomorphic to its bidual, but *not* reflexive, i.e. the canonical embedding Φ is not surjective, but there exists a different isometric isomorphism $\phi : X \cong X''$, $\phi \neq \Phi$. An example of such a Banach space was constructed by R.C. James in 1951 (see [Wer11, Excercise I.4.8] and [Wer11, page 105] for the definition and further references).

(e) The spaces c_0 and $L^1([0,1], \mathcal{L}^1, \lambda^1)$ are not reflexive, since we know from Ex. 3.40(b) that they are not isometrically isomorphic to the dual of any normed space. For c_0, we know $(c_0)' \cong l^1$ from Prop. 2.32(b). In Sec. 4.4 below, we will see that $(c_0)'' \cong (l^1)' \cong l^\infty$. As before, let e_k, $k \in \mathbb{N}$, denote the standard unit vector in $\mathbb{K}^{\mathbb{N}}$. Then $(e_k)_{k\in\mathbb{N}}$ is a sequence in l^1. If $x = (x_k)_{k\in\mathbb{N}}$ is in c_0, then

$$\lim_{k\to\infty} \langle x, e_k \rangle = \lim_{k\to\infty} \sum_{i=1}^{\infty} x_i (e_k)_i = \lim_{k\to\infty} x_k = 0,$$

showing e_k weak*-converges to 0. However, $\alpha := (1)_{k\in\mathbb{N}}$ is an element of l^∞ and

$$\underset{k\in\mathbb{N}}{\forall} \quad \alpha(e_k) = \sum_{i=1}^{\infty} (e_k)_i = 1,$$

showing that e_k does not converge to 0 in the $(c_0)''$-topology on $(c_0)'$, showing the weak*-topology on $(c_0)'$ to be strictly weaker than the weak topology on $(c_0)'$ (i.e. $\mathcal{T}_{\mathrm{w}*} \subsetneq \mathcal{T}'_{\mathrm{w}}$ in terms of the notation from Rem. 4.7(b)).

Definition 4.9. Let $(X, \|\cdot\|)$, $(Y, \|\cdot\|)$ be normed vector spaces over $\mathbb{K}$, $A \in \mathcal{L}(X, Y)$. Then the map

$$A' : Y' \longrightarrow X', \quad \underset{\beta\in Y'}{\forall} \ \underset{x\in X}{\forall} \quad A'(\beta)(x) = \langle x, A'(\beta) \rangle := \beta(A(x)) = \langle A(x), \beta \rangle \tag{4.5}$$

is called the *adjoint operator* or just the *adjoint* of A. Caveat: In Hilbert spaces X, Y, the present adjoint is, in general, not the same as the *Hilbert adjoint* to be defined in Sec. 4.2 below.

—

Note that the dual pairings in (4.5) are defined on different spaces. Still, in particular when dealing with adjoint operators, the dual pairing notation is quite useful, since we see we can move a linear operator across the comma in a dual pairing, provided we replace the operator by its adjoint.

Lemma 4.10. *Let $(X, \|\cdot\|)$, $(Y, \|\cdot\|)$, $(Z, \|\cdot\|)$ be normed vector spaces over $\mathbb{K}$ and consider $A \in \mathcal{L}(X, Y)$, $B \in \mathcal{L}(Y, Z)$. Then*

$$\|B \circ A\| \leq \|B\| \, \|A\| \tag{4.6}$$

holds with respect to the corresponding operator norms.

Proof. Let $x \in X$ with $\|x\| = 1$. If $Ax = 0$, then $\|B(A(x))\| = 0 \leq \|B\| \, \|A\|$. If $Ax \neq 0$, then one estimates

$$\big\|B(Ax)\big\| = \|Ax\| \, \left\|B\left(\frac{Ax}{\|Ax\|}\right)\right\| \leq \|A\| \, \|B\|,$$

thereby establishing the case. ∎

Definition 4.11. Let $(X, \|\cdot\|)$ be a normed vector spaces over $\mathbb{K}$ and let $V \subseteq X$ be a vector subspace. Then the vector space

$$V^\perp := \left\{\alpha \in X' : \underset{x \in V}{\forall} \langle x, \alpha \rangle = 0\right\} \tag{4.7}$$

is called the *annihilator* of V (analogously, one can define the annihilator $V^\perp \subseteq X$ of a subspace V of X').

Theorem 4.12. *Let $(X, \|\cdot\|)$, $(Y, \|\cdot\|)$ be normed vector spaces over $\mathbb{K}$.*

(a) *For each $A \in \mathcal{L}(X, Y)$, the adjoint A' is well-defined by Def.* 4.9, *i.e. $A'(\beta) \in X'$ for each $\beta \in Y'$. Moreover, A' is the unique map on Y' such that* (4.5) *holds.*

(b) *One has that $A \mapsto A'$ is a (linear) isometric isomorphism of $\mathcal{L}(X, Y)$ onto a subspace of $\mathcal{L}(Y', X')$ (not necessarily surjective onto $\mathcal{L}(Y', X')$, see Ex.* 4.13(c) *below).*

(c) $(\mathrm{Id}_X)' = \mathrm{Id}_{X'}$.

(d) *If $(Z, \|\cdot\|)$ is another normed vector space over $\mathbb{K}$, $A \in \mathcal{L}(X, Y)$, $B \in \mathcal{L}(Y, Z)$, then*

$$(B \circ A)' = A' \circ B'. \tag{4.8}$$

(e) *If $\Phi_X : X \longrightarrow X''$ and $\Phi_Y : Y \longrightarrow Y''$ are the canonical embeddings, then*

$$\forall_{A \in \mathcal{L}(X,Y)} \quad A'' \circ \Phi_X = \Phi_Y \circ A \tag{4.9}$$

(thus, we can interpret A'' as an extension of A from X to X''). Moreover, $B \in \mathcal{L}(Y', X')$ is an adjoint of some $A \in \mathcal{L}(X,Y)$ if, and only if, $B'(\Phi_X(X)) \subseteq \Phi_Y(Y)$.

(f) *If $A \in \mathcal{L}(X,Y)$, then $\ker(A') = (A(X))^\perp$.*

(g) *If X, Y are both Banach spaces and $A \in \mathcal{L}(X,Y)$, then $A^{-1} \in \mathcal{L}(Y,X)$ exists if, and only if, $(A')^{-1} \in \mathcal{L}(X', Y')$ exists, and, in that case,*

$$(A')^{-1} = (A^{-1})'. \tag{4.10}$$

Proof. (a): If $A \in \mathcal{L}(X,Y)$, $\beta \in Y'$, then $A'(\beta) = \beta \circ A \in X'$, as both A and β are linear and continuous. For each $\beta \in Y'$, $x \mapsto \langle A(x), \beta \rangle$ uniquely determines a map $A'(\beta)$ on X, i.e. $\beta \mapsto A'(\beta)$ is uniquely determined by (4.5).

(b): Exercise.

(c): One has

$$\forall_{\alpha \in X'} \; \forall_{x \in X} \quad (\mathrm{Id}_X)'(\alpha)(x) = \alpha(x),$$

showing $(\mathrm{Id}_X)' = \mathrm{Id}_{X'}$.

(d): Exercise.

(e): We have to show $A''(\Phi_X(x)) = \Phi_Y(A(x))$ for each $x \in X$. To this end, let $\beta \in Y'$ and compute

$$A''(\Phi_X(x))(\beta) = \big(\Phi_X(x)\big)\big(A'(\beta)\big) = \big(A'(\beta)\big)(x) = \beta(A(x)) = \big(\Phi_Y(A(x))\big)(\beta),$$

proving the desired identity. Now let $B \in \mathcal{L}(Y', X')$. Suppose, there exists $A \in \mathcal{L}(X,Y)$ such that $B = A'$. Then

$$B'(\Phi_X(X)) = A''(\Phi_X(X)) \overset{(4.9)}{=} \Phi_Y(A(X)) \subseteq \Phi_Y(Y).$$

Conversely, assume $B'(\Phi_X(X)) \subseteq \Phi_Y(Y)$. Then we can define

$$A : X \longrightarrow Y, \quad A := \Phi_Y^{-1} \circ B' \circ \Phi_X.$$

In consequence,

$$A'' \circ \Phi_X \overset{(4.9)}{=} \Phi_Y \circ A = B' \circ \Phi_X,$$

implying $A'' = B'$. As we know from (b) that forming the adjoint is an injective map, $A' = B$.

(f): We have

$$\beta \in \ker(A') \quad \Leftrightarrow \quad \underset{x\in X}{\forall} A'(\beta)(x) = \langle A(x), \beta \rangle = 0 \quad \Leftrightarrow \quad \beta \in (A(X))^{\perp}.$$

(g): Exercise. ■

Example 4.13. **(a)** Let $m, n \in \mathbb{N}$, let X be $\mathbb{K}^n$ and Y be $\mathbb{K}^m$, each with the norm topology. Then $\mathcal{L}(X, Y) = \mathbb{K}^{mn}$ and each $A \in \mathcal{L}(X, Y)$ can be represented by and $m \times n$ matrix $A = (a_{kl})_{(k,l)\in\{1,\dots,m\}\times\{1,\dots,n\}}$. We claim that the adjoint A' of A is represented by the transpose matrix $A^{\mathrm{t}} = (a_{lk})_{(k,l)\in\{1,\dots,m\}\times\{1,\dots,n\}}$: Let $\beta = (\beta_1, \dots, \beta_m) \in Y'$ and $x = (x_1, \dots, x_n)^{\mathrm{t}} \in X$, where we interpret β as a row vector and x as a column vector such that the application of the respective linear maps is just matrix multiplication with the representing matrices. Then

$$(\beta A)x = (A^{\mathrm{t}}\beta^{\mathrm{t}})^{\mathrm{t}}x = \beta(Ax), \tag{4.11}$$

which holds due to matrix multiplication being associative. Note that, even for $\mathbb{K} = \mathbb{C}$, the adjoint is just the transpose, without complex conjugation (in contrast to the Hilbert adjoint of Sec. 4.2 below).

(b) Consider the left shift operator

$$A : c_0 \longrightarrow c_0, \quad A(x_1, x_2, \dots) := (x_2, x_3, \dots).$$

We know from Prop. 2.32(b) that $(c_0)' \cong l^1$. We claim that the adjoint is the right shift operator

$$A' : l^1 \longrightarrow l^1, \quad A'(a_1, a_2, \dots) := (0, a_1, a_2, \dots) :$$

Indeed,

$$\underset{\alpha=(a_k)_{k\in\mathbb{N}}\in(c_0)'}{\forall} \quad \underset{x=(x_k)_{k\in\mathbb{N}}\in c_0}{\forall} \quad \alpha(Ax) = \sum_{k=1}^{\infty} a_k x_{k+1} = A'(\alpha)(x).$$

(c) Consider

$$B : l^1 \longrightarrow l^1, \quad B(a_1, a_2, \dots) := \left(\sum_{k=1}^{\infty} a_k, 0, 0, \dots \right).$$

We will see in Sec. 4.4 below that $(l^1)' \cong l^{\infty}$, where the assignment rule is still the same as for the isomorphism $(c_0)' \cong l^1$. Here, we use this to verify

$$B' : l^{\infty} \longrightarrow l^{\infty}, \quad B'(b_1, b_2, \dots) := (b_1, b_1, b_1, \dots) :$$

Indeed,

$$\forall_{\beta=(b_k)_{k\in\mathbb{N}}\in(l^1)'} \quad \forall_{a=(a_k)_{k\in\mathbb{N}}\in l^1} \quad \beta(Ba)=\sum_{k=1}^{\infty} b_1 a_k = B'(\beta)(a).$$

In particular, we see that $B'(c_0)\not\subseteq c_0$. Thus, according to Th. 4.12(e), B is not the adjoint of some $A\in\mathcal{L}(c_0,c_0)$, showing $'\,:\,\mathcal{L}(c_0,c_0)\longrightarrow\mathcal{L}((c_0)',(c_0)')$ is not surjective.

4.2 Hilbert Space, Riesz Representation Theorem I

Let X be a vector space over $\mathbb{K}$ and let $\langle\cdot,\cdot\rangle$ be an inner product (also called a scalar product) on X. We know from [Phi16b, Prop. 1.65] that the inner product induces a norm on X via

$$\|\cdot\|:\,X\longrightarrow\mathbb{R}_0^+,\quad \|x\|:=\sqrt{\langle x,x\rangle}.$$

In [Phi16b, Def. 1.66], we called $\big(X,\langle\cdot,\cdot\rangle\big)$ an *inner product space* or a *pre-Hilbert space*. Moreover, we called an inner product space a *Hilbert space* if, and only if, it was complete (i.e. a Banach space).

Theorem 4.14. *Let $\big(X,\langle\cdot,\cdot\rangle\big)$ be an inner product space over $\mathbb{K}$ with induced norm $\|\cdot\|$. Then the following assertions hold true:*

(a) Cauchy-Schwarz Inequality:

$$\forall_{x,y\in X}\quad |\langle x,y\rangle|\le\|x\|\,\|y\|.$$

(b) *For each $x,y\in X$, the maps*

$$\begin{aligned}\alpha_y:\,X\longrightarrow\mathbb{K},&\quad \alpha_y(a):=\langle a,y\rangle,\\ \beta_x:\,X\longrightarrow\mathbb{K},&\quad \beta_x(a):=\langle x,a\rangle,\end{aligned}$$

are both continuous (α_y is linear, β_x is conjugate-linear).

(c) Parallelogram Law:

$$\forall_{x,y\in X}\quad \|x+y\|^2+\|x-y\|^2=2\big(\|x\|^2+\|y\|^2\big). \tag{4.12}$$

(d) *If $\mathbb{K}=\mathbb{R}$, then*

$$\forall_{x,y\in X}\quad \langle x,y\rangle=\frac{1}{4}\big(\|x+y\|^2-\|x-y\|^2\big).$$

If $\mathbb{K} = \mathbb{C}$, *then*

$$\forall_{x,y \in X} \quad \langle x, y\rangle = \frac{1}{4}\big(\|x+y\|^2 - \|x-y\|^2 + i\,\|x+iy\|^2 - i\,\|x-iy\|^2\big).$$

Proof. (a) was proved as [Phi16b, Th. 1.64].

(b) holds, as (a) says that α_y and $\overline{\beta_x}$ are bounded linear functionals on X.

(c): The computation

$$\begin{aligned}\|x+y\|^2 + \|x-y\|^2 &= \|x\|^2 + \langle x, y\rangle + \langle y, x\rangle + \|y\|^2 + \|x\|^2 - \langle x, y\rangle - \langle y, x\rangle + \|y\|^2 \\ &= 2\big(\|x\|^2 + \|y\|^2\big)\end{aligned}$$

proves (4.12).

(d): If $\mathbb{K} = \mathbb{R}$, then

$$\|x+y\|^2 - \|x-y\|^2 = 4\,\langle x, y\rangle.$$

If $\mathbb{K} = \mathbb{C}$, then

$$\begin{aligned}\|x+y\|^2 - \|x-y\|^2 + i\,\|x+iy\|^2 - i\,\|x-iy\|^2 &= 4\operatorname{Re}\langle x, y\rangle + 4i\operatorname{Re}\langle x, iy\rangle \\ &= 4\operatorname{Re}\langle x, y\rangle + 4\operatorname{Im}\langle x, y\rangle = 4\,\langle x, y\rangle,\end{aligned}$$

proving (d). ∎

One can actually also show (with more effort) that a normed space that satisfies (4.12) must be an inner product space, see, e.g., [Wer11, Th. V.1.7].

Example 4.15. (a) Let $(\Omega, \mathcal{A}, \mu)$ be a measure space, $\Omega \neq \emptyset$. Then $L^2(\mu)$ constitutes a Hilbert space, where $\|\cdot\|_2$ on $L^2(\mu)$ is induced by the inner product

$$\langle\cdot,\cdot\rangle : L^2(\mu) \times L^2(\mu) \longrightarrow \mathbb{K}, \quad \langle f, g\rangle := \int_\Omega f\,\overline{g}\,\mathrm{d}\mu \; : \tag{4.13}$$

First, note that (4.13) is well-defined, since $f\,\overline{g} \in L^1(\mu)$ by the Hölder inequality. We now verify that (4.13), indeed, defines an inner product: If $f \in L^2(\mu)$, $f \neq 0$, then there exists $A \in \mathcal{A}$ with $\mu(A) > 0$ and $\int_A |f|^2\,\mathrm{d}\mu > 0$, implying

$$\langle f, f\rangle = \int_\Omega |f|^2\,\mathrm{d}\mu \geq \int_A |f|^2\,\mathrm{d}\mu > 0.$$

Next, let $f, g, h \in L^2(\mu)$ and $\lambda, \mu \in \mathbb{K}$. One computes

$$\langle \lambda f + \mu g, h\rangle = \int_\Omega (\lambda f + \mu g)\,\overline{h}\,\mathrm{d}\mu = \lambda\int_\Omega f\overline{h}\,\mathrm{d}\mu + \mu\int_\Omega g\overline{h}\,\mathrm{d}\mu = \lambda\langle f, h\rangle + \mu\lambda\langle g, h\rangle.$$

Moreover,

$$\underset{f,g\in L^2(\mu)}{\forall} \quad \langle f,g\rangle = \int_\Omega f\,\overline{g}\,\mathrm{d}\mu = \overline{\int_\Omega \overline{f}\,g\,\mathrm{d}\mu} = \overline{\langle g,f\rangle},$$

showing $\langle\cdot,\cdot\rangle$ to be an inner product on $L^2(\mu)$. That the inner product induces the 2-norm on $L^2(\mu)$ is immediate from the definition of the 2-norm. Finally, $L^2(\mu)$ is a Hilbert space, since it is complete by [Phi17, Th. 2.44(a)].

(b) As a special case of (a), consider $(S,\mathcal{P}(S),\mu)$, where $S\neq\emptyset$ is a set and μ is counting measure on S (cf. [Phi17, Ex. 1.12(b)]) and define

$$l^2(S) := L^2(S,\mathcal{P}(S),\mu).$$

Then $l^2 = l^2(\mathbb{N})$. In general,

$$l^2(S) = \left\{ f:\, S \longrightarrow \mathbb{K}:\ \begin{array}{l} \{s\in S:\ f(s)\neq 0\} \text{ is finite or countable} \\ \text{and } \sum_{s\in S} |f(s)|^2 < \infty \end{array} \right\},$$

where, due to the given absolute convergence, the sum $\sum_{s\in S}|f(s)|^2$ can be evaluated using an arbitrary enumeration of $\{s\in S:\ f(s)\neq 0\}$. If $f,g\in l^2(S)$, then, by the Hölder inequality, $f\overline{g}\in l^1(S) := L^1(S,\mathcal{P}(S),\mu)$ and

$$\langle f,g\rangle = \int_S f\overline{g}\,\mathrm{d}\mu = \sum_{s\in S} f(s)\,\overline{g(s)}$$

is well-defined.

(c) As an example of an inner product space that is not a Hilbert space consider the space c_{00} of sequences in $\mathbb{K}$ that are finally constant and equal to 0 with the 2-norm. Then c_{00} is a vector subspace of l^2, but not complete: Define

$$x := (x_n)_{n\in\mathbb{N}} := (2^{-1}, 2^{-2}, \dots),$$

$$\underset{k\in\mathbb{N}}{\forall} \quad x^k := (x^k_n)_{n\in\mathbb{N}}, \quad x^k_n := \begin{cases} x_n & \text{for } n\leq k, \\ 0 & \text{for } n>k. \end{cases}$$

Then $\lim_{k\to\infty} x^k = x$ in l^2 and $(x^k)_{k\in\mathbb{N}}$ is a Cauchy sequence. However, $x\notin c_{00}$, showing c_{00} not to be complete.

Definition 4.16. Let $\big(X,\langle\cdot,\cdot\rangle\big)$ be an inner product space over $\mathbb{K}$.

(a) $x,y\in X$ are called *orthogonal* or *perpendicular* (denoted $x\perp y$) if, and only if, $\langle x,y\rangle = 0$.

(b) Let $E \subseteq X$. Define the perpendicular space $E^\perp$ to E (called E *perp*) by

$$E^\perp := \left\{ y \in X : \underset{x \in E}{\forall} \langle x, y \rangle = 0 \right\}. \tag{4.14}$$

Lemma 4.17. *Let $\big(X, \langle \cdot, \cdot \rangle\big)$ be an inner product space over $\mathbb{K}$, $E \subseteq X$.*

(a) $E \cap E^\perp \subseteq \{0\}$.

(b) $E \subseteq (E^\perp)^\perp$.

(c) *$E^\perp$ is a closed vector subspace of X.*

Proof. (a): If $x \in E \cap E^\perp$, then $\langle x, x \rangle = 0$, implying $x = 0$.
(b): If $x \in E$ and $y \in E^\perp$, then $\langle x, y \rangle = 0$, showing $x \in (E^\perp)^\perp$.
(c): We have $0 \in E^\perp$ and

$$\underset{\lambda, \mu \in \mathbb{K}}{\forall} \quad \underset{y_1, y_2 \in E^\perp}{\forall} \quad \underset{x \in E}{\forall} \quad \langle x, \lambda y_1 + \mu y_2 \rangle = \overline{\lambda} \langle x, y_1 \rangle + \overline{\mu} \langle x, y_2 \rangle = 0,$$

showing $\lambda y_1 + \mu y_2 \in E^\perp$, i.e. $E^\perp$ is a vector space. Using the notation from Th. 4.14(b), we have

$$E^\perp = \bigcap_{x \in E} \beta_x^{-1}(\{0\}),$$

where each set $\beta_x^{-1}(\{0\}$ is closed (as β_x is continuous), showing $E^\perp$ to be closed as well. ∎

Theorem 4.18 (Projection Theorem)**.** *Let $\big(H, \langle \cdot, \cdot \rangle\big)$ be a Hilbert space over $\mathbb{K}$ with induced norm $\| \cdot \|$. Let $x_0 \in H$. If $C \subseteq H$ is nonempty, closed, and convex, then there exists a unique $y \in C$ such that*

$$\|y - x_0\| = \inf\{\|x - x_0\| : x \in C\}. \tag{4.15}$$

Proof. First, consider $x_0 = 0$. Set $\delta := \inf\{\|x\| : x \in C\}$. For $x, y \in H$, apply the parallelogram law (4.12) to $\frac{1}{2}x$ and $\frac{1}{2}y$ to obtain

$$\frac{1}{4}\|x - y\|^2 = \frac{1}{2}\|x\|^2 + \frac{1}{2}\|y\|^2 - \left\| \frac{x + y}{2} \right\|^2.$$

If $x, y \in C$, then the convexity of C implies $\frac{x+y}{2} \in C$ and, thus,

$$\underset{x, y \in C}{\forall} \quad \|x - y\|^2 \leq 2\|x\|^2 + 2\|y\|^2 - 4\delta^2. \tag{4.16}$$

In particular, if $x, y \in C$ with $\|x\| = \|y\| = \delta$, then $x = y$, proving the uniqueness statement of the theorem. According to the definition of δ, there exists a sequence $(c_n)_{n\in\mathbb{N}}$ in C such that $\delta = \lim_{n\to\infty} \|c_n\|$. Applying (4.16) with $x := c_m$ and $y := c_n$ for $m, n \in \mathbb{N}$, shows $(c_n)_{n\in\mathbb{N}}$ to be a Cauchy sequence in H. As H is complete, there exists $y := \lim_{n\to\infty} c_n \in H$. Since C is closed, we know $y \in C$ as well. By the continuity of the norm, $\|y\| = \lim_{n\to\infty} \|c_n\| = \delta$, proving (4.15). Now let $x_0 \in H$ be arbitrary. Then we know from above that $C - x_0$ contains a unique y_0 such that

$$\|y_0\| = \inf\{\|x\| : x \in C - x_0\} = \inf\{\|x - x_0\| : x \in C\},$$

implying $y := y_0 + x_0$ to be the unique element of C, satisfying (4.15). ■

Lemma 4.19. *Let $\big(X, \langle\cdot,\cdot\rangle\big)$ be an inner product space over $\mathbb{K}$ with induced norm $\|\cdot\|$. Let $C \subseteq X$ be convex, $y \in C$. Then, given $x_0 \in X$,* (4.15) *is equivalent to*

$$\underset{x\in C}{\forall} \quad \operatorname{Re}\langle x_0 - y, x - y\rangle \leq 0. \tag{4.17}$$

Proof. Exercise. ■

Theorem 4.20 (Orthogonal Projection Theorem)**.** *Let $\big(H, \langle\cdot,\cdot\rangle\big)$ be a Hilbert space over $\mathbb{K}$ with induced norm $\|\cdot\|$. Let $V \subseteq H$ be a closed vector subspace of H. Then Th.* 4.18 *gives rise to maps*

$$P_V : H \longrightarrow V, \quad P_{V^\perp} : H \longrightarrow V^\perp,$$

where, given $x_0 \in H$, $P_V(x_0)$ (resp. $P_{V^\perp}(x_0)$) is the unique element $y \in V$ (resp. $y \in V^\perp$), satisfying the equivalent conditions (4.15) *and* (4.17) *with C replaced by V (resp. with C replaced by $V^\perp$). Then the map P_V (resp. $P_{V^\perp}$) is called the* orthogonal projection *onto V (resp. onto $V^\perp$). Moreover, the following assertions hold true:*

(a) *For each $x_0 \in H$, $P_V(x_0)$ and $P_{V^\perp}(x_0)$ are the nearest points to x_0 in V and in $V^\perp$, respectively.*

(b) *Given $x_0 \in H$, $P_V(x_0)$ is the unique element of V, satisfying*

$$\underset{v\in V}{\forall} \quad \langle v, x_0 - P_V(x_0)\rangle = 0 \tag{4.18}$$

(justifying the name orthogonal projection*).*

(c) *P_V and $P_{V^\perp}$ are continuous linear maps with* $\ker(P_V) = V^\perp$ *and* $\ker(P_{V^\perp}) = V$*. If $V \neq \{0\}$, then $\|P_V\| = 1$; if $V \neq H$, then $\|P_{V^\perp}\| = 1$.*

(d) *$P_{V^\perp} = \operatorname{Id} - P_V$.*

(e) *$H = V \oplus V^\perp$.*

(f) $V = (V^\perp)^\perp$.

(g) $\|x\|^2 = \|P_V(x)\|^2 + \|P_{V^\perp}(x)\|^2$ *holds for each* $x \in H$.

Proof. (a) is merely a restatement of (4.15).

(b): Let $x_0 \in H$. According to (4.17),

$$\forall_{v\in V} \quad \mathrm{Re}\langle x_0 - P_V(x_0),\, v - P_V(x_0)\rangle \leq 0,$$

which, as $v \mapsto v - P_V(x_0)$ is a bijection on V, is equivalent to

$$\forall_{v\in V} \quad \mathrm{Re}\langle x_0 - P_V(x_0),\, v\rangle \leq 0. \tag{4.19}$$

Since, for each $v \in V$, (4.19) also holds for v replaced by $-v$ and by $i\,v$ (for $\mathbb{K} = \mathbb{C}$), we see that $P_V(x_0)$ satisfies (4.18). Conversely, if (4.18) holds, then we can use the bijection $v \mapsto v - P_V(x_0)$ on V again to conclude

$$\forall_{v\in V} \quad \langle x_0 - P_V(x_0),\, v - P_V(x_0)\rangle = 0,$$

which implies (4.17) (even with equality and without Re, which is due to V being a vector subspace).

We can restate (b) by saying $P_V(x_0)$ is the unique element of V such that $x_0 - P_V(x_0) \in V^\perp$. Thus, if $\lambda, \mu \in \mathbb{K}$ and $x_1, x_2 \in H$, then

$$\big(\lambda x_1 - \lambda P_V(x_1)\big) + \big(\mu x_2 - \mu P_V(x_2)\big) \in V^\perp,$$

showing

$$P_V(\lambda x_1 + \mu x_2) = \lambda P_V(x_1) + \mu P_V(x_2),$$

i.e. P_V is linear. Next, we obtain

$$P_V(x_0) = 0 \quad \Leftrightarrow \quad x_0 \in V^\perp,$$

proving $\ker(P_V) = V^\perp$. From (4.18), we see that $\mathrm{Id} - P_V$ is a linear map, mapping H into $V^\perp$. Thus, combining

$$\forall_{x_0\in H} \quad x_0 = P_V(x_0) + x_0 - P_V(x_0) = P_V(x_0) + (\mathrm{Id} - P_V)(x_0)$$

with Lem. 4.17(a) proves (e). From Lem. 4.17(b), we know $V \subseteq (V^\perp)^\perp$. If $x \in (V^\perp)^\perp$, write $x = x_1 + x_2$ with $x_1 \in V$, $x_2 \in V^\perp$. Since $x_2 = x - x_1 \in (V^\perp)^\perp \cap V^\perp$, we obtain $x_2 = 0$ and $x = x_1 \in V$, proving (f).

To prove (d), note that one can replace V by $V^\perp$ in the above arguments, i.e. we already know $P_{V^\perp}$ to be a linear map with $\ker(P_{V^\perp}) = (V^\perp)^\perp = V$ and such that, for each $x_0 \in H$, $P_{V^\perp}(x_0)$ is the unique element of $V^\perp$ with $x_0 - P_{V^\perp}(x_0) \in (V^\perp)^\perp = V$. Since $y := (\operatorname{Id} - P_V)(x_0) \in V^\perp$ has the property $x_0 - y = x_0 - x_0 + P_V(x_0) \in V$, the proof of (d) is complete.

(g): Since, for each $x \in X$, $x = P_V(x) + P_{V^\perp}(x)$ as well as $P_V(x) \perp P_{V^\perp}(x)$, (g) is immediate from Pythagoras' theorem of [Phi16b, (1.48)].

If $\|x\| = 1$, then (g) implies

$$\|P_V(x)\| = 1 - \|P_{V^\perp}(x)\| \leq 1,$$

showing $\|P_V\| \leq 1$. If there exists $0 \neq v \in V$, then $P_V(v) = v$, showing $\|P_V\| = 1$. Replacing V by $V^\perp$ shows $\|P_{V^\perp}\| = 1$ for $V \neq H$, completing the proof of (c) and the theorem. ■

Theorem 4.21 (Riesz Representation Theorem)**.** *Let $\big(H, \langle \cdot, \cdot \rangle\big)$ be a Hilbert space over $\mathbb{K}$ with induced norm $\|\cdot\|$. Then the map*

$$\psi : H \longrightarrow H', \quad \psi(y) := \alpha_y, \tag{4.20}$$

where

$$\alpha_y : H \longrightarrow \mathbb{K}, \quad \alpha_y(a) = \langle a, y \rangle,$$

is the map from Th. 4.14(b)*, is bijective, conjugate-linear, and isometric (in particular, each $\alpha \in H'$ can be represented by $y \in H$ with $\|y\| = \|\alpha\|$, and ψ is an isometric isomorphism (i.e. linear) for $\mathbb{K} = \mathbb{R}$).*

Proof. We already know from Th. 4.14(b) that, for each $y \in H$, α_y is linear and continuous, i.e. ψ is well-defined. Moreover,

$$\underset{\lambda,\mu\in\mathbb{K}}{\forall} \quad \underset{y_1,y_2\in H}{\forall} \quad \underset{a\in H}{\forall} \quad \begin{aligned} \psi(\lambda y_1 + \mu y_2)(a) &= \langle a, \lambda y_1 + \mu y_2 \rangle = \overline{\lambda}\langle a, y_1 \rangle + \overline{\mu}\langle a, y_2 \rangle \\ &= (\overline{\lambda}\psi(y_1) + \overline{\mu}\psi(y_2))(a), \end{aligned}$$

showing ψ to be conjugate-linear. By the Cauchy-Schwarz inequality, we have

$$\underset{y,a\in H}{\forall} \quad |\psi(y)(a)| \leq \|a\| \, \|y\|,$$

showing $\|\psi(y)\| \leq \|y\|$. On the other hand, if $y \neq 0$, then

$$\frac{|\psi(y)(y)|}{\|y\|} = \frac{\langle y, y \rangle}{\|y\|} = \|y\|,$$

showing $\|\psi(y)\| = \|y\|$, i.e. ψ is isometric and, in particular, injective (so far, we have actually not used that H is complete). It remains to show, ψ is surjective. Let $\alpha \in H'$, $\alpha \neq 0$, $V := \ker(\alpha)$. Then V is a closed vector subspace of H, $H = V \oplus V^\perp$ by Th. 4.20(e), and, as $\alpha \neq 0$, $V^\perp \neq \{0\}$. Let $x \in H$ and let $z \in V^\perp$ be such that $\|z\| = 1$. Moreover, let $u := (\alpha(x))z - (\alpha(z))x$. Then $\alpha(u) = 0$, i.e. $u \in V$ and $\langle u, z\rangle = 0$. Thus,

$$\alpha(x) = \alpha(x)\langle z, z\rangle - \langle u, z\rangle = \alpha(z)\langle x, z\rangle.$$

In consequence, if $y := \overline{\alpha(z)}\, z$, then

$$\psi(y)(x) = \langle x, y\rangle = \alpha(z)\langle x, z\rangle = \alpha(x).$$

Since $x \in H$ was arbitrary, this proves $\psi(y) = \alpha$, i.e. ψ is sujective. ■

Corollary 4.22. *Let $\big(H, \langle\cdot,\cdot\rangle\big)$ be a Hilbert space over $\mathbb{K}$ with induced norm $\|\cdot\|$.*

(a) *H' is a Hilbert space over $\mathbb{K}$, where, using the map ψ of* (4.20),

$$\langle\cdot,\cdot\rangle : H' \times H' \longrightarrow \mathbb{K}, \quad \langle\alpha, \beta\rangle := \langle\psi^{-1}(\beta), \psi^{-1}(\alpha)\rangle, \tag{4.21}$$

defines an inner product on H', satisfying, with regard to the operator norm on H',

$$\forall_{\alpha \in H'} \quad \|\alpha\| = \sqrt{\langle\alpha, \alpha\rangle}.$$

(b) *H is reflexive.*

(c) *A sequence $(x_k)_{k\in\mathbb{N}}$ converges weakly to $x \in H$ if, and only if,*

$$\forall_{y \in H} \quad \lim_{k\to\infty} \langle x_k, y\rangle = \langle x, y\rangle. \tag{4.22}$$

Proof. (a): As we already know $(H', \|\cdot\|)$ to be a Banach space, we merely need to check that (4.21) defines an inner product that induces the operator norm. If $0 \neq \alpha \in H'$, then $x := \psi^{-1}(\alpha) \neq 0$, i.e. $\langle\alpha, \alpha\rangle = \langle x, x\rangle > 0$. Next, let $\alpha, \beta, \gamma \in H'$ and $\lambda, \mu \in \mathbb{K}$. One computes

$$\begin{aligned}
\langle\lambda\alpha + \mu\beta, \gamma\rangle &= \big\langle\psi^{-1}(\gamma),\, \psi^{-1}(\lambda\alpha + \mu\beta)\big\rangle = \big\langle\psi^{-1}(\gamma),\, \overline{\lambda}\psi^{-1}(\alpha) + \overline{\mu}\psi^{-1}(\beta)\big\rangle \\
&= \lambda\big\langle\psi^{-1}(\gamma),\, \psi^{-1}(\alpha)\big\rangle + \mu\big\langle\psi^{-1}(\gamma),\, \psi^{-1}(\beta)\big\rangle \\
&= \lambda\langle\alpha, \gamma\rangle + \mu\langle\beta, \gamma\rangle.
\end{aligned}$$

Moreover,

$$\langle\alpha, \beta\rangle = \langle\psi^{-1}(\beta), \psi^{-1}(\alpha)\rangle = \overline{\langle\psi^{-1}(\alpha), \psi^{-1}(\beta)\rangle} = \overline{\langle\beta, \alpha\rangle},$$

proving (4.21) to define an inner product on H'. Moreover,

$$\underset{\alpha\in H'}{\forall} \quad \|\alpha\| = \|\psi^{-1}(\alpha)\| = \sqrt{\langle\psi^{-1}(\alpha),\psi^{-1}(\alpha)\rangle} = \sqrt{\langle\alpha,\alpha\rangle},$$

showing that the inner product defined by (4.21) induces the operator norm on H'.

(b): As in (a), let $\psi: H \longrightarrow H'$ be the map of (4.20). Let τ be the corresponding map on H', i.e.

$$\tau: H' \longrightarrow H'', \quad \tau(\alpha) := f_\alpha,$$

where

$$f_\alpha: H' \longrightarrow \mathbb{K}, \quad f_\alpha(\beta) = \langle\beta,\alpha\rangle = \langle\psi^{-1}(\alpha),\psi^{-1}(\beta)\rangle.$$

From Th. 4.21, we know ψ and τ to be surjective. Thus, if we can show that the canonical embedding satisfies

$$\Phi: H \longrightarrow H'', \quad \Phi = \tau\circ\psi, \tag{4.23}$$

then Φ is surjective and H reflexive. Indeed, we have

$$\underset{x\in H}{\forall} \quad \underset{\alpha\in H'}{\forall} \quad \Phi(x)(\alpha) = \alpha(x) = \langle x,\psi^{-1}(\alpha)\rangle = \langle\alpha,\psi(x)\rangle = (\tau(\psi(x)))(\alpha) = (\tau\circ\psi)(x)(\alpha),$$

proving (4.23) and H to be reflexive.

(c) follows by combining (3.17) with Th. 4.21. ■

Remark 4.23. Let $\big(H,\langle\cdot,\cdot\rangle_H\big)$ be a Hilbert space over $\mathbb{K}$. Let $\langle\cdot,\cdot\rangle_{H'}$ be the inner product on H' given by (4.21). Moreover, let $\langle\cdot,\cdot\rangle: H\times H' \longrightarrow \mathbb{K}$ denote the dual pairing according to (4.4). If $\psi: H \longrightarrow H'$ is the map of (4.20), then

$$\underset{x\in H}{\forall} \quad \underset{\alpha\in H'}{\forall} \quad \alpha(x) = \langle x,\alpha\rangle = \langle x,\psi^{-1}(\alpha)\rangle_H = \langle\alpha,\psi(x)\rangle_{H'}. \tag{4.24}$$

Definition 4.24. Let $\big(X,\langle\cdot,\cdot\rangle\big)$ be an inner product space over $\mathbb{K}$ and $S\subseteq X$. Then S is an *orthogonal system* if, and only if, $x\perp y$ for each $x,y\in S$ with $x\neq y$. Moreover, S is called an *orthonormal system* if, and only if, S is an orthogonal system consisting entirely of unit vectors (i.e. $S\subseteq S_1(0)$). Finally, S is called an *orthonormal basis* if, and only if, it is a *maximal* orthonormal system in the sense that, if $S\subseteq T\subseteq X$ and T is an orthonormal system, then $S=T$ (caveat: an orthonormal basis of X is not necessarily a vector space basis of X, see below).

Example 4.25. Consider the Hilbert space $H := L^2([0,2\pi],\mathcal{L}^1,\lambda^1)$. For each $n\in\mathbb{N}$, define

$$f_n: [0,2\pi] \longrightarrow \mathbb{R}, \quad f_n(t) := \frac{\sin nt}{\sqrt{\pi}}.$$

Then $S := \{f_n : n \in \mathbb{N}\}$ constitutes an orthonormal system in H: One computes, for each $m, n \in \mathbb{N}$,

$$\begin{aligned}
\langle f_n, f_n \rangle &= \frac{1}{\pi} \int_0^{2\pi} \sin^2 nt \, \mathrm{d}t \; = \frac{1}{\pi} \left[\frac{t}{2} - \frac{\sin nt \cos nt}{2n} \right]_0^{2\pi} = 1, \\
\langle f_m, f_n \rangle &= \frac{1}{\pi} \int_0^{2\pi} \sin mt \sin nt \, \mathrm{d}t \\
&= \frac{1}{\pi} \left[\frac{\sin mt \cos nt - \cos mt \sin nt}{2(m-n)} - \frac{\sin mt \cos nt - \cos mt \sin nt}{2(m+n)} \right]_0^{2\pi} \\
&= 0 \quad \text{for } m \neq n.
\end{aligned}$$

The set S from above is not an orthonormal basis (cf. [Wer11, Ex. V.4(a)]).

Theorem 4.26. *Let $\big(X, \langle \cdot, \cdot \rangle\big)$ be an inner product space over $\mathbb{K}$ with induced norm $\|\cdot\|$. Let $S \subseteq X$ be an orthonormal system.*

(a) Bessel Inequality*: If $S = \{e_n : n \in \mathbb{N}\}$, then*

$$\forall_{x \in X} \quad \sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2 \leq \|x\|^2. \tag{4.25}$$

(b) *If $S = \{e_n : n \in \mathbb{N}\}$, then*

$$\forall_{x,y \in X} \quad \sum_{n=1}^{\infty} |\langle x, e_n \rangle \langle e_n, y \rangle| < \infty.$$

(c) *If $S = \{e_n : n \in \mathbb{N}\}$, $x \in X$, and $x = \sum_{n=1}^{\infty} \lambda_n e_n$ with $\lambda_n \in \mathbb{K}$, then $\lambda_n = \langle x, e_n \rangle$ for each $n \in \mathbb{N}$ (i.e. the coefficients, called* Fourier coefficients[3]*, are uniquely determined by S and x).*

(d) *For each $x \in X$, the set $S_x := \{e \in S : \langle x, e \rangle \neq 0\}$ is finite or countable.*

Proof. (a): Let $x \in X$. For each $N \in \mathbb{N}$, define $x_N := x - \sum_{n=1}^{N} \langle x, e_n \rangle \, e_n$. Then

$$\forall_{k \in \{1,\dots,N\}} \quad \langle x_N, e_k \rangle = \langle x, e_k \rangle - \langle x, e_k \rangle \langle e_k, e_k \rangle = 0$$

[3]Originally, the term *Fourier coefficients* comes from the Hilbert space $H := L^2([0, 2\pi], \mathcal{L}^1, \lambda^1)$ of Ex. 4.25, using trigonometric functions such as the f_n of Ex. 4.25 (together with corresponding cosine functions) as orthonormal basis functions. For $f \in H$, the expansion of f into a series with respect to such basis vectors is a (traditional) Fourier series and the corresponding coefficients are (traditional) Fourier coefficients.

and, from Pythagoras' theorem [Phi16b, (1.48)],

$$\begin{aligned}\|x\|^2 &= \left\| x_N + \sum_{n=1}^{N} \langle x, e_n \rangle \, e_n \right\|^2 = \|x_N\|^2 + \left\| \sum_{n=1}^{N} \langle x, e_n \rangle \, e_n \right\|^2 = \|x_N\|^2 + \sum_{n=1}^{N} |\langle x, e_n \rangle|^2 \\ &\geq \sum_{n=1}^{N} |\langle x, e_n \rangle|^2.\end{aligned}$$

Letting $N \to \infty$ in the above inequality proves (4.25).

(b): Let $x, y \in X$. According to (a), the sequences $(\langle x, e_n \rangle)_{n\in\mathbb{N}}$, $(\langle y, e_n \rangle)_{n\in\mathbb{N}}$ are in l^2. Then, by Hölder's inequality, $(\langle x, e_n \rangle \langle e_n, y \rangle)_{n\in\mathbb{N}} \in l^1$, proving (b).

(c): If $x = \sum_{n=1}^{\infty} \lambda_n \, e_n$ with $\lambda_n \in \mathbb{K}$, then the orthonormality of the e_k implies

$$\underset{n\in\mathbb{N}}{\forall} \quad \langle x, e_n \rangle = \lambda_n \langle e_n, e_n \rangle = \lambda_n.$$

(d): Let $x \in X$. According to (a), for each $n \in \mathbb{N}$, the set $S_{x,n} := \{e \in S : |\langle x, e \rangle| \geq \frac{1}{n}\}$ must be finite. Thus, $S_x = \bigcup_{n\in\mathbb{N}} S_{x,n}$ must be finite or countable. ■

Example 4.27. Let $\big(H, \langle \cdot, \cdot \rangle\big)$ be a Hilbert space over $\mathbb{K}$ with induced norm $\| \cdot \|$. Let $\{e_n : n \in \mathbb{N}\} \subseteq H$ be an orthonormal system. If $\alpha \in H'$, then, according to the Riesz Representation Th. 4.21, there exists $y \in H$ such that $\alpha(x) = \langle x, y \rangle$ for each $x \in X$. Thus,

$$\underset{n\in\mathbb{N}}{\forall} \quad \alpha(e_n) = \langle e_n, y \rangle \overset{(4.25)}{\to} 0 = \alpha(0),$$

showing $e_n \rightharpoonup 0$. On the other hand $(e_n)_{n\in\mathbb{N}}$ does *not* converge strongly to 0 – actually, as

$$\begin{aligned}\|e_m - e_n\|^2 &= \langle e_m - e_n \, , \, e_m - e_n \rangle \\ &= \|e_m\|^2 - 2\,\mathrm{Re}\langle e_m, e_n \rangle + \|e_n\|^2 = 2 \quad \text{for each} \quad m \neq n,\end{aligned} \tag{4.26}$$

$(e_n)_{n\in\mathbb{N}}$ is not even a Cauchy sequence. A concrete example is given by the orthonormal sine functions f_n from Ex. 4.25.

Definition 4.28. Let $(X, \|\cdot\|)$ be a normed vector space over $\mathbb{K}$ (the following definition actually still makes sense if X is merely a topological space on which an addition with neutral element 0 is defined). Let $(x_i)_{i\in I}$ be a family in X. Then we say that the "series" $\sum_{i\in I} x_i$ *converges unconditionally* to $x \in X$ if, and only if, (i) and (ii) hold, where

(i) The set $I_0 := \{i \in I : x_i \neq 0\}$ is finite or countable.

(ii) For each enumeration $I_0 = \{i_1, i_2, \dots\}$ of I_0, one has $\sum_{n=1}^{\infty} x_{i_n} = x$ (where the sum must be replaced by a finite sum for $\#I_0 < \infty$), i.e. the result of $\sum_{n=1}^{\infty} x_{i_n}$ does not depend on the order of summation.

If $\sum_{i\in I} x_i$ converges unconditionally to $x \in X$, then we write $\sum_{i\in I} x_i = x$.

—

As a caveat it is pointed out that, in contrast to the situation on finite-dimensional spaces, on infinite-dimensional Banach spaces, the condition of absolute convergence is strictly stronger than the condition of unconditional convergence (cf. [Wer11, p. 235]).

Corollary 4.29 (Bessel Inequality). *Let $\big(X, \langle\cdot,\cdot\rangle\big)$ be an inner product space over $\mathbb{K}$ with induced norm $\|\cdot\|$. Let $S \subseteq X$ be an orthonormal system. Then*

$$\underset{x\in X}{\forall} \quad \sum_{e\in S} |\langle x, e\rangle|^2 \leq \|x\|^2, \tag{4.27}$$

where the convergence is unconditional in the sense of Def. 4.28.

Proof. Let $x \in X$. According to Th. 4.26(d), the set $S_x := \{e \in S : \langle x, e\rangle \neq 0\}$ is finite or countable. If S_x is finite, then (4.27) is clear. If S_x is infinite and $(e_n)_{n\in\mathbb{N}}$ is an enumeration of S_x, then $\sum_{n=1}^{\infty} |\langle x, e_n\rangle|^2$ converges absolutely by Th. 4.26(a), i.e. each rearrangement of the series converges to the same number, i.e. $\sum_{e\in S} |\langle x, e\rangle|^2$ converges unconditionally and (4.27) holds by (4.25). ■

Theorem 4.30. *Let $\big(H, \langle\cdot,\cdot\rangle\big)$ be a Hilbert space over $\mathbb{K}$ with induced norm $\|\cdot\|$. Let $S \subseteq X$ be an orthonormal system.*

(a) *For each $x \in H$, the series $\sum_{e\in S}\langle x, e\rangle\, e$ converges unconditionally.*

(b) *The map*

$$P : H \longrightarrow V := \overline{\operatorname{span} S}, \quad P(x) := \sum_{e\in S}\langle x, e\rangle\, e,$$

is the orthogonal projection onto V.

Proof. (a): Fix $x \in H$ and set $S_x := \{e \in S : \langle x, e\rangle \neq 0\}$. Then (a) is clear if S_x is finite. Thus, let S_x be infinite and let $(e_n)_{n\in\mathbb{N}}$ be an enumeration of S_x. We show that the partial sums of $\sum_{n=1}^{\infty}\langle x, e_n\rangle\, e_n$ form a Cauchy sequence in H: The orthonormality of the e_n implies

$$\underset{k\leq l}{\forall} \quad \left\|\sum_{n=k}^{l}\langle x, e_n\rangle\, e_n\right\|^2 = \sum_{n=k}^{l} |\langle x, e_n\rangle|^2,$$

and, since $\sum_{n=1}^{\infty} |\langle x, e_n\rangle|^2$ converges due to the Bessel inequality (4.25), the partial sums of $\sum_{n=1}^{\infty}\langle x, e_n\rangle\, e_n$ are Cauchy. Since H is complete, there exists a limit $y = \sum_{n=1}^{\infty}\langle x, e_n\rangle\, e_n \in H$. Analogously, if $\pi : \mathbb{N} \longrightarrow \mathbb{N}$ is a bijection, then there exists a limit for the rearranged series, $y_\pi = \sum_{n=1}^{\infty}\langle x, e_{\pi(n)}\rangle\, e_{\pi(n)} \in H$. Now

$$\underset{z\in H}{\forall} \quad \langle y, z\rangle = \sum_{n=1}^{\infty}\langle x, e_n\rangle\, \langle e_n, z\rangle \overset{(*)}{=} \sum_{n=1}^{\infty}\langle x, e_{\pi(n)}\rangle\, \langle e_{\pi(n)}, z\rangle = \langle y_\pi, z\rangle,$$

showing $y - y_\pi \in H^\perp = \{0\}$, i.e. $y = y_\pi$, proving (a) (at $(*)$, we used that we may rearrange the series, as it converges absolutely by Th. 4.26(b)).

(b): According to Th. 4.20(b), we have to show

$$\underset{v\in V}{\forall} \quad \underset{x\in H}{\forall} \quad \left\langle x - \sum_{n=1}^{\infty}\langle x, e_n\rangle\, e_n,\, v \right\rangle = 0, \tag{4.28}$$

where $(e_n)_{n\in\mathbb{N}}$ is an enumeration of $S_x = \{e \in S :\, \langle x, e\rangle \neq 0\}$, where (4.28) is equivalent to

$$\underset{e\in S}{\forall} \quad \underset{x\in H}{\forall} \quad F(e, x) := \left\langle x - \sum_{n=1}^{\infty}\langle x, e_n\rangle\, e_n,\, e \right\rangle = 0. \tag{4.29}$$

Fix $x \in H$. Since S forms an orthonormal system, if $e \in S \setminus S_x$, then $F(e, x) = 0 - 0 = 0$ and (4.29) is valid; if $e \in S_x$, then $F(e, x) = \langle x, e\rangle - \langle x, e\rangle \cdot 1 = 0$ and (4.29) is valid again. ■

Theorem 4.31 (Orthonormal Basis). *Let $\big(H, \langle\cdot,\cdot\rangle\big)$ be a Hilbert space over $\mathbb{K}$ with induced norm $\|\cdot\|$.*

(a) *If $S \subseteq H$ is an orthonormal system, then there exists an orthonormal basis B with $S \subseteq B \subseteq H$ (in particular, if $H \neq \{0\}$, then H has an orthonormal basis).*

(b) *Let $S \subseteq H$ be an orthonormal system. Then the following statements are equivalent:*

(i) *S is an orthonormal basis.*

(ii) *$S^\perp = \{0\}$.*

(iii) *$H = \overline{\operatorname{span} S}$.*

(iv) *$x = \sum_{e\in S}\langle x, e\rangle\, e$ holds for each $x \in H$.*

(v) *$\langle x, y\rangle = \sum_{e\in S}\langle x, e\rangle\, \langle e, y\rangle$ holds for each $x, y \in H$.*

(vi) *The* Parseval Identity *holds, i.e.*

$$\underset{x\in H}{\forall} \quad \|x\|^2 = \sum_{e\in S} |\langle x, e\rangle|^2.$$

(c) *Let $(K, \langle \cdot, \cdot \rangle)$ be another Hilbert space. Let $S \subseteq H$ be an orthonormal basis of H and let $T \subseteq K$ be an orthonormal basis of K. Then H and K are isometrically isomorphic if, and only if, $\# S = \# T$ (i.e. if, and only if, there exists a bijective map $\phi : S \longrightarrow T$).*

Proof. (a) follows from Zorn's lemma: Let

$$\mathcal{P} := \{T \subseteq H : S \subseteq T \text{ and } T \text{ is orthonormal system}\},$$

partially ordered by "$\subseteq$". Let $\mathcal{C} \subseteq \mathcal{P}$ be a chain and define $T_{\mathcal{C}} := \bigcup_{T \in \mathcal{C}} T$. Then $S \subseteq T_{\mathcal{C}}$ and, since $\mathcal{C}$ is a chain, $T_{\mathcal{C}}$ is an orthonormal system, showing $T_{\mathcal{C}} \in \mathcal{P}$. Clearly, $T_{\mathcal{C}}$ is an upper bound for $\mathcal{C}$. Thus, Zorn's lemma applies and $\mathcal{P}$ must have a maximal element, i.e. there exists an orthonormal basis containing S. As an aside, we remark that, if H is separable, then H has a countable orthonormal basis (see below), which can be constructed without using Zorn's lemma via Gram-Schmidt orthogonalization (cf. Th. E.1 in the Appendix).

(b): "(i)$\Rightarrow$(ii)": If there is $0 \neq x \in S^{\perp}$, then $S \dot\cup \{x/\|x\|\}$ is an orthonormal system, i.e. S is not maximal.

"(ii)$\Rightarrow$(iii)": Let $V := \overline{\operatorname{span} S}$. Then (ii) implies $V^{\perp} = \{0\}$. Thus, since V is a closed vector space,

$$V \overset{\text{Th. 4.20(f)}}{=} (V^{\perp})^{\perp} = \{0\}^{\perp} = H.$$

"(iii)$\Rightarrow$(iv)": If (iii) holds, then, by Th. 4.30(b), the map $x \mapsto \sum_{e \in S} \langle x, e \rangle \, e$ must be the identity on H.

"(iv)$\Rightarrow$(v)": Plug the formula for x, given by (iv), into $\langle x, y \rangle$ to obtain (v) (note the unconditional convergence due to Th. 4.26(b)).

"(v)$\Rightarrow$(vi)": Set $x = y$ in (v).

"(vi)$\Rightarrow$(i)": If S is not an orthonormal basis, then there exists $x \in H \setminus S$ such that $S \dot\cup \{x\}$ is an orthonormal system. Then $\|x\| = 1$, but $\sum_{e \in S} |\langle x, e \rangle|^2 = 0$, i.e. (vi) does not hold.

(c): First, assume $\phi : S \longrightarrow T$ to be bijective and define

$$\Phi : H \longrightarrow K, \quad x = \sum_{e \in S} \lambda_e \, e \mapsto \Phi(x) := \sum_{e \in S} \lambda_e \, \phi(e).$$

We need to check that Φ is well-defined, where one can argue analogous to the proof of Th. 4.30(a): Fix $x = \sum_{e \in S} \lambda_e \, e \in H$ and set $S_x := \{e \in S : \langle x, e \rangle \neq 0\}$. Note that

$$\underset{e \in S}{\forall} \quad \lambda_e = \langle x, e \rangle.$$

We need to show that $\sum_{e\in S}\lambda_e\,\phi(e)$ converges unconditionally to some $\Phi(x)\in K$. This is clear if S_x is finite. Thus, let S_x be infinite and let $(e_n)_{n\in\mathbb{N}}$ be an enumeration of S_x. We show that the partial sums of $\sum_{n=1}^{\infty}\lambda_{e_n}\,\phi(e_n)$ form a Cauchy sequence in K: The orthonormality of the $\phi(e_n)$ implies

$$\underset{k\leq l}{\forall}\quad \left\|\sum_{n=k}^{l}\lambda_{e_n}\,\phi(e_n)\right\|^2=\sum_{n=k}^{l}|\lambda_{e_n}|^2=\sum_{n=k}^{l}|\langle x,e_n\rangle|^2,$$

and, since $\sum_{n=1}^{\infty}|\langle x,e_n\rangle|^2$ converges due to the Bessel inequality (4.25), the partial sums of $\sum_{n=1}^{\infty}\lambda_{e_n}\,\phi(e_n)$ are Cauchy. Since K is complete, there exists a limit $y=\sum_{n=1}^{\infty}\lambda_{e_n}\,\phi(e_n)\in K$. But then, for each $n\in\mathbb{N}$, $\lambda_{e_n}=\langle y,\phi(e_n)\rangle$ must be the Fourier coefficient of y with respect to T, also implying unconditional convergence. Thus, Φ is well defined. If $\lambda,\mu\in\mathbb{K}$ and $x,y\in H$, then

$$\begin{aligned}\Phi(\lambda x+\mu y)&=\sum_{e\in S}\langle\lambda x+\mu y,\,e\rangle\,\phi(e)\\&=\lambda\sum_{e\in S}\langle x,e\rangle\,\phi(e)+\mu\sum_{e\in S}\langle y,\,e\rangle\,\phi(e)=\lambda\Phi(x)+\mu\Phi(y),\end{aligned}$$

showing Φ to be linear. Also

$$\underset{x,y\in H}{\forall}\quad \langle\Phi(x),\Phi(y)\rangle=\left\langle\sum_{e\in S}\langle x,e\rangle\,\phi(e),\,\sum_{e\in S}\langle y,e\rangle\,\phi(e)\right\rangle=\sum_{e\in S}\langle x,e\rangle\,\langle e,y\rangle=\langle x,y\rangle,$$

showing Φ to be isometric and injective. If

$$y=\sum_{e\in T}\langle y,e\rangle\,e\in K,\quad x:=\sum_{e\in T}\langle y,e\rangle\,\phi^{-1}(e)\in H,$$

then $\Phi(x)=y$, showing Φ to be surjective. Now, conversely, assume $\Phi:\,H\longrightarrow K$ to be an isometric isomorphism. As one can recover the inner product from the norm by Th. 4.14(d), we then also have

$$\underset{x,y\in H}{\forall}\quad \langle\Phi(x),\Phi(y)\rangle=\langle x,y\rangle.$$

Thus, if S is an orthonormal basis of H, then $\Phi(S)$ is an orthonormal basis of K. In consequence, we may now consider $K=H$, $\Phi=\mathrm{Id}$, i.e. it only remains to show that, if S,T are both orthonormal bases of H, then there exists a bijection $\phi:\,S\longrightarrow T$. If $\#S=n\in\mathbb{N}$, then $\dim H<\infty$, i.e. we know $\#T=n$ from Linear Algebra. Suppose S is infinite. For each $s\in S$, let $T_s:=\{t\in T:\,\langle s,t\rangle\neq 0\}$. Then each T_s is finite or countable and there exists a bijection $\psi:\,S\longrightarrow U:=\bigcup_{s\in S}T_s$. However, if $t\in T$, then $0\neq t$, i.e. $t\in U$. Thus, there exists an injective map $\phi_1:\,T\longrightarrow S$. One can now switch the roles of S and T to also obtain an injective map $\phi_2:\,S\longrightarrow T$. Then there exists a bijection $\phi:\,S\longrightarrow T$ by the Schröder-Bernstein theorem (cf. [Phi16a, Th. A.56]). ■

Corollary 4.32. *Let $\big(H, \langle\cdot,\cdot\rangle\big)$ be a Hilbert space over $\mathbb{K}$ with induced norm $\|\cdot\|$.*

(a) *If $S \subseteq H$ is an orthonormal basis, then $H \cong l^2(S)$, i.e. H isometrically isomorphic to $l^2(S)$, the space defined in Ex.* 4.15(b).

(b) *The following statements are equivalent for $H \neq \{0\}$:*

(i) *H is separable.*

(ii) *H has a finite or countable orthonormal basis.*

(iii) *Each orthonormal basis of H is finite or countable.*

(iv) *$H \cong \mathbb{K}^n$ with $n \in \mathbb{N}$ or $H \cong l^2$.*

(c) Riesz-Fischer*: $l^2 \cong L^2([0,1], \mathcal{L}^1, \lambda^1)$.*

Proof. (a) is immediate from Th. 4.31(c), since $T := \{\chi_{\{s\}} : s \in S\}$ forms an orthonormal basis of $l^2(S)$.

(b): The equivalence of (ii) – (iv) is immediate from Th. 4.31(c).

"(ii)⇒(i)": If S is an orthonormal basis of H, then, clearly, linear combinations of elements from S with coefficients from $\mathbb{Q}$ (for $\mathbb{K} = \mathbb{R}$) or from $\mathbb{Q} + i\mathbb{Q}$ (for $\mathbb{K} = \mathbb{C}$) are dense in H. Thus, if S is finite or countable, then H is separable.

"(i)⇒(ii)": If S is an orthonormal basis of H, then $\|e - f\| = \sqrt{2}$ for each $e, f \in S$ with $e \neq f$. Thus, if S is an uncountable orthonormal basis of H and $A \subseteq H$ is countable, then, for each $a \in A$, $\{x \in H : \|x - a\| < \frac{\sqrt{2}}{2}\}$ can contain at most one element of S (due to the triangle inequality). Thus, A can not be dense in H and H can not be separable.

(c) follows from (b), since l^2 and $L^2([0,1], \mathcal{L}^1, \lambda^1)$ both are infinite-dimensional Hilbert spaces and both separable by [Phi17, Th. 2.47(e)]. ■

Remark 4.33. Let $\big(H, \langle\cdot,\cdot\rangle\big)$ be a Hilbert space over $\mathbb{K}$ and let $B \subseteq H$ be an orthonormal basis. If $\dim H < \infty$, then B is a vector space basis[4] of H as well. We show that this *only* occurs for $\dim H < \infty$: If B is not finite and $x \in H$ is such that $B_x := \{e \in B : \langle x, e\rangle \neq 0\}$ is infinite (clearly, such $x \in H$ always exist), then $x \notin \operatorname{span} B$ (as the Fourier coefficients are unique). Thus, B can not be a vector space basis of H.

Definition 4.34. Let $\big(H_1, \langle\cdot,\cdot\rangle\big)$, $\big(H_2, \langle\cdot,\cdot\rangle\big)$ be Hilbert spaces over $\mathbb{K}$. Moreover, let $A \in \mathcal{L}(H_1, H_2)$, let $A' \in \mathcal{L}(H_2', H_1')$ be the adjoint operator according to Def. 4.9, and

[4]Sometimes called a *Hamel basis* in this context.

let $\psi_1 : H_1 \longrightarrow H_1'$, $\psi_2 : H_2 \longrightarrow H_2'$ be the maps given by the Riesz Representation Th. 4.21. Then the map

$$A^* : H_2 \longrightarrow H_1, \quad A^* := \psi_1^{-1} \circ A' \circ \psi_2, \tag{4.30}$$

is called the *Hilbert adjoint* of A.

Corollary 4.35. *Let $\big(H_1, \langle\cdot,\cdot\rangle\big)$, $\big(H_2, \langle\cdot,\cdot\rangle\big)$ be Hilbert spaces over $\mathbb{K}$, where $\|\cdot\|$ denotes the induced norms.*

(a) *For each $A \in \mathcal{L}(H_1, H_2)$, one has $A^* \in \mathcal{L}(H_2, H_1)$, and A^* is the unique map $H_2 \longrightarrow H_1$ such that*

$$\underset{x\in H_1}{\forall} \quad \underset{y\in H_2}{\forall} \quad \langle Ax, y\rangle = \langle x, A^*y\rangle. \tag{4.31}$$

(b) *One has that $A \mapsto A^*$ is a conjugate-linear isometric bijection of $\mathcal{L}(H_1, H_2)$ onto $\mathcal{L}(H_2, H_1)$.*

(c) $(\mathrm{Id}_{H_1})^* = \mathrm{Id}_{H_1}$.

(d) *If $\big(H_3, \langle\cdot,\cdot\rangle\big)$ is another Hilbert space over $\mathbb{K}$, $A \in \mathcal{L}(H_1, H_2)$, $B \in \mathcal{L}(H_2, H_3)$, then*

$$(B \circ A)^* = A^* \circ B^*.$$

(e) *One has*

$$\underset{A\in\mathcal{L}(H_1,H_2)}{\forall} \quad A^{**} = A.$$

(f) *If $A \in \mathcal{L}(H_1, H_2)$, then $\ker(A^*) = (A(H_1))^\perp$.*

(g) *If $A \in \mathcal{L}(H_1, H_2)$, then $A^{-1} \in \mathcal{L}(H_2, H_1)$ exists if, and only if, $(A^*)^{-1} \in \mathcal{L}(H_1, H_2)$ exists, and, in that case,*

$$(A^*)^{-1} = (A^{-1})^*.$$

(h) *$A \in \mathcal{L}(H_1, H_2)$ is isometric if, and only if, $A^* = A^{-1}$.*

Proof. (a): Let $A \in \mathcal{L}(H_1, H_2)$. Then $A^* \in \mathcal{L}(H_2, H_1)$, since each of the maps ψ_1^{-1}, A', ψ_2 is continuous, A' is linear, and ψ_1^{-1} and ψ_2 are both conjugate-linear. Moreover, we know A' is the unique map on H_2' such that

$$\underset{\beta\in H_2'}{\forall} \underset{x\in H_1}{\forall} \quad A'(\beta)(x) = \beta(A(x)).$$

Thus,

$$\underset{x\in H_1}{\forall} \quad \underset{y\in H_2}{\forall} \quad \langle x, A^*y\rangle = \langle x, (\psi_1^{-1} \circ A' \circ \psi_2)(y)\rangle = A'(\psi_2(y))(x) = \psi_2(y)(Ax) = \langle Ax, y\rangle,$$

proving (4.31). For each $y \in H_2$, $x \mapsto \langle Ax, y\rangle$ uniquely determines a continuous linear functional $\alpha_y : H_1 \longrightarrow \mathbb{K}$. Then the Riesz Representation Th. 4.21 and (4.31) imply $A^*(y) = \psi_1^{-1}(\alpha_y)$, showing A^* to be uniquely determined by (4.31).

(b): If $A, B \in \mathcal{L}(H_1, H_2)$ and $\lambda \in \mathbb{K}$, then, for each $y \in H_2$,

$$\begin{aligned}(A+B)^*(y) &= (\psi_1^{-1} \circ (A+B)' \circ \psi_2)(y) = (\psi_1^{-1} \circ (A' + B') \circ \psi_2)(y)\\ &= (\psi_1^{-1} \circ A' \circ \psi_2)(y) + (\psi_1^{-1} \circ B' \circ \psi_2)(y) = (A^* + B^*)(y)\end{aligned}$$

and

$$(\lambda A)^*(y) = (\psi_1^{-1} \circ (\lambda A)' \circ \psi_2)(y) = \overline{\lambda}(\psi_1^{-1} \circ A' \circ \psi_2)(y) = (\overline{\lambda} A^*)(y),$$

showing $A \mapsto A^*$ to be conjugate-linear. Moreover, $A \mapsto A^*$ is isometric, since the maps $\psi_1, \psi_2, A \mapsto A'$ all are isometric; $A \mapsto A^*$ is surjective, since H_1 and H_2 are reflexive.

(c): One has $(\mathrm{Id}_{H_1})^* = \psi_1^{-1} \circ (\mathrm{Id}_{H_1})' \circ \psi_1 = \psi_1^{-1} \circ \mathrm{Id}_{H_1'} \circ \psi_1 = \mathrm{Id}_{H_1}$.

(d): Let $\psi_3 : H_3 \longrightarrow H_3'$ be given by Th. 4.21. Then

$$A^* \circ B^* = \psi_1^{-1} \circ A' \circ \psi_2 \circ \psi_2^{-1} \circ B' \circ \psi_3 = \psi_1^{-1} \circ (B \circ A)' \circ \psi_3 = (B \circ A)^*.$$

(e): According to (a), A^{**} is the unique map $H_1 \longrightarrow H_2$ such that

$$\underset{x \in H_1}{\forall} \quad \underset{y \in H_2}{\forall} \quad \langle A^* y, x\rangle = \langle y, A^{**} x\rangle.$$

Comparing with (4.31) yields $A = A^{**}$.

(f): We have

$$y \in \ker(A^*) \quad \Leftrightarrow \quad \underset{x \in H_1}{\forall} \langle x, A^* y\rangle = 0 \quad \Leftrightarrow \quad \underset{x \in H_1}{\forall} \langle Ax, y\rangle = 0 \quad \Leftrightarrow \quad y \in (A(H_1))^{\perp}.$$

(g): One has

$$\begin{aligned} A^{-1} \in \mathcal{L}(H_2, H_1) \text{ exists} \quad &\Leftrightarrow \quad (A^{-1})' = (A')^{-1} \in \mathcal{L}(H_1', H_2') \text{ exists}\\ &\Leftrightarrow \quad (A^*)^{-1} = (\psi_1^{-1} \circ A' \circ \psi_2)^{-1} \in \mathcal{L}(H_1, H_2) \text{ exists.}\end{aligned}$$

Moreover, if $A^{-1} \in \mathcal{L}(H_2, H_1)$ exists, then

$$(A^*)^{-1} = \psi_2^{-1} \circ (A^{-1})' \circ \psi_1 = (A^{-1})^*.$$

(h): If A is isometric, then

$$\underset{x \in H_1}{\forall} \quad \underset{y \in H_2}{\forall} \quad \langle Ax, y\rangle = \langle Ax, AA^{-1} y\rangle = \langle x, A^{-1} y\rangle,$$

such that (a) implies $A^* = A^{-1}$. Conversely, if $A^* = A^{-1}$, then

$$\underset{u,v \in H_1}{\forall} \quad \langle Au, Av\rangle = \langle u, A^* A v\rangle = \langle u, v\rangle,$$

proving A to be isometric. ■

Example 4.36. (a) Let $m, n \in \mathbb{N}$, let X be $\mathbb{K}^n$ and Y be $\mathbb{K}^m$, each with the norm topology. Then $\mathcal{L}(X,Y) = \mathbb{K}^{mn}$ and each $A \in \mathcal{L}(X,Y)$ can be represented by and $m \times n$ matrix $A = (a_{kl})_{(k,l)\in\{1,\dots,m\}\times\{1,\dots,n\}}$. We know from Ex. 4.13(a) that the adjoint A' of A is represented by the transpose matrix A^{t}. Now consider $\mathbb{K}^n$ and $\mathbb{K}^m$ with the standard inner product. The map $\psi_1 : X \longrightarrow X'$ according to Th. 4.21 is given by

$$\psi_1 \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := (\overline{x}_1, \dots, \overline{x}_n) :$$

Indeed, if $\alpha = (\alpha_1, \dots, \alpha_n) \in X'$ and $x = (x_1, \dots, x_n)^{\mathrm{t}} \in X$, then

$$\alpha(x) = \sum_{k=1}^{n} \alpha_k x_k = \langle x, \overline{\alpha}^{\mathrm{t}} \rangle.$$

We claim that the Hilbert adjoint A^* of A is represented by the conjugate transpose matrix $A^* = (\overline{a}_{lk})_{(k,l)\in\{1,\dots,m\}\times\{1,\dots,n\}}$: If $\psi_2 : Y \longrightarrow Y'$ is according to Th. 4.21, then

$$\begin{aligned} \underset{y\in Y}{\forall} \quad A^* y &= (\psi_1^{-1} \circ A' \circ \psi_2)(y) = \psi_1^{-1}\big((\overline{y}_1, \dots, \overline{y}_m)A\big) \\ &= \psi_1^{-1}\left(\sum_{k=1}^{m} a_{k1}\overline{y}_k, \dots, \sum_{k=1}^{m} a_{kn}\overline{y}_k\right) = \left(\sum_{k=1}^{m} \overline{a}_{k1} y_k, \dots, \sum_{k=1}^{m} \overline{a}_{kn} y_k\right)^{\mathrm{t}}, \end{aligned}$$

showing A^* to be represented by $\overline{A}^{\mathrm{t}}$.

(b) As in Ex. 4.13(b), consider the left shift operator, but this time on the Hilbert space l^2:

$$A : l^2 \longrightarrow l^2, \quad A(x_1, x_2, \dots) := (x_2, x_3, \dots).$$

The Hilbert adjoint A^* of A is, once again, the right shift operator

$$A^* : l^2 \longrightarrow l^2, \quad A^*(y_1, y_2, \dots) := (0, y_1, y_2, \dots) :$$

Indeed,

$$\underset{y=(y_k)_{k\in\mathbb{N}}\in l^2}{\forall} \quad \underset{x=(x_k)_{k\in\mathbb{N}}\in l^2}{\forall} \quad \langle Ax, y \rangle = \sum_{k=1}^{\infty} \overline{y}_k x_{k+1} = \langle x, A^* y \rangle,$$

i.e. the right shift is the Hilbert adjoint by (4.31).

4.3 Complex Measures, Radon-Nikodym Theorem

To prove the Riesz representation theorem for L^p-spaces, which says that, for every measure space $(\Omega, \mathcal{A}, \mu)$, one has an isometric isomorphism $(L^p(\mu))' \cong L^q(\mu)$ if $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$, one needs the Radon-Nikodym theorem of measure theory, constituting itself an extremely important result. While we thoroughly studied $[0, \infty]$-valued measures in [Phi17], in preparation for the Radon-Nikodym theorem, we will now have to study so-called *complex* measures, which are $\mathbb{C}$-valued measures.

Definition 4.37. Let $(\Omega, \mathcal{A})$ be a measurable space. A map $\mu : \mathcal{A} \longrightarrow \mathbb{C}$ is called a *complex measure* on $(\Omega, \mathcal{A})$ if, and only if, μ satisfies the following conditions (i) and (ii):

(i) $\mu(\emptyset) = 0$.

(ii) μ is *countably additive* (also called *σ-additive*), i.e., if $(A_k)_{k\in\mathbb{N}}$ is a sequence in $\mathcal{A}$ consisting of (pairwise) disjoint sets, then

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k). \tag{4.32}$$

If μ is a complex measure on $(\Omega, \mathcal{A})$, then the triple $(\Omega, \mathcal{A}, \mu)$ is called a *complex measure space*. A *signed measure* is a complex measure that is $\mathbb{R}$-valued; the corresponding measure space is then also called a *signed measure space*. Note that (i) is separately stated for emphasis only, as it, clearly, follows from σ-additivity (as $\mu(\emptyset) = \infty$ is not allowed). Let $\mathcal{M}_{\mathbb{C}}(\Omega, \mathcal{A})$ denote the set of complex measures on $(\Omega, \mathcal{A})$, let $\mathcal{M}_{\mathbb{R}}(\Omega, \mathcal{A})$ denote the set of signed measures on $(\Omega, \mathcal{A})$.

Remark 4.38. In the present context, we will call the $[0, \infty]$-valued measures of [Phi17] *positive measures*. Thus, a positive measure is a complex measure if, and only if, it is finite; a complex measure is a positive measure if, and only if, it is $\mathbb{R}_0^+$-valued.

Lemma 4.39. *Let $\sum_{j=1}^{\infty} a_j$ be a series in $\mathbb{C}$. Then $\sum_{j=1}^{\infty} a_j$ converges absolutely if, and only if, both $\sum_{j=1}^{\infty} \operatorname{Re} a_j$ and $\sum_{j=1}^{\infty} \operatorname{Im} a_j$ converge absolutely.*

Proof. Recall from [Phi16a, Th. 5.9(d)] that

$$\underset{z\in\mathbb{C}}{\forall} \quad \max\{|\operatorname{Re} z|, |\operatorname{Im} z|\} \leq |z| \leq |\operatorname{Re} z| + |\operatorname{Im} z|.$$

Thus, if $\sum_{j=1}^{\infty} a_j$ converges absolutely, then, as $\sum_{j=1}^{\infty} |a_j|$ dominates both $\sum_{j=1}^{\infty} |\operatorname{Re} a_j|$ and $\sum_{j=1}^{\infty} |\operatorname{Im} a_j|$, these series converge absolutely as well. Conversely, if both series $\sum_{j=1}^{\infty} \operatorname{Re} a_j$ and $\sum_{j=1}^{\infty} \operatorname{Im} a_j$ converge absolutely, then $\sum_{j=1}^{\infty}(|\operatorname{Re} a_j| + |\operatorname{Im} a_j|)$ converges absolutely as well, implying absolute convergence of $\sum_{j=1}^{\infty} a_j$. ∎

Remark 4.40. We know from [Phi16a, Th. 7.95(a)] that, if the series $\sum_{j=1}^{\infty} a_j$ converges absolutely, then every rearrangement converges to the same limit. If $(\Omega, \mathcal{A}, \mu)$ is a complex measure space and $(A_k)_{k\in\mathbb{N}}$ is a sequence in $\mathcal{A}$ consisting of disjoint sets, then Def. 4.37(ii) implies that the convergence in (4.32) is *absolute*: Otherwise, by Lem. 4.39, $\sum_{k=1}^{\infty} \operatorname{Re}\mu(A_k)$ or $\sum_{k=1}^{\infty} \operatorname{Im}\mu(A_k)$ were to converge, but not converge absolutely and, then, Def. 4.37(ii) could not hold due to the Riemann rearrangement theorem [Phi16a, Th. 7.93].

Example 4.41. (a) Let $(\Omega, \mathcal{A})$ be a measurable space. If $\mu_1, \mu_2, \mu_3, \mu_4$ are finite postive measures on $(\Omega, \mathcal{A})$, then

$$\mu : \mathcal{A} \longrightarrow \mathbb{C}, \quad \mu := \mu_1 - \mu_2 + i(\mu_3 - \mu_4),$$

constitutes a complex measure on $(\Omega, \mathcal{A})$: The σ-additivity of μ is clear, since each μ_j is σ-additive and finite.

(b) Let $(\Omega, \mathcal{A}, \mu)$ be a positive measure space and let $f : \Omega \longrightarrow \mathbb{C}$ be integrable. Then

$$f\mu : \mathcal{A} \longrightarrow \mathbb{C}, \quad (f\mu)(A) := \int_A f \,\mathrm{d}\mu\,,$$

defines a complex measure on $(\Omega, \mathcal{A})$ in consequence of (a): Since

$$f = (\operatorname{Re} f)^+ - (\operatorname{Re} f)^- + i\big((\operatorname{Im} f)^+ - (\operatorname{Im} f)^-\big),$$

we know from [Phi17, Prop. 2.62], that each of the measures $(\operatorname{Re} f)^+\mu$, $(\operatorname{Re} f)^-\mu$, $(\operatorname{Im} f)^+\mu$, $(\operatorname{Im} f)^-\mu$ is positive, and each of the measures $(\operatorname{Re} f)^+\mu$, $(\operatorname{Re} f)^-\mu$, $(\operatorname{Im} f)^+\mu$, $(\operatorname{Im} f)^-\mu$ is finite, since f is integrable.

Definition 4.42. Let $(\Omega, \mathcal{A}, \mu)$ be a measure space (positive or complex). One calls the function

$$|\mu| : \mathcal{A} \longrightarrow [0, \infty],$$
$$|\mu|(A) := \sup\left\{\sum_{k=1}^{\infty} |\mu(A_k)| : (A_k)_{k\in\mathbb{N}} \text{ is disjoint sequence in } \mathcal{A} \text{ with } A = \bigcup_{k\in\mathbb{N}} A_k\right\},$$

the *total variation* of the measure μ.

Remark 4.43. (a) Clearly, if $(\Omega, \mathcal{A}, \mu)$ is a positive measure space, then $|\mu| = \mu$.

(b) Even thought the notation $|\mu|$ for the total variation of μ is customary, one has to use some care: While $|\mu|(A) \geq |\mu(A)|$ is clear from the definition of $|\mu|$, the inequality can be strict: For example, if $\Omega = \{1, 2\}$, $\mu(\{1\}) = 1$, $\mu(\{2\}) = -1$, then $|\mu(\Omega)| = 0$, but $|\mu|(\Omega) = 2$.

—

We will show in Th. 4.45 below that the total variation of a complex measure always constitutes a finite positive measure. In preparation, we provide the following lemma:

Lemma 4.44. *Let $N \in \mathbb{N}$ and $z_1, \dots, z_N \in \mathbb{C}$. Then there exists a set $J \subseteq \{1, \dots, N\}$ of indices such that*

$$\left| \sum_{k \in J} z_k \right| \geq \frac{1}{\pi} \sum_{k=1}^{N} |z_k|. \tag{4.33}$$

Proof. For each $k \in \{1, \dots, N\}$, choose $\alpha_k \in \mathbb{R}$ such that $z_k = |z_k| e^{i\alpha_k}$. For each $\theta \in [0, 2\pi]$, define $J(\theta) := \big\{k \in \{1, \dots, N\} : \cos(\alpha_k - \theta) > 0\big\}$ and, using $|e^{-i\theta}| = 1$, estimate

$$\left| \sum_{k \in J(\theta)} z_k \right| = \left| \sum_{k \in J(\theta)} e^{-i\theta} z_k \right| \geq \operatorname{Re} \sum_{k \in J(\theta)} e^{-i\theta} z_k = \sum_{k=1}^{N} |z_k| \cos^+(\alpha_k - \theta).$$

Now note that the right-hand side in the above inequality constitutes a continuous function of θ and, thus, must assume its maximum on the compact interval $[0, 2\pi]$, say, at $\theta_0 \in [0, 2\pi]$. Since

$$\int_0^{2\pi} \cos^+(\alpha_k - t)\,\mathrm{d}t = 2 \int_0^{\pi/2} \cos t\,\mathrm{d}t = 2[\sin t]_0^{\pi/2} = 2(1-0) = 2,$$

one obtains

$$\begin{aligned} \left| \sum_{k \in J(\theta_0)} z_k \right| &\geq \sum_{k=1}^{N} |z_k| \cos^+(\alpha_k - \theta_0) = \frac{1}{2\pi} \int_0^{2\pi} \left(\sum_{k=1}^{N} |z_k| \cos^+(\alpha_k - \theta_0) \right) \mathrm{d}t \\ &\geq \frac{1}{2\pi} \int_0^{2\pi} \left(\sum_{k=1}^{N} |z_k| \cos^+(\alpha_k - t) \right) \mathrm{d}t = \frac{1}{\pi} \sum_{k=1}^{N} |z_k|, \end{aligned}$$

thereby proving the lemma. ∎

Theorem 4.45. *Let $(\Omega, \mathcal{A}, \mu)$ be a complex measure space. Then the total variation $|\mu|$ of μ, as defined in Def. 4.42, constitutes a finite positive measure on $(\Omega, \mathcal{A})$.*

Proof. As $\mu(\emptyset) = 0$ implies $|\mu|(\emptyset) = 0$, to prove $|\mu|$ constitutes a positive measure, we have to show it is σ-additive. Thus, let $A \in \mathcal{A}$ and let $(A_k)_{k \in \mathbb{N}}$ be a disjoint sequence in $\mathcal{A}$ such that $A = \bigcup_{k \in \mathbb{N}} A_k$. If $t_k \in \mathbb{R}$ is such that $t_k < |\mu|(A_k)$, then, by the definition of $|\mu|$, there exists a disjoint sequence $(A_{kl})_{l \in \mathbb{N}}$ in $\mathcal{A}$ such that $A_k = \bigcup_{l \in \mathbb{N}} A_{kl}$

and $t_k < \sum_{l=1}^{\infty} |\mu(A_{kl})|$. Since $(A_{kl})_{(k,l)\in\mathbb{N}^2}$ is a countable disjoint family in $\mathcal{A}$ such that $A = \bigcup_{(k,l)\in\mathbb{N}^2} A_{kl}$, one obtains

$$\sum_{k=1}^{\infty} \max\{0, t_k\} \leq \sum_{(k,l)\in\mathbb{N}^2} |\mu(A_{kl})| \leq |\mu|(A),$$

implying

$$\sum_{k=1}^{\infty} |\mu|(A_k) = \sup\left\{\sum_{k=1}^{\infty} \max\{0, t_k\} : \underset{k\in\mathbb{N}}{\forall}\ t_k < |\mu|(A_k)\right\} \leq |\mu|(A).$$

To prove the opposite inequality, let $(E_l)_{l\in\mathbb{N}}$ also be a disjoint sequence in $\mathcal{A}$ such that $A = \bigcup_{l\in\mathbb{N}} E_l$. Then, for each $k \in \mathbb{N}$, $(A_k \cap E_l)_{l\in\mathbb{N}}$ is a disjoint sequence in $\mathcal{A}$ such that $A_k = \bigcup_{l\in\mathbb{N}}(A_k \cap E_l)$, and, for each $l \in \mathbb{N}$, $(A_k \cap E_l)_{k\in\mathbb{N}}$ is a disjoint sequence in $\mathcal{A}$ such that $E_l = \bigcup_{k\in\mathbb{N}}(A_k \cap E_l)$. Thus,

$$\sum_{l=1}^{\infty} |\mu(E_l)| = \sum_{l=1}^{\infty} \left|\sum_{k=1}^{\infty} \mu(A_k \cap E_l)\right| \leq \sum_{l=1}^{\infty}\sum_{k=1}^{\infty} |\mu(A_k \cap E_l)| \leq \sum_{k=1}^{\infty} |\mu|(A_k).$$

As we may take the sup on the left-hand side, we obtain $|\mu|(A) \leq \sum_{k=1}^{\infty} |\mu|(A_k)$, completing the proof of the σ-additivity of $|\mu|$ and showing $|\mu|$ to be a positive measure.

It remains to show $|\mu|$ is finite, which we will accomplish arguing by contraposition. Consider an arbitrary set $E \in \mathcal{A}$ such that $|\mu|(E) = \infty$ and set $t := \pi(1+|\mu(E)|)$. Then $t \in \mathbb{R}^+$ and there exists a disjoint sequence $(E_k)_{k\in\mathbb{N}}$ in $\mathcal{A}$ such that $E = \bigcup_{k\in\mathbb{N}} E_k$ and $t < \sum_{k=1}^{N} |\mu(E_k)|$ for some $N \in \mathbb{N}$. We apply Lem. 4.44 with $z_k := \mu(E_k)$ to obtain a set $J \subseteq \{1, \dots, N\}$ such that

$$\left|\mu\left(\bigcup_{k\in J} E_k\right)\right| = \left|\sum_{k\in J} \mu(E_k)\right| \geq \frac{1}{\pi}\sum_{k=1}^{N} |\mu(E_k)| > \frac{t}{\pi} \geq 1.$$

Letting $A := \bigcup_{k\in J} E_k$ and $B := E \setminus A$, this yields

$$|\mu(B)| = |\mu(E) - \mu(A)| \geq |\mu(A)| - |\mu(E)| > \frac{t}{\pi} - |\mu(E)| = 1.$$

Thus, we have decomposed E into disjoint sets A, B with $|\mu|(A) \geq |\mu(A)| > 1$ and $|\mu|(B) \geq |\mu(B)| > 1$. Since we already know $|\mu|$ to be a positive measure and $|\mu|(E) = \infty$, $|\mu|(A)$ or $|\mu|(B)$ must be infinite. If $|\mu|(\Omega) = \infty$, then we can now use an inductive construction to obtain a disjoint sequence $(A_k)_{k\in\mathbb{N}}$ in $\mathcal{A}$ such that $|\mu(A_k)| > 1$ for each $k \in \mathbb{N}$. But then $\sum_{k=1}^{\infty} \mu(A_k)$ does not converge absolutely and μ can not be a complex measure. ■

One of the advantages of complex (or signed) measures over positive measures is that they form vector spaces:

Remark 4.46. Let $(\Omega, \mathcal{A})$ be a measurable space. From the rules for convergent series in $\mathbb{K}$, it is immediate that $X := \mathcal{M}_{\mathbb{K}}(\Omega, \mathcal{A})$ forms a vector space over $\mathbb{K}$ (a subspace of the vector space $\mathbb{K}^{\mathcal{A}}$ of $\mathbb{K}$-valued functions on $\mathcal{A}$). We can make X into a normed space over $\mathbb{K}$ by defining

$$\|\cdot\| : X \longrightarrow \mathbb{R}_0^+, \quad \|\mu\| := |\mu|(\Omega) :$$

Let $\mu \in X$. Then $\mu = 0$ implies $\|\mu\| = 0$. If $\mu \neq 0$, then there exists $A \in \mathcal{A}$ with $\mu(A) \neq 0$ implying $\|\mu\| = |\mu|(\Omega) \geq |\mu(A)| > 0$. If $\lambda \in \mathbb{K}$ and $(A_k)_{k\in\mathbb{N}}$ is disjoint sequence in $\mathcal{A}$ with $\Omega = \bigcup_{k\in\mathbb{N}} A_k$, then

$$\sum_{k=1}^{\infty} |\lambda\mu(A_k)| = |\lambda| \sum_{k=1}^{\infty} |\mu(A_k)|,$$

showing $\|\lambda\mu\| = |\lambda\mu|(\Omega) = |\lambda|\,|\mu|(\Omega) = |\lambda|\,\|\mu\|$. If $\mu, \nu \in X$ and $(A_k)_{k\in\mathbb{N}}$ as before, then

$$\sum_{k=1}^{\infty} |(\mu+\nu)(A_k)| \leq \sum_{k=1}^{\infty} |\mu(A_k)| + \sum_{k=1}^{\infty} |\nu(A_k)|,$$

showing $\|\mu+\nu\| = |\mu+\nu|(\Omega) \leq |\mu|(\Omega) + |\nu|(\Omega) = \|\mu\| + \|\nu\|$.

Definition and Remark 4.47. (a) Let $(\Omega, \mathcal{A}, \mu)$ be a signed measure space. The measures

$$\mu^+ := \frac{1}{2}\big(|\mu| + \mu\big), \quad \mu^- := \frac{1}{2}\big(|\mu| - \mu\big)$$

are called the *positive variation* and *negative variation* of μ, respectively. Both measures μ^+, μ^- are actually positive, since $|\mu|(A) \geq |\mu(A)|$ for each $A \in \mathcal{A}$. Clearly,

$$\mu = \mu^+ - \mu^-, \quad |\mu| = \mu^+ + \mu^-,$$

where the decomposition $\mu = \mu^+ - \mu^-$ into the difference of the two positive measures μ^+, μ^- is known as the *Jordan decomposition* of μ.

(b) Let $(\Omega, \mathcal{A}, \mu)$ be a complex measure space. Then, clearly, both $\operatorname{Re}\mu$ and $\operatorname{Im}\mu$ are signed measures. Thus, by (a), we obtain the decomposition

$$\mu = (\operatorname{Re}\mu)^+ - (\operatorname{Re}\mu)^- + i\,\big((\operatorname{Im}\mu)^+ - (\operatorname{Im}\mu)^-\big).$$

Definition 4.48. Let $(\Omega, \mathcal{A})$ be a measurable space. Let λ, μ be measures on $(\Omega, \mathcal{A})$, where μ is positive and λ is arbitrary (i.e. positive or complex; note that a positive measure may be infinite).

(a) We call λ *absolutely continuous* with respect to μ (denoted $\lambda \ll \mu$) if, and only if,

$$\forall_{A\in\mathcal{A}} \quad \Big(\mu(A) = 0 \;\Rightarrow\; \lambda(A) = 0\Big).$$

(b) λ is said to be *concentrated* on $A \in \mathcal{A}$ if, and only if,

$$\forall_{E\in\mathcal{A}} \quad \lambda(E) = \lambda(E \cap A). \tag{4.34}$$

(c) If ν is another arbitrary measure on $(\Omega, \mathcal{A})$, then λ, ν are called *mutually singular* (denoted $\lambda \perp \nu$) if, and only if, there exist disjoint sets $A, B \in \mathcal{A}$ such that λ is concentrated on A and ν is concentrated on B.

Proposition 4.49. *Let $(\Omega, \mathcal{A})$ be a measurable space. Let $\alpha, \lambda, \mu, \nu$ be measures on $(\Omega, \mathcal{A})$, where μ is assumed to be positive.*

(a) *λ is concentrated on $A \in \mathcal{A}$ if, and only if,*

$$\forall_{E\in\mathcal{A}} \quad \Big(E \cap A = \emptyset \;\Rightarrow\; \lambda(E) = 0\Big). \tag{4.35}$$

(b) *If λ is concentrated on $A \in \mathcal{A}$, then so is $|\lambda|$.*

(c) *If $A, B \in \mathcal{A}$ and λ is concentrated on A as well as concentrated on B, then λ is concentrated on $A \cap B$.*

(d) *If $\lambda \perp \nu$, then $|\lambda| \perp |\nu|$.*

(e) *If α, λ are complex with $\alpha \perp \nu$ and $\lambda \perp \nu$, then $\alpha + \lambda \perp \nu$.*

(f) *If α, λ are complex with $\alpha \ll \mu$ and $\lambda \ll \mu$, then $\alpha + \lambda \ll \mu$.*

(g) *If $\lambda \ll \mu$, then $|\lambda| \ll \mu$.*

(h) *If $\alpha \ll \mu$ and $\lambda \perp \mu$, then $\alpha \perp \mu$.*

(i) *If $\lambda \ll \mu$ and $\lambda \perp \mu$, then $\lambda = 0$.*

Proof. (a): It is immediate that (4.34) implies (4.35). Conversely, assume (4.35). If $E \in \mathcal{A}$, then

$$\lambda(E) = \lambda(E \cap A) + \lambda(E \setminus A) \overset{(4.35)}{=} \lambda(E \cap A) + 0 = \lambda(E \cap A),$$

proving the validity of (4.34).

(b): We use (4.35): Let $E \in \mathcal{A}$ with $E \cap A = \emptyset$. If $(E_k)_{k\in\mathbb{N}}$ is a disjoint sequence in $\mathcal{A}$ such that $E = \bigcup_{k\in\mathbb{N}} E_k$, then $E_k \cap A = \emptyset$ and $\mu(E_k) = 0$ for each $k \in \mathbb{N}$, implying $|\mu|(E) = 0$.

(c): Let $E \in \mathcal{A}$. Then

$$\lambda(E) = \lambda(E \cap A) = \lambda(E \cap A \cap B),$$

showing λ to be concentrated on $A \cap B$.

(d) is immediate from (b).

(e): There exist disjoint $A_1, B_1 \in \mathcal{A}$ such that α is concentrated on A_1 and ν concentrated on B_1. Likewise, there exist disjoint $A_2, B_2 \in \mathcal{A}$ such that λ is concentrated on A_2 and ν concentrated on B_2. Then, by (c), ν is concentrated on $B_1 \cap B_2$. Moreover, for each $E \in \mathcal{A}$,

$$\begin{aligned}
(\alpha + \lambda)(E) &= \alpha(E \cap A_1) + \lambda(E \cap A_2) \\
&= \alpha\big(E \cap (A_1 \cup A_2) \cap A_1\big) + \lambda\big(E \cap (A_1 \cup A_2) \cap A_2\big) \\
&= (\alpha + \lambda)\big(E \cap (A_1 \cup A_2)\big),
\end{aligned}$$

showing $\alpha + \lambda$ to be concentrated on $A_1 \cup A_2$. Since $A_1 \cup A_2$ and $B_1 \cap B_2$ are disjoint, we have $\alpha + \lambda \perp \nu$.

(f) is clear.

(g): Let $E \in \mathcal{A}$ with $\mu(E) = 0$. If $(E_k)_{k\in\mathbb{N}}$ is a disjoint sequence in $\mathcal{A}$ such that $E = \bigcup_{k\in\mathbb{N}} E_k$, then $\mu(E_k) = \lambda(E_k) = 0$ for each $k \in \mathbb{N}$, implying $|\lambda|(E) = 0$ and $|\lambda| \ll \mu$.

(h): Assume $\alpha \ll \mu$ and $\lambda \perp \mu$. Due to $\lambda \perp \mu$, there exists $A \in \mathcal{A}$ such that λ is concentrated on A and $\mu(A) = 0$. Then $\alpha(E) = 0$ for each $E \in \mathcal{A}$ with $E \subseteq A$, implying

$$\underset{E\in\mathcal{A}}{\forall} \quad \alpha(E) = \alpha(E \cap A) + \alpha(E \cap A^c) = \alpha(E \cap A^c).$$

Thus, α is concentrated on A^c, showing $\alpha \perp \mu$.

(i): If $\lambda \ll \mu$ and $\lambda \perp \mu$, then (h) implies $\lambda \perp \lambda$. Thus, λ is concentrated on two disjoint sets. Then (c) implies λ to be concentrated on $\emptyset$, i.e. $\lambda = 0$. ■

Example 4.50. Let $(\Omega, \mathcal{A}, \mu)$ be a positive measure space and let $f : \Omega \longrightarrow \mathbb{C}$ be integrable. We noted in Ex. 4.41(b) that

$$f\mu : \mathcal{A} \longrightarrow \mathbb{C}, \quad (f\mu)(A) := \int_A f \,\mathrm{d}\mu\,,$$

defines a complex measure on $(\Omega, \mathcal{A})$. It is then immediate that $f\mu \ll \mu$. The Radon-Nikodym Th. 4.53(b) below will show that, for σ-finite μ, *every* complex measure that is absolutely continuous with respect to μ is obtained in this way.

—

In preparation for Th. 4.53, we need two more lemmas:

Lemma 4.51. *Let $(\Omega, \mathcal{A}, \mu)$ be a finite positive measure space, let $f \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$, and let $C \subseteq \mathbb{C}$ be closed. If*

$$\underset{A\in\mathcal{A}}{\forall} \quad \left(\mu(A) > 0 \;\Rightarrow\; z_A(f) := \frac{1}{\mu(A)} \int_A f \,\mathrm{d}\mu \,\in C\right), \tag{4.36}$$

then $f(x) \in C$ for μ-almost every $x \in \Omega$.

Proof. Let $z \in \mathbb{C} \setminus C$ and $r \in \mathbb{R}^+$ such that $B_r(z) \subseteq \mathbb{C} \setminus C$. Let $A := f^{-1}(B_r(z))$. We claim $\mu(A) = 0$: Indeed, if $\mu(A) > 0$, then

$$|z_A(f) - z| = \frac{1}{\mu(A)} \left| \int_A (f - z) \,\mathrm{d}\mu \right| \leq \frac{1}{\mu(A)} \int_A |f - z| \,\mathrm{d}\mu \,\leq r,$$

in contradiction to (4.36). Since there are sequences $(r_k)_{k\in\mathbb{N}}$ in $\mathbb{R}^+$ and $(z_k) \in \mathbb{C} \setminus C$ such that $\mathbb{C} \setminus C = \bigcup_{k\in\mathbb{N}} B_{r_k}(z_k)$, we obtain $\mu(f^{-1}(\mathbb{C} \setminus C)) = 0$, proving the lemma. ∎

Lemma 4.52. *Let $(\Omega, \mathcal{A}, \mu)$ be a positive σ-finite measure space.*

(a) *There exists $w \in \mathcal{L}^1(\mu)$ such that $0 < w < 1$.*

(b) *There exists a finite positive measure ν on $(\Omega, \mathcal{A})$ such that the set of μ-null sets is identical with the sets of ν-null sets.*

Proof. (a): As μ is σ-finite, there exists a sequence $(A_k)_{k\in\mathbb{N}}$ in $\mathcal{A}$ such that $\Omega = \bigcup_{k\in\mathbb{N}} A_k$ and $\mu(A_k) < \infty$ for each $k \in \mathbb{N}$. Define

$$\underset{k\in\mathbb{N}}{\forall} \quad w_k : \Omega \longrightarrow [0, 1[, \quad w_k := \frac{2^{-k}}{1 + \mu(A_k)} \chi_{A_k}$$

as well as

$$w : \Omega \longrightarrow]0, 1[, \quad w := \sum_{k=1}^{\infty} w_k.$$

Then, as a series of nonnegative measurable functions, w is measurable. Moreover, $0 < w$ is clear, whereas $w < 1$ is seen by estimating w from above by the the geometric series. Similarly,

$$\int_\Omega w \, \mathrm{d}\mu = \sum_{k=1}^\infty \int_\Omega w_k = \sum_{k=1}^\infty \frac{2^{-k}\, \mu(A_k)}{1+\mu(A_k)} < \sum_{k=1}^\infty 2^{-k} = 1,$$

showing w to be integrable.

(b): Let $w \in \mathcal{L}^1(\mu)$ be as in (a) and set $\nu := w\mu$. ∎

Theorem 4.53. *Let $(\Omega, \mathcal{A})$ be a measurable space. Let λ, μ be measures on $(\Omega, \mathcal{A})$, where μ is positive and σ-finite, and λ is complex.*

(a) Lebesgue Decomposition: *There exists a unique pair $(\lambda_{\mathrm{a}}, \lambda_{\mathrm{s}})$ of complex measures on $(\Omega, \mathcal{A})$ such that*

$$\lambda = \lambda_{\mathrm{a}} + \lambda_{\mathrm{s}} \quad \wedge \quad \lambda_{\mathrm{a}} \ll \mu \quad \wedge \quad \lambda_{\mathrm{s}} \perp \mu. \tag{4.37}$$

The pair $(\lambda_{\mathrm{a}}, \lambda_{\mathrm{s}})$ is called the Lebesgue decomposition *of λ relative to μ. Moreover, if λ is positive and finite, then so are $\lambda_{\mathrm{a}}, \lambda_{\mathrm{s}}$.*

(b) Radon-Nikodym: *There exists a unique $h \in L^1(\mu)$ such that*

$$\forall_{A \in \mathcal{A}} \quad \lambda_{\mathrm{a}}(A) = \int_A h \, \mathrm{d}\mu \tag{4.38}$$

(one then calls h the Radon-Nikodym derivative *of λ_{a} with respect to μ).*

Proof. We start with the uniqueness statements. Suppose $(\lambda'_{\mathrm{a}}, \lambda'_{\mathrm{s}})$ is another Lebesgue decomposition of λ relative to μ. Then

$$\lambda'_{\mathrm{a}} - \lambda_{\mathrm{a}} = \lambda'_{\mathrm{s}} - \lambda_{\mathrm{s}}, \qquad \lambda'_{\mathrm{a}} - \lambda_{\mathrm{a}} \overset{\text{Prop. 4.49(f)}}{\ll} \mu, \qquad \lambda'_{\mathrm{s}} - \lambda_{\mathrm{s}} \overset{\text{Prop. 4.49(e)}}{\perp} \mu,$$

and $\lambda'_{\mathrm{a}} = \lambda_{\mathrm{a}}$ as well as $\lambda'_{\mathrm{s}} = \lambda_{\mathrm{s}}$ follows from Prop. 4.49(i). The uniqueness of the Radon-Nikodym derivative $h \in L^1(\mu)$ is due to [Phi17, Th. 2.18(d)].

The following argument will show both the existence of the Lebesgue decomposition and the existence of the Radon-Nikodym derivative. First, assume λ to be positive as well. As μ is positive and σ-finite, we use Lem. 4.52(a) to obtain $w \in \mathcal{L}^1(\mu)$ such that $0 < w < 1$, $\nu := w\mu$. Then the measure $\varphi := \lambda + \nu$ is still positive and finite. If $f : \Omega \longrightarrow [0, \infty]$ is $\mathcal{A}$-measurable, then

$$\int_\Omega f \, \mathrm{d}\varphi = \int_\Omega f \, \mathrm{d}\lambda + \int_\Omega f \, \mathrm{d}\nu \tag{4.39}$$

(as (4.39) holds for $f = \chi_A$ with $A \in \mathcal{A}$, it then holds for simple f, and then also for nonnegative measurable f). For each $f \in L^2(\varphi)$, we know $f \in L^1(\varphi)$ (as φ is finite, cf. [Phi17, Th. 2.42]) and we estimate

$$\left|\int_\Omega f \,\mathrm{d}\lambda\right| \leq \int_\Omega |f| \,\mathrm{d}\lambda \leq \int_\Omega |f| \,\mathrm{d}\varphi \leq \left(\int_\Omega |f|^2 \,\mathrm{d}\varphi\right)^{\frac{1}{2}} \big(\varphi(\Omega)\big)^{\frac{1}{2}},$$

where the last estimate holds by the Cauchy-Schwarz inequality. Thus,

$$\alpha : L^2(\varphi) \longrightarrow \mathbb{K}, \quad \alpha(f) := \int_\Omega f \,\mathrm{d}\lambda\,,$$

defines a bounded linear functional on $L^2(\varphi)$. By the Riesz Representation Th. 4.21, there exists a unique $g \in L^2(\varphi)$ such that

$$\forall_{f \in L^2(\varphi)} \quad \alpha(f) = \int_\Omega f \,\mathrm{d}\lambda = \int_\Omega fg \,\mathrm{d}\varphi\,. \tag{4.40}$$

Instead of $g \in L^2(\varphi)$, we will now consider a representative $g \in \mathcal{L}^2(\varphi)$ (still denoted by g for simplicity of notation). From (4.40), we obtain

$$\forall_{A \in \mathcal{A}} \quad \lambda(A) = \int_\Omega \chi_A \,\mathrm{d}\lambda = \int_A g \,\mathrm{d}\varphi\,,$$

implying, since $0 \leq \lambda \leq \varphi$,

$$\forall_{A \in \mathcal{A}} \quad \left(\varphi(A) > 0 \;\Rightarrow\; 0 \leq \frac{1}{\varphi(A)} \int_A g \,\mathrm{d}\varphi = \frac{\lambda(A)}{\varphi(A)} \leq 1\right).$$

Thus, Lem. 4.51 yields $g(x) \in [0,1]$ for φ-almost every $x \in \Omega$. By changing g on a φ-null set, we see that there exists $g \in \mathcal{L}^2(\varphi)$ such that (4.40) holds and

$$\forall_{x \in \Omega} \quad 0 \leq g(x) \leq 1.$$

We can combine (4.39) and (4.40) to obtain

$$\forall_{f \in \mathcal{L}^2(\varphi)} \quad \int_\Omega (1-g) f \,\mathrm{d}\lambda = \int_\Omega fg \,\mathrm{d}\nu = \int_\Omega fgw \,\mathrm{d}\mu\,. \tag{4.41}$$

Set

$$E := \{x \in \Omega :\; 0 \leq g(x) < 1\}, \quad F := \{x \in \Omega :\; g(x) = 1\}$$

and define measures

$$\lambda_{\mathrm{a}} : \mathcal{A} \longrightarrow \mathbb{R}_0^+, \quad \lambda_{\mathrm{a}}(A) := \lambda(A \cap E), \tag{4.42a}$$

$$\lambda_{\mathrm{s}} : \mathcal{A} \longrightarrow \mathbb{R}_0^+, \quad \lambda_{\mathrm{s}}(A) := \lambda(A \cap F). \tag{4.42b}$$

Next, we use $f := \chi_F$ in (4.41) to conclude

$$0 = \int_F (1-g)\,\mathrm{d}\lambda = \int_F \mathrm{d}\nu = \nu(F) = \mu(F),$$

showing $\lambda_s \perp \mu$. Since φ is finite and g is bounded, for each $A \in \mathcal{A}$ and each $n \in \mathbb{N}$, we may apply (4.41) with

$$f := (1 + g + \cdots + g^n)\,\chi_A$$

to obtain

$$\forall_{n\in\mathbb{N}} \quad \int_A (1-g^{n+1})\,\mathrm{d}\lambda = \int_A g(1+g+\cdots+g^n)w\,\mathrm{d}\mu\,. \tag{4.43}$$

If $x \in F$, then $1 - g^{n+1}(x) = 0$ for each $n \in \mathbb{N}$; if $x \in E$, then $\lim_{n\to\infty}(1 - g^{n+1}(x)) = 1$. Thus, by the dominated convergence theorem [Phi17, Th. 2.20],

$$\lim_{n\to\infty} \int_A (1-g^{n+1})\,\mathrm{d}\lambda = \lambda(A \cap E) = \lambda_a(A). \tag{4.44}$$

As the nonnegatvie integrands on the right-hand side of (4.43) increase monotonically, we obtain a measurable pointwise limit

$$h := \lim_{n\to\infty} g(1+g+\cdots+g^n)w$$

and the monotone convergence theorem [Phi17, Th. 2.7] together with (4.43) and (4.44) implies

$$\lambda_a(A) = \lim_{n\to\infty} \int_A g(1+g+\cdots+g^n)w\,\mathrm{d}\mu = \int_A h\,\mathrm{d}\mu\,,$$

thereby proving (4.38). Since $\int_\Omega h\,\mathrm{d}\mu = \lambda(\Omega \cap E) < \infty$, we have $h \in L^1(\mu)$, concluding the proof of (b). Since (4.38) also implies $\lambda_a \ll \mu$, the proof of (a) is also complete. Finally, if λ is complex, then we decompose λ according to Def. and Rem. 4.47(b) and apply the positive case to $(\operatorname{Re}\lambda)^+$, $(\operatorname{Re}\lambda)^-$, $(\operatorname{Im}\lambda)^+$, and $(\operatorname{Im}\lambda)^-$, which, in combination with Prop. 4.49(e),(f), establishes the case. ■

4.4 L^p-Spaces, Riesz Representation Theorem II

Proposition 4.54. *Let $(\Omega, \mathcal{A}, \mu)$ be a positive measure space and $p, q \in [1, \infty]$ such that $\frac{1}{p} + \frac{1}{q} = 1$. For each $g \in L^q(\mu)$, define the map*

$$\alpha_g : L^p(\mu) \longrightarrow \mathbb{K}, \quad \alpha_g(f) := \int_\Omega fg\,\mathrm{d}\mu\,. \tag{4.45}$$

(a) *For each $g \in L^q(\mu)$, the map α_g is continuous and linear with $\|\alpha_g\| \le \|g\|_q$.*

(b) *If $1 < p < \infty$ or $p = 1$ and μ is σ-finite, then, for each $g \in L^q(\mu)$, one has $\|\alpha_g\| = \|g\|_q$.*

Proof. Let $g \in L^q(\mu)$.

(a): According to the Hölder inequality of [Phi17, Th. 2.7], $fg \in L^1(\mu)$ for $f \in L^p(\mu)$ and α_g is well-defined. The linearity of α_g is immediate from the linearity of the integral. Using the Hölder inequality again, we estimate

$$\forall_{f \in L^p(\mu)} \quad |\alpha_g(f)| = \left| \int_\Omega f g \,\mathrm{d}\mu \right| \leq \int_\Omega |fg| \,\mathrm{d}\mu \leq \|f\|_p \, \|g\|_q,$$

showing both the continuity of α_g and $\|\alpha_g\| \leq \|g\|_q$.

(b): In view of (a), it remains to show $\|\alpha_g\| \geq \|g\|_q$. Let $g \in \mathcal{L}^q(\mu)$ denote a representative of $g \in L^q(\mu)$. If $\|g\|_q = 0$, then there is nothing to prove. Thus, assume $\|g\|_q > 0$. Let $1 < p < \infty$. Define

$$f : \Omega \longrightarrow \mathbb{K}, \quad f(x) := \begin{cases} |g(x)|^{q-1} \, \frac{\overline{g(x)}}{|g(x)|} & \text{for } g(x) \neq 0, \\ 0 & \text{for } g(x) = 0. \end{cases}$$

Then f is measurable with

$$\int_\Omega |f|^p \,\mathrm{d}\mu = \int_\Omega |g|^{p(q-1)} \,\mathrm{d}\mu = \int_\Omega |g|^q \,\mathrm{d}\mu < \infty,$$

showing $f \in \mathcal{L}^p(\mu)$. Moreover,

$$\alpha_g(f) = \int_\Omega |g|^q \,\mathrm{d}\mu = \left(\int_\Omega |g|^q \,\mathrm{d}\mu \right)^{1/q} \left(\int_\Omega |g|^q \,\mathrm{d}\mu \right)^{1-1/q} = \|g\|_q \, \|f\|_p,$$

showing $\|\alpha_g\| \geq \|g\|_q$ as desired. Now let μ be σ-finite, $p = 1$, $q = \infty$. Recalling $\|g\|_\infty > 0$, we consider an arbitary $0 < s < \|g\|_\infty$. Then $\mu(\{|g| \geq s\}) > 0$ and, since μ is σ-finite,

$$\exists_{A_s \in \mathcal{A}} \quad \Big(A_s \subseteq \{|g| \geq s\} \ \wedge \ 0 < \mu(A_s) < \infty \Big).$$

Define

$$f_s : \Omega \longrightarrow \mathbb{K}, \quad f_s(x) := \begin{cases} \chi_{A_s}(x) \, \frac{\overline{g(x)}}{|g(x)|} & \text{for } g(x) \neq 0, \\ 0 & \text{for } g(x) = 0. \end{cases}$$

Then f_s is measurable with $\|f_s\|_1 = \int_\Omega |f_s| \,\mathrm{d}\mu = \mu(A_s) < \infty$, showing $f_s \in \mathcal{L}^1(\mu)$. Moreover,

$$\alpha_g(f_s) = \int_{A_s} |g| \,\mathrm{d}\mu \geq s \, \mu(A_s) = s \, \|f_s\|_1.$$

Since $s \in]0, \|g\|_\infty[$ was arbitrary, we obtain $\|\alpha_g\| \geq \|g\|_\infty$ also in this case. ∎

Theorem 4.55 (Riesz Representation Theorem). *Let $(\Omega, \mathcal{A}, \mu)$ be a positive measure space, $p \in [1, \infty[$, $q \in]1, \infty]$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Consider the map*

$$\varphi : L^q(\mu) \longrightarrow (L^p(\mu))', \quad \varphi(g) := \alpha_g, \tag{4.46}$$

where α_g is as in (4.45), i.e.

$$\alpha_g : L^p(\mu) \longrightarrow \mathbb{K}, \quad \alpha_g(f) := \int_\Omega f g \,\mathrm{d}\mu \,.$$

If $1 < p < \infty$ or $p = 1$ and μ is σ-finite, then φ is a (linear) isometric isomorphism.

Proof. The map φ is well-defined by Prop. 4.54(a) and its linearity is, once again, an immediate consequence of the linearity of the integral. Let $1 < p < \infty$ or $p = 1$ and μ is σ-finite. Then φ is isometric (and, thus, injective) by Prop. 4.54(b). In consequence, it only remains to show φ is surjective.

First, let μ be finite. Given $\alpha \in (L^p(\mu))'$, we will construct $g \in L^q(\mu)$ such that $\alpha_g = \alpha$, using the Radon-Nikodym Th. 4.53(b). If $A \in \mathcal{A}$, then $\mu(\Omega) < \infty$ implies $\chi_A \in L^p(\mu)$. Thus, we can define

$$\lambda : \mathcal{A} \longrightarrow \mathbb{K}, \quad \lambda(A) := \alpha(\chi_A).$$

We verify λ to be a complex measure: Let $(A_k)_{k \in \mathbb{N}}$ be a disjoint sequence in $\mathcal{A}$ and let $A := \bigcup_{k \in \mathbb{N}} A_k$. Then

$$\lim_{n\to\infty} \|\chi_{\bigcup_{k=1}^n A_k} - \chi_A\|_p^p = \lim_{n\to\infty} \mu\left(\bigcup_{k=n+1}^{\infty} A_k\right) = 0,$$

showing $\chi_{\bigcup_{k=1}^n A_k} \to \chi_A$ in $L^p(\mu)$. Thus, we can apply the linearity and continuity of α to obtain

$$\sum_{k=1}^{\infty} \lambda(A_k) = \lim_{n\to\infty} \sum_{k=1}^{n} \lambda(A_k) = \lim_{n\to\infty} \sum_{k=1}^{n} \alpha(\chi_{A_k}) \overset{\alpha \text{ lin.}}{=} \lim_{n\to\infty} \alpha(\chi_{\bigcup_{k=1}^n A_k}) = \alpha(\chi_A) = \lambda(A),$$

proving λ to be σ-additive and a complex measure. Next, we note $\lambda \ll \mu$: Indeed, if $A \in \mathcal{A}$ with $\mu(A) = 0$, then $\|\chi_A\|_p = 0$, i.e. $\lambda(A) = \alpha(\chi_A) = 0$. We now apply the Radon-Nikodym Th. 4.53(b) to obtain $g \in L^1(\mu)$, satisfying

$$\underset{A \in \mathcal{A}}{\forall} \quad \lambda(A) = \int_A g \,\mathrm{d}\mu \,.$$

We need to show $g \in L^q(\mu)$ and $\alpha = \alpha_g$, i.e.

$$\alpha(f) = \int_\Omega f g \,\mathrm{d}\mu \, = \alpha_g(f) \tag{4.47}$$

holds for each $f \in L^p(\mu)$. Due to the continuity of α and α_g, it suffices to show they agree on a dense subset of $L^p(\mu)$. From [Phi17, Th. 2.47(a)], we know the set of simple functions

$$\mathcal{S} := \operatorname{span}\{\chi_A :\ A \in \mathcal{A}\}$$

to be dense in $L^p(\mu)$ for $p < \infty$ (as μ is finite). Since $g \in L^1(\mu)$, so far, we know $\alpha_g(f)$ to be defined for each $f \in L^\infty(\mu)$ and, in particular, for each $f \in \mathcal{S}$ (as μ is finite, which also implies $L^\infty(\mu) \subseteq L^p(\mu)$). If $A \in \mathcal{A}$, then

$$\alpha(\chi_A) = \lambda(A) = \int_A g \,\mathrm{d}\mu = \alpha_g(\chi_A),$$

i.e. (4.47) holds for each $f = \chi_A$, $A \in \mathcal{A}$, and, in consequence, for each $f \in \mathcal{S}$. It remains to show $g \in L^q(\mu)$. To this end, it will be useful to verify (4.47) holds for each $f \in L^\infty(\mu)$. Let $f \in \mathcal{L}^\infty(\mu)$ be a representative that is bounded everywhere by $\|f\|_\infty$. Then, by [Phi17, Th. 1.90], there exists a sequence $(\phi_k)_{k\in\mathbb{N}}$ in $\mathcal{S}$ that satisfies $\|\phi_k\|_\infty \leq \|f\|_\infty$ for each $k \in \mathbb{N}$ and converges *uniformly* to f (i.e. $\phi_k \to f$ both in $L^\infty(\mu)$ and pointwise). Again using μ to be finite, $\lim_{k\to\infty} \|\phi_k - f\|_p = \lim_{k\to\infty} \|\phi_k - f\|_\infty = 0$, implying $\lim_{k\to\infty} \alpha(\phi_k) = \alpha(f)$. Since $\phi_k g \to fg$ pointwise with $|\phi_k g| \leq \|f\|_\infty |g|$, we can apply dominated convergence to obtain

$$\alpha(f) = \lim_{k\to\infty} \alpha(\phi_k) = \lim_{k\to\infty} \alpha_g(\phi_k) = \lim_{k\to\infty} \int_\Omega \phi_k g \,\mathrm{d}\mu = \int_\Omega fg \,\mathrm{d}\mu = \alpha_g(f),$$

proving (4.47) to hold for each $f \in L^\infty(\mu)$. For the verification that $g \in L^q(\mu)$, let $g \in \mathcal{L}^1(\mu)$ be a representative. Consider $q = \infty$. For each $A \in \mathcal{A}$, define

$$f_A :\ \Omega \longrightarrow \mathbb{K}, \quad f_A(x) := \begin{cases} \chi_A(x)\, \frac{\overline{g(x)}}{|g(x)|} & \text{for } g(x) \neq 0, \\ \chi_A(x) & \text{for } g(x) = 0. \end{cases}$$

Then $f_A \in \mathcal{L}^\infty(\mu)$ and we use (4.47) to obtain

$$\int_A |g| \,\mathrm{d}\mu = \int_\Omega f_A g \,\mathrm{d}\mu = \alpha(f_A) \leq \|\alpha\| \, \|f_A\|_1 = \int_A \|\alpha\| \,\mathrm{d}\mu\,.$$

Then [Phi17, Th. 2.18(d)] yields $|g| \leq \|\alpha\|$ μ-almost everywhere, in particular, $g \in L^\infty(\mu)$. Consider $1 < q < \infty$. For each $s \in \mathbb{R}^+$, let $A_s := \{|g| \leq s\}$ and define

$$f_s :\ \Omega \longrightarrow \mathbb{K}, \quad f_s(x) := \begin{cases} \chi_{A_s}(x)\, |g(x)|^{q-1}\, \frac{\overline{g(x)}}{|g(x)|} & \text{for } g(x) \neq 0, \\ \chi_{A_s}(x)\, |g(x)|^{q-1} & \text{for } g(x) = 0. \end{cases}$$

Then $f_s \in \mathcal{L}^\infty(\mu)$ and we use (4.47) to obtain

$$\int_{A_s} |g|^q \,\mathrm{d}\mu = \int_\Omega f_s g \,\mathrm{d}\mu = \alpha(f_s) \leq \|\alpha\| \, \|f_s\|_p = \|\alpha\| \left(\int_{A_s} |g|^q \,\mathrm{d}\mu \right)^{1/p}.$$

Thus,

$$\underset{s\in\mathbb{R}^+}{\forall} \quad \left(\int_{A_s} |g|^q \,\mathrm{d}\mu\right)^{1/q} \leq \|\alpha\|.$$

If we take the limit $s \to \infty$ (by use of the monotone convergence theorem), we obtain $\|g\|_q \leq \|\alpha\|$, in particular, $g \in L^q(\mu)$.

Second, let μ be σ-finite. As before, given $\alpha \in (L^p(\mu))'$, we need to find $g \in L^q(\mu)$ such that $\alpha_g = \alpha$. Using Lem. 4.52(a), choose $w \in \mathcal{L}^1(\mu)$ such that $0 < w < 1$, $\nu := w\mu$. Define

$$\tilde{\alpha} : L^p(\nu) \longrightarrow \mathbb{K}, \quad \tilde{\alpha}(F) := \alpha(w^{1/p}F).$$

Clearly, $\tilde{\alpha}$ is well-defined, linear, and bounded, i.e. $\tilde{\alpha} \in (L^p(\nu))'$. As ν is finite, we already know there exists $G \in L^q(\nu)$ such that

$$\underset{F\in L^p(\nu)}{\forall} \quad \tilde{\alpha}(F) = \int_\Omega FG \,\mathrm{d}\nu = \int_\Omega FGw \,\mathrm{d}\mu\,.$$

Set $g := w^{1/q}G$ (i.e. $g := G$ for $q = \infty$). Then

$$\begin{aligned} &\underset{1<q<\infty}{\forall} \quad \int_\Omega |g|^q \,\mathrm{d}\mu = \int_\Omega |G|^q \,\mathrm{d}\nu = \|G\|_q < \infty \quad &&\Rightarrow \quad g \in L^q(\mu),\\ &\|g\|_\infty = \|G\|_\infty < \infty &&\Rightarrow \quad g \in L^\infty(\mu). \end{aligned}$$

Since

$$\underset{f\in L^p(\mu)}{\forall} \quad \begin{aligned} \alpha(f) = \tilde{\alpha}(w^{-1/p}f) &= \int_\Omega w^{-1/p} fG \,\mathrm{d}\nu = \int_\Omega w^{-1/p}w^{-1/q} fg \,\mathrm{d}\nu\\ &= \int_\Omega w^{-1} fgw \,\mathrm{d}\mu = \alpha_g(f), \end{aligned}$$

we have established the case.

Third, and last, let μ be arbitrary, $1 < p < \infty$. Once again, given $\alpha \in (L^p(\mu))'$, we need to find $g \in L^q(\mu)$ such that $\alpha_g = \alpha$. For each $A \in \mathcal{A}$, we can restrict μ to $\mathcal{A}|A$, obtaining $\mu_A := \mu\restriction_{\mathcal{A}|A}$, and we can consider $L^p_A := L^p(A, \mathcal{A}|A, \mu_A)$ as a subspace of $L^p(\mu)$ by extending $f \in L^p_A$ by 0 to A^c. Let $\alpha_A := \alpha\restriction_{L^p_A}$. Define

$$\mathcal{B} := \{A \in \mathcal{A} : \mu_A \text{ is } \sigma\text{-finite}\}.$$

From the σ-finite case, for each $B \in \mathcal{B}$, there exists a unique $g_B \in L^q_B$ such that $\|g_B\|_q = \|\alpha_B\|$ and

$$\underset{f\in L^p_B}{\forall} \quad \alpha_B(f) = \int_B fg_B \,\mathrm{d}\mu\,. \tag{4.48}$$

Note that, if $A \in \mathcal{A}$, then restriction $f \mapsto f\restriction_A$ is well-defined as a map from $L^r(\mu)$ to L^r_A: If $f_1, f_2 \in \mathcal{L}^r(\mu)$ both are representatives of f, then $f_1\restriction_A$ and $f_2\restriction_A$ are representatives of the same element of $L^r(\mu)$, which we define to be $f\restriction_A$. Now, if $B_1, B_2 \in \mathcal{B}$ are disjoint, then the uniqueness of g_{B_1} and g_{B_2} implies $g_{B_1} = g_{B_1\cup B_2}\restriction_{B_1}$, $g_{B_2} = g_{B_1\cup B_2}\restriction_{B_2}$. In consequence (here we use $1 < q < \infty$),

$$\|\alpha_{B_1\cup B_2}\|^q = \|g_{B_1\cup B_2}\|_q^q = \|g_{B_1}\|_q^q + \|g_{B_2}\|_q^q = \|\alpha_{B_1}\|^q + \|\alpha_{B_2}\|^q. \tag{4.49}$$

Since

$$\sigma := \sup\{\|\alpha_B\| : B \in \mathcal{B}\} \leq \|\alpha\| < \infty,$$

there exists a sequence $(B_k)_{k\in\mathbb{N}}$ in $\mathcal{B}$ such that $\sigma = \lim_{k\to\infty} \|\alpha_{B_k}\|$. Let $B := \bigcup_{k\in\mathbb{N}} B_k$. Clearly, $B \in \mathcal{B}$. Since (4.49) implies

$$\underset{C,D\in\mathcal{B}}{\forall} \quad \Big(C \subseteq D \ \Rightarrow\ \|\alpha_C\| \leq \|\alpha_D\|\Big),$$

we conclude $\sigma = \|\alpha_B\|$. Moreover, as $B \in \mathcal{B}$, there exists $g_B \in L^q_B$ such that(4.48) holds. Note that $\sigma = \|\alpha_B\|$ together with (4.49) implies

$$\underset{C\in\mathcal{B}}{\forall} \quad \alpha_{C\setminus B} \equiv 0.$$

Now let $f \in \mathcal{L}^p(\mu)$, $C := \{f \neq 0\}$. Given $n \in \mathbb{N}$, define $C_n := \{|f| > \frac{1}{n}\}$. Then

$$\mu(C_n) = \int_{C_n} \frac{n^p}{n^p}\,\mathrm{d}\mu \leq \int_\Omega n^p|f|^p\,\mathrm{d}\mu < \infty,$$

showing $C_n \in \mathcal{B}$. Since $C = \bigcup_{n\in\mathbb{N}} C_n$, we have $C \in \mathcal{B}$ as well. Thus,

$$\alpha(\chi_{B^c} f) = \alpha(\chi_{B^c\setminus C} f) + \alpha(\chi_{C\cap B^c} f) = 0 + \alpha_{C\setminus B}(f\restriction_{C\setminus B}) = 0,$$

and

$$\alpha(f) = \alpha(\chi_B f) + \alpha(\chi_{B^c} f) = \alpha_B(f\restriction_B) = \int_B f g_B\,\mathrm{d}\mu = \int_\Omega f g_B\,\mathrm{d}\mu = \alpha_{g_B}(f)$$

completes the proof. ■

Example 4.56. Let $(\Omega, \mathcal{A}, \mu)$ be a positive finite measure space. Consider the continuous linear function

$$A : \mathbb{K} \longrightarrow L^1(\mu), \quad A(s) := f_s \equiv s.$$

We want to find the adjoint $A' : (L^1(\mu))' \longrightarrow \mathbb{K}$. From the Riesz Representation Th. 4.55, we know $(L^1(\mu))' \cong L^\infty(\mu)$, and we know the adjoint is uniquely determined by the condition

$$\underset{g\in L^\infty(\mu)}{\forall} \quad \underset{s\in\mathbb{K}}{\forall} \quad A'(g)\cdot s = A'(g)(s) = g(A(s)) = s\int_\Omega g\,\mathrm{d}\mu\,. \tag{4.50}$$

Thus, $A'(g) = \int_\Omega g\,\mathrm{d}\mu$ for each $g \in L^\infty(\mu)$.

4.5 Borel Measures on Locally Compact Hausdorff Spaces and Riesz Representation Theorems III and IV

Recall from [Phi17, Def. 2.48(a)] that, given a topological space $(X, \mathcal{T})$ and a function $f : X \longrightarrow \mathbb{K}$, the *support* of f is defined by

$$\operatorname{supp} f := \overline{\{x \in X : f(x) \neq 0\}},$$

and that $C_c(X)$ denotes the set of continuous functions from X into $\mathbb{K}$ with *compact* support.

Definition 4.57. Let $(X, \mathcal{T})$ be a topological space. Define

$$\begin{aligned} C^+(X) &:= \{f \in C(X) : f \text{ is } \mathbb{R}_0^+\text{-valued}\}, \\ C_c^+(X) &:= \{f \in C_c(X) : f \text{ is } \mathbb{R}_0^+\text{-valued}\}. \end{aligned}$$

We call a function $F : C(X) \longrightarrow \mathbb{K}$ (resp. a function $F : C_c(X) \longrightarrow \mathbb{K}$) *positive* if, and only if, $F(f) \in \mathbb{R}_0^+$ for each $f \in C^+(X)$ (resp. for each $f \in C_c^+(X)$.

Remark 4.58. In the situation of Def. 4.57, let F be positive and linear.

(a) F is *monotone* in the sense that $g \leq f$ implies $F(g) \leq F(f)$: Indeed, if $g \leq f$, then $f - g \geq 0$, implying $F(f) - F(g) = F(f - g) \geq 0$ and $F(g) \leq F(f)$.

(b) If $-F$ is also positive, then $F \equiv 0$: If $f \geq 0$, then $F(f) \geq 0$ and $-F(f) \geq 0$, showing $F(f) = 0$. Thus, for arbitrary f: $F(f) = F(f^+ - f^-) = F(f^+) - F(f^-) = 0$, proving $F \equiv 0$.

Definition 4.59. Let $(X, \mathcal{T})$ be a topological space and let $\mathcal{B} := \sigma(\mathcal{T})$ denote the corresponding Borel sets. Let $(X, \mathcal{A}, \mu)$ be a measure space (positive or complex).

(a) μ is called a *Borel measure* on X if, and only if, $\mathcal{B} \subseteq \mathcal{A}$.

(b) If μ is a positive Borel measure, then it is called *locally finite* if, and only if, for each $x \in X$, there exists an open neighborhood U of x such that $\mu(U) < \infty$.

Lemma 4.60. *Let $(X, \mathcal{T})$ be a T_2 topological space and let $(X, \mathcal{A}, \mu)$ be a positive measure space. Assume μ to be a locally finite Borel measure.*

(a) *If $K \subseteq X$ is compact, then $\mu(K) < \infty$.*

(b) $C_c(X) \subseteq \mathcal{L}^1(\mu)$.

Proof. (a): Since $(X, \mathcal{T})$ is T_2, the compact set K is closed, i.e. $K \in \mathcal{A}$ and $\mu(K)$ is defined. Since μ is locally finite, for each $x \in K$, there exists an open neighborhood U_x of x such that $\mu(U_x) < \infty$. Since K is compact, there exist finitely many $x_1, \dots, x_n$, $n \in \mathbb{N}$, such that $K \subseteq \bigcup_{i=1}^n U_{x_i}$. Thus, $\mu(K) \subseteq \sum_{i=1}^n \mu(U_{x_i}) < \infty$.

(b): If $f \in C_c(X)$, then

$$\int_X |f| \, d\mu \leq \|f\|_\infty \, \mu(\operatorname{supp} f) \overset{(a)}{<} \infty,$$

proving $f \in \mathcal{L}^1(\mu)$. ∎

Example 4.61. Let $(X, \mathcal{T})$ be a T_2 topological space and let $(X, \mathcal{A}, \mu)$ be a positive locally finite Borel measure space. Then $C_c(X) \subseteq \mathcal{L}^1(\mu)$ by Lem. 4.60(b) and we can define

$$\alpha : C_c(X) \longrightarrow \mathbb{K}, \quad \alpha(f) := \int_X f \, d\mu \,. \tag{4.51}$$

It is then clear from the properties of the integral that (4.51) defines a linear functional that is positive in the sense of Def. 4.57.

—

Example 4.61 now raises the question if *every* positive linear functional on $C_c(X)$ can be written in the form (4.51) with a suitable Borel measure μ (for the time being, there is no continuity involved). This turns out to be a difficult question in general and there are several subtleties. We will show in Th. 4.63 below that the answer is positive if $(X, \mathcal{T})$ is a locally compact T_2 space (one can actually even obtain such a representation for the larger class of so-called *completely regular* spaces, cf. [Els07, Sec. VIII.2]). It also turns out that, in general, a functional can be represented by several different Borel measures. To select a specific measure, it is customary to impose regularity properties. Many different variants can be found in the literature (again, we refer to [Els07, Sec. VIII.2] and references therein). Here, we will mostly follow [Rud87, Th. 2.14]. The following Prop. 4.62 will be employed in the proof of Th. 4.63.

Proposition 4.62. *Let the topological space $(X, \mathcal{T})$ be locally compact and T_2. If $O, K \subseteq X$ such that O is open, K is compact, and $K \subseteq O$, then*

$$\underset{f \in C_c(X)}{\exists} \Big(0 \leq f \leq 1 \ \wedge \ f{\restriction}_K \equiv 1 \ \wedge \ \operatorname{supp} f \subseteq O \Big).$$

Proof. According to Prop. 2.5(a), there exists an open set $V \subseteq X$ such that $\overline{V}$ is compact and

$$K \subseteq V \subseteq \overline{V} \subseteq O.$$

Since $(X,\mathcal{T})$ is T_2, K is closed. As a compact T_2 space, $(\overline{V},\mathcal{T}_{\overline{V}})$ is normal (in particular, T_4) by [Phi16b, Prop. 3.30]. Thus, we can apply the Tietze-Urysohn theorem [Phi16b, Th. 3.11] to the closed disjoint sets K and $A := \overline{V} \setminus V$. If $K = \emptyset$, then we set $f :\equiv 0$. Now assume $K \neq \emptyset$. If $A \neq \emptyset$, then [Phi16b, Th. 3.11] provides a continuous map $f : \overline{V} \longrightarrow [0,1]$ with $f{\restriction}_A \equiv 0$ and $f{\restriction}_K \equiv 1$. If we extend f to all of X by setting $f{\restriction}_{X\setminus\overline{V}} :\equiv 0$, then f is still continuous. Also

$$\operatorname{supp} f \subseteq \overline{V} \subseteq O \tag{4.52}$$

with $\overline{V}$ compact. In particular, $\operatorname{supp} f$ is itself compact and f satisfies all required conditions. If $A = \emptyset$, then $V = \overline{V}$ and V is both open and closed. In this case, we define

$$f : X \longrightarrow [0,1], \quad f(x) := \begin{cases} 1 & \text{for } x \in V, \\ 0 & \text{for } x \notin V. \end{cases}$$

Then the sets that occur as preimages under f are precisely $\emptyset, V, X \setminus V, X$, which are all open, showing f to be continuous. Note that (4.52) also holds, such that f, clearly, once again, satisfies all required conditions. ■

Theorem 4.63 (Riesz Representation Theorem)**.** *Let the topological space $(X,\mathcal{T})$ be locally compact and T_2. Moreover, let $\alpha : C_c(X) \longrightarrow \mathbb{K}$ be a positive linear functional. Then there exists a unique positive locally finite Borel measure space $(X,\mathcal{B},\mu)$, $\mathcal{B} = \sigma(\mathcal{T})$, such that α is given by* (4.51) *and such that μ has the following regularity properties:*

(i) *$\mu(K) < \infty$ for each compact $K \subseteq X$.*

(ii) *It holds that*

$$\underset{A\in\mathcal{B}}{\forall} \quad \mu(A) = \inf\{\mu(O) : A \subseteq O,\ O \text{ open}\}.$$

(iii) *It holds that*

$$\underset{A\in\mathcal{B}}{\forall} \quad \Big((A \text{ open} \vee \mu(A) < \infty) \ \Rightarrow\ \mu(A) = \sup\{\mu(K) : K \subseteq A,\ K \text{ compact}\} \Big).$$

One can also obtain a complete *Borel measure space $(X,\mathcal{A},\mu)$ with the above properties (with $\mathcal{B}$ replaced by $\mathcal{A}$) by letting $(X,\mathcal{A},\mu)$ be the completion of $(X,\mathcal{B},\mu)$ in the sense of* [Phi17, Def. 1.51]*. In the above situation, we then say that α is* represented *by the measure $\mu : \mathcal{B} \longrightarrow [0,\infty]$ (resp. by the measure $\mu : \mathcal{A} \longrightarrow [0,\infty]$).*

Proof. As before, since $(X,\mathcal{T})$ is T_2, each compact set K is closed, i.e. $K \in \mathcal{B}$, $\mu(K)$ is defined. In particular, statements (i) and (iii) make sense.

Uniqueness: If μ satisfies (iii), then its values on compact sets uniquely determine its values on all open sets. If it satisfies (ii) as well, then it is uniquely determined on each $A \in \mathcal{B}$ (and on each $A \in \mathcal{A}$ in case of the completion due to [Phi17, Th. 1.50(a)]). Thus, it suffices to show that if $\mu, \nu : \mathcal{B} \longrightarrow [0, \infty]$ are measures satisfying the conditions of the theorem, then $\mu(K) = \nu(K)$ for each compact $K \subseteq X$. Let μ, ν be such measures and $K \subseteq X$ compact. Fix $\epsilon \in \mathbb{R}^+$. As a consequence of (i) and (ii), there exists an open $O \subseteq X$ with $K \subseteq O$ and $\nu(O) < \nu(K) + \epsilon$. Due to Prop. 4.62,

$$\underset{f \in C_c(X)}{\exists} \quad \Big(0 \leq f \leq 1 \ \wedge \ f{\restriction_K} \equiv 1 \ \wedge \ \operatorname{supp} f \subseteq O\Big).$$

Thus,

$$\mu(K) = \int_X \chi_K \,\mathrm{d}\mu \leq \int_X f \,\mathrm{d}\mu \overset{(4.51)}{=} \alpha(f) \overset{(4.51)}{=} \int_X f \,\mathrm{d}\nu \leq \int_X \chi_O \,\mathrm{d}\nu = \nu(O) < \nu(K) + \epsilon,$$

implying $\mu(K) \leq \nu(K)$, as $\epsilon \in \mathbb{R}^+$ was arbitrary. As we can switch the roles of μ and ν, we obtain $\mu(K) = \nu(K)$, completing the proof that μ is unique on $\mathcal{B}$.

It remains to show the existence of μ, which requires some work. The idea is to define an outer measure $\mu : \mathcal{P}(X) \longrightarrow [0, \infty]$ (cf. [Phi17, Def. 1.32]) such that the restriction of μ to $\mathcal{A}$ (i.e. to the completion of $\mathcal{B}$) is a measure with all desired properties. We first define μ on open sets O by letting

$$\mu(O) := \sup\big\{\alpha(f) : f \in C_c(X),\, 0 \leq f \leq 1,\, \operatorname{supp} f \subseteq O\big\}. \tag{4.53}$$

Note that $f \equiv 0 \in C_c(X)$ with $\operatorname{supp} f = \emptyset \subseteq O$ for each O, such that $\mu(O)$ is well-defined by (4.53). If $O_1, O_2 \subseteq X$ are open with $O_1 \subseteq O_2$, then $\mu(O_1) \leq \mu(O_2)$ is immediate from (4.53). Thus, if we define

$$\mu : \mathcal{P}(X) \longrightarrow [0, \infty], \quad \mu(A) := \inf\big\{\mu(O) : A \subseteq O,\, O \text{ open}\big\}, \tag{4.54}$$

then the values given by (4.53) and (4.54) are the same.

Claim 1. μ is an outer measure on X.

Proof. We have to show $\mu(\emptyset) = 0$, μ is monotone, and σ-subadditive. Since, for each $f \in C_c(X)$, we have $\operatorname{supp} f = \emptyset$ if, and only if, $f \equiv 0$, $\mu(\emptyset) = 0$ is clear from (4.53) as well as $\mu \geq 0$. If $A_1, A_2 \subseteq X$ with $A_1 \subseteq A_2$, then $\mu(A_1) \leq \mu(A_2)$ is immediate from (4.54), proving monotonicity. To prove σ-subadditivity, we first consider open sets $O_1, \dots, O_N \subseteq X$, $N \in \mathbb{N}$. Let $f \in C_c(X)$ with $0 \leq f \leq 1$ and $\operatorname{supp} f \subseteq \bigcup_{i=1}^N O_i$. Then $(O_1, \dots, O_N)$ is an open cover of the compact set $\operatorname{supp} f$ and Th. F.1 provides a corresponding partition of unity, i.e. $\varphi_1, \dots, \varphi_N \in C_c(X)$ such that

$$\underset{i \in \{1,\dots,N\}}{\forall} \quad \Big(0 \leq \varphi_i \leq 1 \ \wedge \ \operatorname{supp} \varphi_i \subseteq O_i\Big), \quad \left(\sum_{i=1}^N \varphi_i\right){\restriction_{\operatorname{supp} f}} \equiv 1,$$

implying

$$\forall_{i\in\{1,\dots,N\}} \quad \Big(\operatorname{supp}(\varphi_i f) \subseteq O_i,\ 0 \le \varphi_i f \le 1\Big), \quad f = \sum_{i=1}^{N} (\varphi_i f),$$

and

$$\alpha(f) = \sum_{i=1}^{N} \alpha(\varphi_i f) \le \sum_{i=1}^{N} \mu(O_i).$$

As $f \in C_c(X)$ with $0 \le f \le 1$ and $\operatorname{supp} f \subseteq \bigcup_{i=1}^{N} O_i$ was arbitrary, we obtain

$$\mu\left(\bigcup_{i=1}^{N} O_i\right) \le \sum_{i=1}^{N} \mu(O_i). \tag{4.55}$$

Now let $(A_k)_{k\in\mathbb{N}}$ be a sequence of subsets of X, $A := \bigcup_{k\in\mathbb{N}} A_k$. We need to prove

$$\mu(A) \le \sum_{k=1}^{\infty} \mu(A_k). \tag{4.56}$$

If there exists $k \in \mathbb{N}$ with $\mu(A_k) = \infty$, then (4.56) holds. Thus, we now assume $\mu(A_k) < \infty$ for each $k \in \mathbb{N}$. Fix $\epsilon \in \mathbb{R}^+$. Then, by (4.54),

$$\forall_{k\in\mathbb{N}} \quad \exists_{O_k \subseteq X} \quad \Big(O_k \text{ open} \ \wedge\ A_k \subseteq O_k \ \wedge\ \mu(O_k) < \mu(A_k) + \epsilon\, 2^{-k}\Big).$$

Then $O := \bigcup_{k\in\mathbb{N}} O_k$ is open. Consider $f \in C_c(X)$ with $0 \le f \le 1$ and $\operatorname{supp} f \subseteq O$. Since $\operatorname{supp} f$ is compact, there exists $N \in \mathbb{N}$ such that $\operatorname{supp} f \subseteq \bigcup_{k=1}^{N} O_k$. Then

$$\alpha(f) \overset{(4.53)}{\le} \mu\left(\bigcup_{k=1}^{N} O_k\right) \overset{(4.55)}{\le} \sum_{k=1}^{N} \mu(O_k) \le \sum_{k=1}^{\infty} \mu(A_k) + \epsilon.$$

As $f \in C_c(X)$ with $0 \le f \le 1$ and $\operatorname{supp} f \subseteq O$ was arbitrary, we obtain $\mu(O) \le \sum_{k=1}^{\infty} \mu(A_k) + \epsilon$. Since $A \subseteq O$, this implies

$$\mu(A) \le \mu(O) \le \sum_{k=1}^{\infty} \mu(A_k) + \epsilon.$$

Since $\epsilon \in \mathbb{R}^+$ was arbitrary, (4.56) holds. ▲

We now define the following collections of subsets of X:

$$\mathcal{E} := \Big\{A \subseteq X :\ \mu(A) < \infty \ \wedge\ \mu(A) = \sup\{\mu(K) :\ K \subseteq A,\ K \text{ compact}\}\Big\}, \tag{4.57}$$

$$\mathcal{M} := \left\{A \subseteq X :\ \forall_{K \subseteq X \text{ compact}} \ A \cap K \in \mathcal{E}\right\}. \tag{4.58}$$

It now suffices to show that $\mathcal{M}$ is a σ-algebra with $\mathcal{B} \subseteq \mathcal{M}$ and that $\mu{\restriction}_{\mathcal{M}}$ is a complete measure, satisfying (4.51) and (i) – (iii) with $\mathcal{B}$ replaced by $\mathcal{M}$ (then, in particular, $\mathcal{B} \subseteq \mathcal{A} \subseteq \mathcal{M}$).

Claim 2. If $K \subseteq X$ is compact, then $K \in \mathcal{E}$ and

$$\mu(K) = \inf\{\alpha(f) : f \in C_c(X),\, 0 \le f \le 1,\, f{\restriction}_K \equiv 1\}. \tag{4.59}$$

Proof. Let $K \subseteq X$ be compact and let $f \in C_c(X)$ with $0 \le f \le 1$ and $f{\restriction}_K \equiv 1$. Define

$$\forall_{s \in]0,1[} \quad O_s := f^{-1}(]s,\infty[).$$

Then each O_s is open, $K \subseteq O_s$, and

$$\forall_{s \in]0,1[} \quad \forall_{\substack{g \in C_c(X),\\ 0 \le g \le 1}} \quad \Big(\operatorname{supp} g \subseteq O_s \;\Rightarrow\; sg \le f \;\Rightarrow\; g \le s^{-1} f\Big).$$

Thus,

$$\forall_{s \in]0,1[} \quad \mu(K) \le \mu(O_s) \overset{(4.53)}{=} \sup\big\{\alpha(g) : g \in C_c(X),\, 0 \le g \le 1,\, \operatorname{supp} g \subseteq O_s\big\} \le s^{-1}\alpha(f).$$

For $s \to 1$, the above inequality yields

$$\mu(K) \le \alpha(f),$$

showing $\mu(K) < \infty$ and $K \in \mathcal{E}$. Now fix $\epsilon \in \mathbb{R}^+$. By (4.54), there exists $O \subseteq X$ open with $K \subseteq O$ and $\mu(O) < \mu(K) + \epsilon$. As before, Prop. 4.62 yields

$$\exists_{f \in C_c(X)} \quad \Big(0 \le f \le 1 \;\wedge\; f{\restriction}_K \equiv 1 \;\wedge\; \operatorname{supp} f \subseteq O\Big).$$

Then $\mu(K) \le \alpha(f) \le \mu(O) < \mu(K) + \epsilon$, proving (4.59), as $\epsilon \in \mathbb{R}^+$ was arbitrary. ▲

Claim 3. If $O \subseteq X$ is open with $\mu(O) < \infty$, then $O \in \mathcal{E}$.

Proof. Let $O \subseteq X$ be open with $\mu(O) < \infty$. Let $s \in \mathbb{R}$ with $s < \mu(O)$. Then, by (4.53), there exists $f \in C_c(X)$ with $0 \le f \le 1$, $\operatorname{supp} f \subseteq O$, and $s < \alpha(f)$. Let $K := \operatorname{supp} f$. If $W \subseteq X$ is open with $K \subseteq W$, then $\alpha(f) \le \mu(W)$, again by (4.53). Now (4.54) yields $s < \alpha(f) \le \mu(K)$. Since $K \subseteq O$ and $s < \mu(K)$, the proof of $O \in \mathcal{E}$ is complete. ▲

Claim 4. Let $(E_k)_{k \in \mathbb{N}}$ be disjoint sequence in $\mathcal{E}$, $E := \bigcup_{k \in \mathbb{N}} E_k$. Then

$$\mu(E) = \sum_{k=1}^{\infty} \mu(E_k). \tag{4.60}$$

If $\mu(E) < \infty$, then $E \in \mathcal{E}$.

Proof. We first show that, if $K_1, \dots, K_N \subseteq X$ are disjoint and compact, $N \in \mathbb{N}$, $K := \bigcup_{i=1}^{N}$, then

$$\mu(K) = \sum_{i=1}^{\infty} \mu(K_i) : \tag{4.61}$$

It suffices to consider $N = 2$, since, then, (4.61) follows by induction. So let $N = 2$ and fix $\epsilon \in \mathbb{R}^+$. Due to Prop. 4.62,

$$\underset{f \in C_c(X)}{\exists} \Big(0 \leq f \leq 1 \ \wedge \ f{\restriction}_{K_1} \equiv 1 \ \wedge \ \operatorname{supp} f \subseteq X \setminus K_2 \Big).$$

Moreover, by (4.59),

$$\underset{g \in C_c(X)}{\exists} \Big(0 \leq g \leq 1, \ \wedge \ g{\restriction}_{K} \equiv 1 \ \wedge \ \alpha(g) < \mu(K) + \epsilon \Big).$$

Then $fg, (1-f)g \in C_c(X)$, $0 \leq fg, (1-f)g$, $(fg){\restriction}_{K_1} \equiv 1$, $((1-f)g){\restriction}_{K_2} \equiv 1$, such that (4.59) and the linearity of α imply

$$\mu(K_1) + \mu(K_2) \leq \alpha(fg) + \alpha(g - fg) = \alpha(g) < \mu(K) + \epsilon.$$

As $\epsilon \in \mathbb{R}^+$ was arbitrary, we obain $\mu(K_1) + \mu(K_2) \leq \mu(K)$. Since we already know μ to be σ-subadditive, (4.61) is proved. Next, notice that the σ-subadditivity of μ also implies (4.60) in the case, where $\mu(E) = \infty$. In the case, where $\mu(E) < \infty$, we, again, fix $\epsilon \in \mathbb{R}^+$. For each $k \in \mathbb{N}$, since $E_k \in \mathcal{E}$, there exists a compact $K_k \subseteq X$, $K_k \subseteq E_k$, such that

$$\mu(K_k) > \mu(E_k) - 2^{-k}\,\epsilon.$$

If we now let $H_k := \bigcup_{i=1}^{k} K_i$, then (4.61) implies

$$\mu(E) \geq \mu(H_k) = \sum_{i=1}^{k} \mu(K_i) > \sum_{i=1}^{k} \mu(E_i) - \epsilon.$$

As the previous inequality holds for each $k \in \mathbb{N}$ and each $\epsilon \in \mathbb{R}^+$, we have $\mu(E) \geq \sum_{k=1}^{\infty} \mu(E_k)$. As the opposite inequality holds due to the σ-subadditivity of μ, (4.60) is proved. ▲

Claim 5. For each $E \in \mathcal{E}$ and each $\epsilon \in \mathbb{R}^+$, there exist $O, K \subseteq X$ such that O is open, K is compact, $K \subseteq E \subseteq O$, and $\mu(O \setminus K) < \epsilon$.

Proof. Let $E \in \mathcal{E}$ and $\epsilon \in \mathbb{R}^+$. From the definition of μ, we optain an open $O \subseteq X$, $E \subseteq O$, and from the definition of $\mathcal{E}$, we obtain a compact $K \subseteq E$, satisfying

$$\mu(O) - \frac{\epsilon}{2} < \mu(E) < \mu(K) + \frac{\epsilon}{2}.$$

As $O \setminus K$ is open, we have $O \setminus K \in \mathcal{E}$ by Cl. 3. Thus, by Cl. 3,

$$\mu(K) + \mu(O \setminus K) = \mu(O) < \mu(K) + \epsilon,$$

thereby establishing the case. ▲

Claim 6. If $E, F \in \mathcal{E}$, then $E \setminus F \in \mathcal{E}$, $E \cup F \in \mathcal{E}$, and $E \cap F \in \mathcal{E}$.

Proof. Let $E, F \in \mathcal{E}$ and fix $\epsilon \in \mathbb{R}^+$. According to Cl. 5, there exist $O_1, O_2, K_1, K_2 \subseteq X$ such that O_1, O_2 are open, K_1, K_2 are compact, $K_1 \subseteq E \subseteq O_1$, $K_2 \subseteq F \subseteq O_2$, $\mu(O_1 \setminus K_1) < \epsilon$ and $\mu(O_2 \setminus K_2) < \epsilon$. Then, since

$$E \setminus F \subseteq O_1 \setminus K_2 \subseteq (O_1 \setminus K_1) \cup (K_1 \setminus O_2) \cup (V_2 \setminus K_2)$$

the σ-subadditivity of μ shows

$$\mu(E \setminus F) \leq \epsilon + \mu(K_1 \setminus O_2) + \epsilon. \tag{4.62}$$

As $K_1 \setminus O_2$ is compact, (4.62) and the definition of $\mathcal{E}$ show $E \setminus F \in \mathcal{E}$. Then $E \cup F = (E \setminus F) \dot{\cup} F \in \mathcal{E}$ by Cl. 4 and $E \cap F = E \setminus (E \setminus F) \in \mathcal{E}$ as well. ▲

Claim 7. $\mathcal{M}$ is a σ-algebra and $\mathcal{B} \subseteq \mathcal{M}$.

Proof. Let $K \subseteq X$ be compact. Suppose $A \in \mathcal{M}$. Then $A \cap K \in \mathcal{E}$ and $A^c \cap K = K \setminus (A \cap K) \in \mathcal{E}$ by Cl. 6. Thus, $A^c \in \mathcal{M}$ by the definition of $\mathcal{M}$. Now let $(A_k)_{k \in \mathbb{N}}$ be a sequence in $\mathcal{M}$ and $A := \bigcup_{k \in \mathbb{N}} A_k$. Set $E_1 := A_1 \cap K$ and

$$\underset{k \geq 2}{\forall} \quad E_k := (A_k \cap K) \setminus \bigcup_{i=1}^{k-1} E_i.$$

Then, Cl. 6, $(B_k)_{k \in \mathbb{N}}$ is a disjoint sequence in $\mathcal{E}$, implying $A \cap K = \bigcup_{k \in \mathbb{N}} E_k \in \mathcal{E}$ by Cl. 4. Thus, $A \in \mathcal{M}$ by the definition of $\mathcal{M}$. Now let $C \subseteq X$ be closed. Then $C \cap K \in \mathcal{E}$, since $C \cap K$ is compact. As $\mathcal{B}$ is generated by the closed subsets of X, we obtain $\mathcal{B} \subseteq \mathcal{M}$ and the claim is proved. ▲

Claim 8. We have $\mathcal{E} = \{E \in \mathcal{M} : \mu(E) < \infty\}$.

Proof. Denote $\mathcal{F} := \{E \in \mathcal{M} : \mu(E) < \infty\}$ Let $E \in \mathcal{E}$. Then $\mu(E) < \infty$ and, if $K \subseteq X$ is compact, then $E \cap K \in \mathcal{E}$, showing $E \in \mathcal{F}$. Conversely, let $E \in \mathcal{M}$ with $\mu(E) < \infty$ and fix $\epsilon \in \mathbb{R}^+$. Since $\mu(E) < \infty$, there exists an open set $O \subseteq X$ such that $E \subseteq O$. Since $O \in \mathcal{E}$, there exists a compact set $K \subseteq O$ with $\mu(O \setminus K) < \epsilon$. Moreover, since $E \cap K \in \mathcal{E}$, there exists a compact set $H \subseteq E \cap K$ with

$$\mu(E \cap K) < \mu(H) + \epsilon.$$

Thus, we use $E \subseteq (E \cap K) \cup (O \setminus K)$ to obtain

$$\mu(E) \leq \mu(E \cap K) + \mu(O \setminus K) < \mu(H) + 2\epsilon,$$

showing $E \in \mathcal{E}$ as desired. ▲

Claim 9. $\mu{\restriction}_{\mathcal{M}}$ is a locally finite complete measure.

Proof. To see that $\mu{\restriction}_{\mathcal{M}}$ is a measure, one merely observes that μ is σ-additive on $\mathcal{M}$ as a consequence of Claims 4 and 8. Moreover, for each $x \in X$, $\{x\}$ is compact, i.e. $\mu{\restriction}_{\mathcal{M}}$ is locally finite as a consequence of Cl. 2 and (4.54). To see that $\mu{\restriction}_{\mathcal{M}}$ is complete, let $A \in \mathcal{M}$ with $\mu(A) = 0$ and $E \subseteq A$. Then $\mu(E) = 0$ and also $A \in \mathcal{E}$ by Cl. 8. Since $0 = \mu(E) = \sup\{\mu(K) : K \subseteq E, K$ compact, we have $E \in \mathcal{E} \subseteq \mathcal{M}$, proving completeness. ▲

Claim 10. μ satisfies (4.51), i.e.

$$\underset{f \in C_c(X)}{\forall} \quad \alpha(f) = \int_X f \,\mathrm{d}\mu\,.$$

Proof. Let $f \in C_c(X)$. As $f = \operatorname{Re} f + i \operatorname{Im} f$, it suffices to prove (4.51) for $\mathbb{R}$-valued f. Next, we observe it suffices to show

$$\underset{f \in C_c(X,\mathbb{R})}{\forall} \quad \alpha(f) \leq \int_X f \,\mathrm{d}\mu\; : \tag{4.63}$$

If (4.63) holds, then

$$\underset{f \in C_c(X,\mathbb{R})}{\forall} \quad -\alpha(f) = \alpha(-f) \leq \int_X (-f) \,\mathrm{d}\mu\; = -\int_X f \,\mathrm{d}\mu\,,$$

showing (4.63) to imply (4.51). It remains to prove (4.63). Let $f \in C_c(X,\mathbb{R})$, $K :=$ $\operatorname{supp} f$. As f is continuous, $f(X) = f(K) \cup \{0\}$ is compact. In particular, there exist $a, b \in \mathbb{R}$ such that $a < b$ and $f(X) \subseteq [a,b]$. Fix $\epsilon \in \mathbb{R}^+$. Now choose $s_0, s_1, \dots, s_n \in \mathbb{R}$, $n \in \mathbb{N}$, such that

$$s_0 < a < s_1 < \cdots < s_n = b \quad \wedge \quad \underset{i \in \{1,\dots,n\}}{\forall} \quad s_i - s_{i-1} < \epsilon\, : \tag{4.64}$$

If $\delta := (b-a)/(\epsilon/2) \notin \mathbb{N}$, then let $n := \min\{i \in \mathbb{N} :\, b - i(\epsilon/2) < a\}$ and $s_0 := b - n(\epsilon/2)$. If $\delta \in \mathbb{N}$, then $\gamma := (b-a)/(\epsilon/\sqrt{2}) \notin \mathbb{N}$ (otherwise, $\sqrt{2} = 2\gamma/\delta \in \mathbb{Q}$). In this case, let $n := \min\{i \in \mathbb{N} :\, b - i(\epsilon/\sqrt{2}) < a\}$ and $s_0 := b - n(\epsilon/\sqrt{2})$. Next, define

$$\underset{i \in \{1,\dots,n\}}{\forall} \quad E_i := K \cap \{x \in X :\, s_{i-1} < f(x) \leq s_i\}$$

and observe that each $E_i \in \mathcal{B}$ due to the continuity of f. Moreover, $E_1, \dots, E_n$ are, clearly, disjoint with $K = \bigcup_{i=1}^n E_i$. Since each $E_i \in \mathcal{E}$, there exist open sets $O_1, \dots, O_n \subseteq X$, satisfying

$$\forall_{i \in \{1,\dots,n\}} \quad \Big(E_i \subseteq O_i \subseteq f^{-1}(]-\infty, s_i + \epsilon[) \ \wedge \ \mu(O_i) < \mu(E_i) + \frac{\epsilon}{n} \Big). \tag{4.65}$$

As the $O_1, \dots, O_N$ cover K, we can apply Th. F.1 to obtain a corresponding partition of unity, i.e. $\varphi_1, \dots, \varphi_n \in C_{\mathrm{c}}(X)$ such that

$$\forall_{i \in \{1,\dots,n\}} \quad \Big(0 \leq \varphi_i \leq 1 \ \wedge \ \operatorname{supp} \varphi_i \subseteq O_i \Big), \qquad \left(\sum_{i=1}^n \varphi_i \right) \restriction_K \equiv 1,$$

implying $f = \sum_{i=1}^n (\varphi_i f)$. From Cl. 2, we obtain

$$\mu(K) \leq \alpha \left(\sum_{i=1}^n \varphi_i \right) = \sum_{i=1}^n \alpha(\varphi_i). \tag{4.66}$$

Next, we observe that (4.64) and (4.65) imply

$$\forall_{i \in \{1,\dots,n\}} \quad \Big(\varphi_i f \leq (s_i + \epsilon)\, \varphi_i \ \wedge \ s_i - \epsilon < s_{i-1} < f \restriction_{E_i} \Big), \tag{4.67}$$

and, using the monotonicity of α, we estimate

$$\begin{aligned}
\alpha(f) \quad &= \quad \sum_{i=1}^n \alpha(\varphi_i f) \overset{(4.67)}{\leq} \sum_{i=1}^n (s_i + \epsilon)\, \alpha(\varphi_i) \\
&= \quad \sum_{i=1}^n (|a| + s_i + \epsilon)\, \alpha(\varphi_i) - \sum_{i=1}^n |a|\, \alpha(\varphi_i) \\
&\overset{(4.53),(4.66)}{\leq} \quad \sum_{i=1}^n (|a| + s_i + \epsilon)\, \mu(O_i) - |a|\, \mu(K) \\
&\overset{(4.65)}{\leq} \quad \sum_{i=1}^n (|a| + s_i + \epsilon) \left(\mu(E_i) + \frac{\epsilon}{n} \right) - |a|\, \mu(K) \\
&= \quad \sum_{i=1}^n (s_i - \epsilon)\, \mu(E_i) + 2\epsilon\, \mu(K) + \frac{\epsilon}{n} \sum_{i=1}^n (|a| + s_i + \epsilon) \\
&\overset{(4.67)}{\leq} \quad \int_X f \,\mathrm{d}\mu + \epsilon\big(2\mu(K) + |a| + b + \epsilon \big).
\end{aligned} \tag{4.68}$$

As $\epsilon \in \mathbb{R}^+$ was arbitrary, (4.68) proves (4.63) and the claim. ▲

Now (i) holds by Cl. 2, (ii) (with $\mathcal{B}$ replaced by $\mathcal{M}$) holds by (4.54), (iii) (with $\mathcal{B}$ replaced by $\mathcal{M}$) holds by Cl. 8, and μ satisfies (4.51) by Cl. 9. Since $\mathcal{B} \subseteq \mathcal{M}$ and $\mu{\restriction}_{\mathcal{M}}$ is complete, $\mathcal{A} \subseteq \mathcal{M}$, concluding the proof. ■

Corollary 4.64. *Let the topological space $(X, \mathcal{T})$ be locally compact and T_2. Moreover, let $\alpha : C_c(X) \longrightarrow \mathbb{K}$ be a positive linear functional. Then α is continuous with respect to $\|\cdot\|_\infty$ on $C_c(X)$ if, and only if, the measure μ, representing α and given by Th. 4.63, is finite. Moreover, in that case, $\|\alpha\| = \mu(X)$.*

Proof. If μ is finite, then

$$\forall_{f \in C_c(X)} \quad |\alpha(f)| \overset{(4.51)}{=} \left| \int_X f \,\mathrm{d}\mu \right| \leq \int_X |f| \,\mathrm{d}\mu \leq \|f\|_\infty \, \mu(X),$$

implying α to be continuous with $\|\alpha\| \leq \mu(X)$. On the other hand, if α is continuous, then

$$\forall_{f \in C_c(X)} \quad |\alpha(f)| \leq \|\alpha\| \, \|f\|_\infty.$$

Let $K \subseteq X$ be compact. Due to Prop. 4.62,

$$\exists_{f_K \in C_c(X)} \quad \Big(0 \leq f_K \leq 1 \ \wedge \ f_K{\restriction}_K \equiv 1\Big).$$

Then

$$\mu(K) = \int_X \chi_K \,\mathrm{d}\mu \leq \int_X f_K \,\mathrm{d}\mu \overset{(4.51)}{=} \alpha(f_K) \leq \|\alpha\| \, \|f_K\|_\infty = \|\alpha\|,$$

implying

$$\mu(X) \overset{\text{Th. 4.63(iii)}}{=} \sup\big\{\mu(K) : K \subseteq X, \ K \text{ compact}\big\} \leq \|\alpha\|,$$

showing μ to be finite and, in combination with $\|\alpha\| \leq \mu(X)$ from above, $\|\alpha\| = \mu(X)$. ■

Corollary 4.65. *Let the topological space $(K, \mathcal{T})$ be compact and T_2. Moreover, let $\alpha : C(K) \longrightarrow \mathbb{K}$ be a positive linear functional.*

(a) *Then the measure μ, representing α and given by Th. 4.63, is finite and satisfies*

$$\forall_{A \in \mathcal{A}} \quad \mu(A) = \inf\{\mu(O) : A \subseteq O, \ O \text{ open}\}, \tag{4.69a}$$

$$\forall_{A \in \mathcal{A}} \quad \mu(A) = \sup\{\mu(H) : H \subseteq A, \ H \text{ compact}\}, \tag{4.69b}$$

where, as in Th. 4.63, $(X, \mathcal{A}, \mu)$ is the completion of $(X, \mathcal{B}, \mu)$.

(b) *α is continuous.*

Proof. If $(K, \mathcal{T})$ is compact, then it is locally compact and Th. 4.63 applies, yielding the measure μ, representing α. Then μ is finite by Th. 4.63(i). In consequence, α is continuous by Cor. 4.64, proving (b). Moreover, (4.69a) is the same as Th. 4.63(ii), and (4.69b) is implied by Th. 4.63(iii), thereby proving (a). ■

While Th. 4.63 and Cor. 4.64 provided representations of positive (resp. continuous positive) linear functionals on $C_c(X)$ for X being a compact T_2 space, one would also like to obtain representations for continuous linear functionals that are not necessarily positive (i.e. a representation of the dual of $(C_c(X), \|\cdot\|_\infty)$). It turns out we can build on Th. 4.63 and Cor. 4.64 to achieve the desired representation in the following Th. 4.76. First, we still need some preparations:

Proposition 4.66. *Let $(X, \mathcal{A})$ be a measurable space and let $\mu, \nu : \mathcal{A} \longrightarrow [0, \infty]$ be positive measures. If $\mu \leq \nu$ and $f : X \longrightarrow \mathbb{K}$ is ν-integrable, then it is μ-integrable.*

Proof. Let $f : X \longrightarrow \mathbb{K}$ be ν-integrable. If $A \in \mathcal{A}$ and $f = \chi_A$, then $\int_X f \, d\mu = \mu(A) \leq \nu(A) = \int_X f \, d\nu < \infty$, showing f to be μ-integrable. Thus, the assertion also holds if f is a nonnegative simple function. If f is an arbitrary ν-integrable nonnegative function, then let $(\phi_k)_{k \in \mathbb{N}}$ be an increasing sequence of ν-integrable nonnegative simple functions with $\phi_k \uparrow f$. Then, by the monotone convergence theorem,

$$\int_X f \, d\mu = \lim_{k \to \infty} \int_X \phi_k \, d\mu \leq \lim_{k \to \infty} \int_X \phi_k \, d\nu = \int_X f \, d\nu < \infty,$$

showing f to be μ-integrable also in this case. Finally, if f is an arbitrary ν-integrable function, then $|f|$ is ν-integrable, implying $|f|$ to be μ-integrable, which, in turn, implies f to be μ-integrable. ■

Remark 4.67. (a) Let $(X, \mathcal{A}, \mu)$ be a signed measure space, where $|\mu|$ denotes the total variation of μ. From Def. and Rem. 4.47(a), we recall the decomposition $|\mu| = \mu^+ + \mu^-$, where the positive measures μ^+, μ^- were also defined in Def. and Rem. 4.47(a). Thus, if $f : X \longrightarrow \mathbb{K}$ is $|\mu|$-integrable, then Prop. 4.66 implies f to be μ^+-integrable and μ^--integrable.

(b) Let $(X, \mathcal{A}, \mu)$ be a complex measure space. Since, for each $z \in \mathbb{C}$, one has the estimate $\max\{|\operatorname{Re} z|, |\operatorname{Im} z|\} \leq |z|$, Def. 4.42 implies, for the total variations, $\max\{|\operatorname{Re} \mu|, |\operatorname{Im} \mu|\} \leq |\mu|$. Thus, if $f : X \longrightarrow \mathbb{K}$ is $|\mu|$-integrable, then (a) and Prop. 4.66 imply f to be $(\operatorname{Re} \mu)^+$-, $(\operatorname{Re} \mu)^-$-, $(\operatorname{Im} \mu)^+$-, and $(\operatorname{Im} \mu)^-$-integrable.

—

Given a complex measure μ, recall the decomposition

$$\mu = (\operatorname{Re} \mu)^+ - (\operatorname{Re} \mu)^- + i \left((\operatorname{Im} \mu)^+ - (\operatorname{Im} \mu)^- \right).$$

from Def. and Rem. 4.47(b). This gives rise to the following definition:

Definition 4.68. Let $(X, \mathcal{A}, \mu)$ be a complex measure space. For each $f \in \mathcal{L}^1(|\mu|)$, define the integral

$$\int_X f \,\mathrm{d}\mu := \int_X f \,\mathrm{d}(\operatorname{Re}\mu)^+ - \int_X f \,\mathrm{d}(\operatorname{Re}\mu)^- + i\left(\int_X f \,\mathrm{d}(\operatorname{Im}\mu)^+ - \int_X f \,\mathrm{d}(\operatorname{Im}\mu)^-\right),$$

where all integrals in this definition are well-defined by Rem. 4.67(b).

Proposition 4.69. *Let $(X, \mathcal{A}, \mu)$ be a complex measure space.*

(a) *The integral, as defined in Def. 4.68, is linear on $\mathcal{L}^1(|\mu|)$.*

(b) *One has*

$$\underset{f \in \mathcal{L}^\infty(|\mu|)}{\forall} \quad \left|\int_X f \,\mathrm{d}\mu\right| \leq \int_X |f| \,\mathrm{d}|\mu| .$$

Note: Using the Radon-Nikodym theorem, one can even show the above inequality to hold for each $f \in \mathcal{L}^1(|\mu|)$ (cf. [Rud87, Th. 6.12, Eq. 6.18(1)]*). In the following, we will only need the inequality for $f \in \mathcal{L}^\infty(|\mu|)$, which allows a more elementary proof.*

Proof. (a): One computes, for each $s \in \mathbb{K}$, $f \in C_c(X)$,

$$\begin{aligned}
\int_X sf \,\mathrm{d}\mu &= s\int_X f \,\mathrm{d}(\operatorname{Re}\mu)^+ - s\int_X f \,\mathrm{d}(\operatorname{Re}\mu)^- \\
&\quad + i\left(s\int_X f \,\mathrm{d}(\operatorname{Im}\mu)^+ - s\int_X f \,\mathrm{d}(\operatorname{Im}\mu)^-\right) \\
&= s\int_X f \,\mathrm{d}\mu
\end{aligned}$$

and, for each $f, g \in C_c(X)$,

$$\begin{aligned}
\int_X (f+g) \,\mathrm{d}\mu &= \int_X (f+g) \,\mathrm{d}(\operatorname{Re}\mu)^+ - \int_X (f+g) \,\mathrm{d}(\operatorname{Re}\mu)^- \\
&\quad + i\left(\int_X (f+g) \,\mathrm{d}(\operatorname{Im}\mu)^+ - \int_X (f+g) \,\mathrm{d}(\operatorname{Im}\mu)^-\right) \\
&= \int_X f \,\mathrm{d}\mu + \int_X g \,\mathrm{d}\mu ,
\end{aligned}$$

proving the linearity of the integral.

(b): Let $A \in \mathcal{A}$. Then

$$\int_X \chi_A \,\mathrm{d}\mu = (\operatorname{Re}\mu)^+(A) - (\operatorname{Re}\mu)^-(A) + i\,\big((\operatorname{Im}\mu)^+(A) - (\operatorname{Im}\mu)^-(A)\big) = \mu(A).$$

Thus, if $f = \sum_{k=1}^N \lambda_k \chi_{A_k}$, $N \in \mathbb{N}$, with $\lambda_1, \dots, \lambda_N \in \mathbb{C}$ and $A_1, \dots, A_N \in \mathcal{A}$ disjoint, then

$$\left|\int_X f \,\mathrm{d}\mu\right| \overset{\text{(a)}}{\leq} \sum_{k=1}^N |\lambda_k||\mu(A_k)| \leq \sum_{k=1}^N |\lambda_k||\mu|(A_k) = \int_X |f| \,\mathrm{d}|\mu|\,,$$

proving (b) to hold for each simple function $f \in \mathcal{S}(|\mu|)$. Now let $f \in \mathcal{L}^\infty(|\mu|)$ be arbitrary. Then, by [Phi17, Th. 1.90], there exists a sequence $(\phi_k)_{k\in\mathbb{N}}$ in $\mathcal{S}(|\mu|)$ that satisfies $\|\phi_k\|_\infty \leq \|f\|_\infty$ for each $k \in \mathbb{N}$ and converges pointwise (and even uniformly) to f. Thus, the dominated convergence theorem applies, yielding

$$\left|\int_X f \,\mathrm{d}\mu\right| = \left|\lim_{k\to\infty} \int_X \phi_k \,\mathrm{d}\mu\right| \leq \lim_{k\to\infty} \int_X |\phi_k| \,\mathrm{d}|\mu| = \int_X |f| \,\mathrm{d}|\mu|\,,$$

completing the proof of (b). ■

Proposition 4.70. *Let $(X, \mathcal{A})$ be a measurable space.*

(a) *If μ, ν are complex measures on $(X, \mathcal{A})$, then*

$$\underset{f\in\mathcal{L}^1(|\mu|)\cap\mathcal{L}^1(|\nu|)}{\forall} \quad \int_X f \,\mathrm{d}(\mu+\nu) = \int_X f \,\mathrm{d}\mu + \int_X f \,\mathrm{d}\nu\,.$$

(b) *If μ is a complex measure on $(X, \mathcal{A})$ and $s \in \mathbb{C}$, then*

$$\underset{f\in\mathcal{L}^1(|\mu|)}{\forall} \quad \int_X f \,\mathrm{d}(s\mu) = s\int_X f \,\mathrm{d}\mu\,.$$

Proof. (a): It is an exercise to prove the case, where μ, ν are positive measures. Let $\lambda := \mu + \nu$, $f \in \mathcal{L}^1(|\mu|) \cap \mathcal{L}^1(|\nu|)$. If μ, ν are signed measures, then

$$\lambda^+ + \mu^- + \nu^- = \mu^+ + \nu^+ + \lambda^-$$

and, thus, the case of positive measures yields

$$\int_X f \,\mathrm{d}\lambda^+ + \int_X f \,\mathrm{d}\mu^- + \int_X f \,\mathrm{d}\nu^- = \int_X f \,\mathrm{d}\mu^+ + \int_X f \,\mathrm{d}\nu^+ + \int_X f \,\mathrm{d}\lambda^-$$

and

$$\begin{aligned}\int_X f \, \mathrm{d}(\mu+\nu) &= \int_X f \, \mathrm{d}\lambda = \int_X f \, \mathrm{d}\lambda^+ - \int_X f \, \mathrm{d}\lambda^- \\ &= \int_X f \, \mathrm{d}\mu^+ + \int_X f \, \mathrm{d}\nu^+ - \int_X f \, \mathrm{d}\mu^- - \int_X f \, \mathrm{d}\nu^- = \int_X f \, \mathrm{d}\mu + \int_X f \, \mathrm{d}\nu \,.\end{aligned}$$

The case of general complex measures now also follows, since $\operatorname{Re}\lambda = \operatorname{Re}\mu + \operatorname{Re}\nu$ and $\operatorname{Im}\lambda = \operatorname{Im}\mu + \operatorname{Im}\nu$.

(b): First, we consider the case, where μ is a positive measure, $s \in \mathbb{R}_0^+$: If $A \in \mathcal{A}$, then

$$\int_X \chi_A \, \mathrm{d}(s\mu) = s\,\mu(A) = s \int_X \chi_A \, \mathrm{d}\mu \,.$$

Thus, as the integral is linear, the claimed equality holds for each simple function $f \in \mathcal{S}^+(\mathcal{A})$. If f is nonnegative and measurable, then there exists a sequence $(f_k)_{k\in\mathbb{N}}$ in $\mathcal{S}^+(\mathcal{A})$ with $f_k \uparrow f$, implying

$$\int_X f \, \mathrm{d}(s\mu) = \lim_{k\to\infty} \int_X f_k \, \mathrm{d}(s\mu) = s \lim_{k\to\infty} \int_X f_k \, \mathrm{d}\mu = s \int_X f \, \mathrm{d}\mu \,.$$

The general case of $f \in \mathcal{L}^1(\mu)$ now also follows, since

$$f = (\operatorname{Re} f)^+ - (\operatorname{Re} f)^- + i\,\big((\operatorname{Im} f)^+ - (\operatorname{Im} f)^-\big).$$

Now let μ be a general complex measure, $f \in \mathcal{L}^1(|\mu|)$. If $s \in \mathbb{R}_0^+$, then

$$\begin{aligned}\int_X f \, \mathrm{d}(s\mu) &= \int_X f \, \mathrm{d}(\operatorname{Re}(s\mu))^+ - \int_X f \, \mathrm{d}(\operatorname{Re}(s\mu))^- \\ &\quad + i\left(\int_X f \, \mathrm{d}(\operatorname{Im}(s\mu))^+ - \int_X f \, \mathrm{d}(\operatorname{Im}(s\mu))^-\right) \\ &= \int_X f \, \mathrm{d}s(\operatorname{Re}\mu)^+ - \int_X f \, \mathrm{d}s(\operatorname{Re}\mu)^- \\ &\quad + i\left(\int_X f \, \mathrm{d}s(\operatorname{Im}\mu)^+ - \int_X f \, \mathrm{d}s(\operatorname{Im}\mu)^-\right) \\ &= s\int_X f \, \mathrm{d}\mu \,.\end{aligned}$$

The case $s = -1$ is obtained from

$$\begin{aligned}\int_X f \, \mathrm{d}(-\mu) &= -\int_X f \, \mathrm{d}(\operatorname{Re}\mu)^+ + \int_X f \, \mathrm{d}(\operatorname{Re}\mu)^- \\ &\quad - i\left(\int_X f \, \mathrm{d}(\operatorname{Im}\mu)^+ - \int_X f \, \mathrm{d}(\operatorname{Im}\mu)^-\right) \\ &= -\int_X f \, \mathrm{d}\mu \,.\end{aligned}$$

Since $i\mu = i \operatorname{Re}\mu - \operatorname{Im}\mu$, the case $s = i$ is obtained from

$$\begin{aligned}\int_X f \,\mathrm{d}(i\mu) &= i\left(\int_X f \,\mathrm{d}(\operatorname{Re}\mu)^+ - \int_X f \,\mathrm{d}(\operatorname{Re}\mu)^-\right) \\ &\quad - \left(\int_X f \,\mathrm{d}(\operatorname{Im}\mu)^+ - \int_X f \,\mathrm{d}(\operatorname{Im}\mu)^-\right) \\ &= i\int_X f \,\mathrm{d}\mu\,,\end{aligned}$$

completing the proof. ∎

In the following Def. 4.71, we define so-called *regular Borel measures*. As a caveat, it is pointed out that one finds several variations of this definition in the literature, so one should always verify what precisely is meant by a regular Borel measure in any given text.

Definition 4.71. Let $(X, \mathcal{T})$ be a T_2 topological space and let $(X, \mathcal{A}, \mu)$ be a Borel measure space (positive or complex).

(a) If μ is positive, then μ is called *outer regular* if, and only if,

$$\underset{A\in\mathcal{A}}{\forall} \quad \mu(A) = \inf\{\mu(O) :\, A \subseteq O,\, O \text{ open}\}.$$

If μ is complex, then μ is called *outer regular* if, and only if, each of the positive measures $(\operatorname{Re}\mu)^+$, $(\operatorname{Re}\mu)^-$, $(\operatorname{Im}\mu)^+$, $(\operatorname{Im}\mu)^-$ is outer regular.

(b) If μ is positive, then μ is called *inner regular* if, and only if, $\mu(K) < \infty$ for each compact $K \subseteq X$ and

$$\underset{A\in\mathcal{A}}{\forall} \quad \mu(A) = \sup\{\mu(K) :\, K \subseteq A,\, K \text{ compact}\}.$$

If μ is complex, then μ is called *inner regular* if, and only if, each of the positive measures $(\operatorname{Re}\mu)^+$, $(\operatorname{Re}\mu)^-$, $(\operatorname{Im}\mu)^+$, $(\operatorname{Im}\mu)^-$ is inner regular.

(c) μ is called *regular* if, and only if, μ is both inner and outer regular.

Moreover, let $\mathcal{M}_{\mathbb{K},\mathrm{r}}(X, \mathcal{A})$ denote the set of all $\mathbb{K}$-valued regular measures on $\mathcal{A}$.

Example 4.72. (a) Let $n \in \mathbb{N}$. If $A \in \mathcal{L}^n$ (i.e. if $A \subseteq \mathbb{R}^n$ is Lebesgue-measurable), then $(A, \mathcal{L}^n, \lambda^n)$ is a regular Borel measure due to [Phi17, Th. 1.61(a)] (also cf. [Phi17, Ex. 1.59]).

(b) Let $(X,\mathcal{T})$ be a T_2 topological space and $a \in X$. Then the Dirac measure

$$\delta_a : \mathcal{P}(X) \longrightarrow \{0,1\}, \quad \delta_a(A) := \begin{cases} 1 & \text{if } a \in A, \\ 0 & \text{if } a \notin A, \end{cases}$$

is a regular Borel measure: Let $A \subseteq X$. Then

$$\delta_a(A) := \begin{cases} 1 = \delta_a(\{a\}) = \delta_a(X) & \text{if } a \in A, \\ 0 = \delta_a(\emptyset) = \delta_a(X \setminus \{a\}) & \text{if } a \notin A, \end{cases}$$

where $\{a\}$, $\emptyset$ are compact and X, $X \setminus \{a\}$ are open, showing δ_a to be regular.

Proposition 4.73. *Let $(X,\mathcal{T})$ be a T_2 topological space.*

(a) *If $(X,\mathcal{A},\mu)$ is a complex Borel measure space, then μ is regular if, and only if, for each $A \in \mathcal{A}$ and for each $\epsilon \in \mathbb{R}^+$, there exist $K, O \subseteq X$, where K is compact, O is open, $K \subseteq A \subseteq O$, and $|\mu|(O \setminus K) < \epsilon$.*

(b) *If $(X,\mathcal{A})$ is a measurable space with $\mathcal{B} \subseteq \mathcal{A}$, then $\mathcal{M}_{\mathbb{K},\mathrm{r}}(X,\mathcal{A})$ is a vector subspace of the normed vector space $\mathcal{M}_{\mathbb{K}}(X,\mathcal{A})$ over $\mathbb{K}$.*

Proof. (a): Suppose μ is regular. Let $A \in \mathcal{A}$ and $\epsilon \in \mathbb{R}^+$. First, assume μ to be positive, i.e. $\mu = |\mu|$. By the regularity of μ, there exist $K, O \subseteq X$, where K is compact, O is open, $K \subseteq A \subseteq O$, $\mu(O \setminus A) < \frac{\epsilon}{2}$, and $\mu(A \setminus K) < \frac{\epsilon}{2}$. Then

$$\mu(O \setminus K) = \mu(O \setminus A) + \mu(A \setminus K) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

as desired. If μ is $\mathbb{K}$-valued, then the positive measures $\mu_1 := (\operatorname{Re}\mu)^+$, $\mu_2 := (\operatorname{Re}\mu)^-$, $\mu_3 := (\operatorname{Im}\mu)^+$, $\mu_4 := (\operatorname{Im}\mu)^-$ are all regular. For each $k \in \{1,\dots,4\}$, choose $K_k, O_k \subseteq X$, where K_k is compact, O_k is open, $K_k \subseteq A \subseteq O_k$, and $\mu_k(O_k \setminus K_k) < \frac{\epsilon}{4}$. Let $K := \bigcup_{k=1}^4 K_k$, $O := \bigcap_{k=1}^4 O_k$. Then K is compact, O is open, $K \subseteq A \subseteq O$, and

$$|\mu|(O \setminus K) \le \sum_{k=1}^{k} \mu_k(O \setminus K) < 4\,\frac{\epsilon}{4} = \epsilon,$$

as desired. For the converse, assume μ to satisfy the condition of (a). Let $A \in \mathcal{A}$ and set

$$M := \inf\{|\mu|(O) : A \subseteq O,\ O \text{ open}\}, \quad m := \sup\{|\mu|(K) : K \subseteq A,\ K \text{ compact}\}.$$

The condition of (a), clearly, implies $m = M$. Since $m \le |\mu|(A) \le M$, this yields $m = |\mu|(A) = M$, showing $|\mu|$ to be regular. Now let $\nu \in \{(\operatorname{Re}\mu)^+, (\operatorname{Re}\mu)^-, (\operatorname{Im}\mu)^+, (\operatorname{Im}\mu)^-\}$.

Then $\nu = |\nu| \leq |\mu|$ (cf. Rem. 4.67(b)). Thus, if $A \in \mathcal{A}$, $\epsilon \in \mathbb{R}^+$ and K, O are given according to the condition of (a), then $\nu(O \setminus K) \leq |\mu|(O \setminus K) < \epsilon$. Thus, by what we have already shown above, ν is regular, showing μ to be regular as well.

(b): Clearly, $0 \in \mathcal{M}_{\mathbb{K},\mathrm{r}}(X, \mathcal{A})$. Let $\mu, \nu \in \mathcal{M}_{\mathbb{K},\mathrm{r}}(X, \mathcal{A})$. We have to show $\mu + \nu$ is regular. Let $A \in \mathcal{A}$ and $\epsilon \in \mathbb{R}^+$. As μ, ν are regular, by (a), there exist $K_\mu, K_\nu, O_\mu, O_\nu \subseteq X$, where K_μ, K_ν are compact, O_μ, O_ν are open, $K_\mu \subseteq A \subseteq O_\mu$, $K_\nu \subseteq A \subseteq O_\nu$, $|\mu|(O_\mu \setminus K_\mu) < \frac{\epsilon}{2}$, and $|\nu|(O_\nu \setminus K_\nu) < \frac{\epsilon}{2}$. Let $K := K_\mu \cup K_\nu$, $O := O_\mu \cap O_\nu$. Then K is compact, O is open, $K \subseteq A \subseteq O$, and

$$|\mu + \nu|(O \setminus K) \leq |\mu|(O \setminus K) + |\nu|(O \setminus K) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

showing $\mu + \nu$ to be regular by (a). Now let $0 \neq s \in \mathbb{K}$ and let A, ϵ, μ be as before. This time, choose $K, O \subseteq X$, where K is compact, O is open, $K \subseteq A \subseteq O$, and $|\mu|(O \setminus K) < \frac{\epsilon}{|s|}$. Then

$$|s\mu|(O \setminus K) = |s||\mu|(O \setminus K) < |s| \frac{\epsilon}{|s|} = \epsilon,$$

showing $s\mu$ to be regular and completing the proof. ■

Example 4.74. Let $(X, \mathcal{T})$ be a T_2 topological space and let $(X, \mathcal{A}, \mu)$ be a Borel measure with μ being $\mathbb{K}$-valued. As $C_c(X) \subseteq \mathcal{L}^1(|\mu|)$ by Lem. 4.60(b) and we can define

$$\alpha_\mu : C_c(X) \longrightarrow \mathbb{K}, \quad \alpha_\mu(f) := \int_X f \, \mathrm{d}\mu \,. \tag{4.70}$$

Then, by Prop. 4.69(a),(b) α_μ constitutes a continuous linear functional on the normed space $(C_c(X), \|\cdot\|_\infty)$, where

$$\underset{f \in \mathcal{L}^\infty(|\mu|)}{\forall} \quad |\alpha_\mu(f)| = \left| \int_X f \, \mathrm{d}\mu \right| \leq \int_X |f| \, \mathrm{d}|\mu| \; \leq \|f\|_\infty \, |\mu|(X) = \|f\|_\infty \, \|\mu\|$$

shows

$$\|\alpha_\mu\| \leq \|\mu\|. \tag{4.71}$$

Proposition 4.75. *Let $(X, \mathcal{T})$ be a topological space.*

(a) *Let $\alpha : (C_c(X, \mathbb{R}), \|\cdot\|_\infty) \longrightarrow \mathbb{R}$ be linear and continuous. Then there exists a unique decomposition*

$$\alpha = \alpha^+ - \alpha^- \tag{4.72}$$

of α into positive *linear maps $\alpha^+, \alpha^- : C_c(X, \mathbb{R}) \longrightarrow \mathbb{R}$ that is* minimal *in the sense that, for each decomposition $\alpha = \beta - \gamma$ with positive linear maps $\beta, \gamma : C_c(X, \mathbb{R}) \longrightarrow \mathbb{R}$, the map $\beta - \alpha^+ = \gamma - \alpha^-$ is positive. Moreover, α^+, α^- are then continuous and*

$$\underset{f \in C_c^+(X, \mathbb{R})}{\forall} \quad \alpha^+(f) = \sup\{\alpha(h) : \, h \in C_c^+(X, \mathbb{R}), \, 0 \leq h \leq f\}. \tag{4.73}$$

(b) *Let $\alpha : (C_c(X,\mathbb{C}), \|\cdot\|_\infty) \longrightarrow \mathbb{C}$ be linear and continuous. Then there exists a unique decomposition*

$$\alpha = \alpha^+ - \alpha^- + i\,(\tilde{\alpha}^+ - \tilde{\alpha}^-) \tag{4.74}$$

of α into positive *linear maps $\alpha^+, \alpha^-, \tilde{\alpha}^+, \tilde{\alpha}^- : C_c(X,\mathbb{C}) \longrightarrow \mathbb{C}$ that is* minimal *in the sense that, for each decomposition $\alpha = \beta - \gamma + i\,(\tilde{\beta} - \tilde{\gamma})$ with positive linear maps $\beta, \gamma, \tilde{\beta}, \tilde{\gamma} : C_c(X,\mathbb{C}) \longrightarrow \mathbb{C}$, the maps $\beta - \alpha^+ = \gamma - \alpha^-$ and $\tilde{\beta} - \tilde{\alpha}^+ = \tilde{\gamma} - \tilde{\alpha}^-$ are positive. Moreover, α^+, α^- $\tilde{\alpha}^+$, $\tilde{\alpha}^-$ are then all continuous.*

Proof. (a): Uniqueness: If the decompositions $\alpha = \alpha^+ - \alpha^-$ and $\alpha = \gamma - \beta$ of α into positive linear maps $\alpha^+, \alpha^-, \beta, \gamma : C_c(X,\mathbb{R}) \longrightarrow \mathbb{R}$ are both minimal in the stated sense, then $\beta - \alpha^+ = \gamma - \alpha^-$ is positive as well as $(-\beta - \alpha^+) = -(\gamma - \alpha^-)$. Thus, $\beta = \alpha^+$ and $\gamma = \alpha^-$ by Rem. 4.58(b). Existence: For $f \in C_c^+(X,\mathbb{R})$, we *define* α^+ by (4.73) and note $0 \leq \alpha^+(f) < \infty$: $\alpha(0) = 0$ yields $\alpha^+(f) \geq 0$, whereas $\alpha^+(f) < \infty$, since the continuous linear map α maps the bounded set $\{\alpha(h) : h \in C_c^+(X,\mathbb{R}),\, 0 \leq h \leq f\} \subseteq B_{\|f\|_\infty}(0)$ into a bounded set. Next, we show

$$\underset{f,g \in C_c^+(X,\mathbb{R})}{\forall} \quad \alpha^+(f+g) = \alpha^+(f) + \alpha^+(g) : \tag{4.75}$$

Let $f, g \in C_c^+(X,\mathbb{R})$. If $h_f, h_g \in C_c^+(X,\mathbb{R})$ with $0 \leq h_f \leq f$ and $0 \leq h_g \leq g$, then $h_f + h_g \in C_c^+(X,\mathbb{R})$ with $0 \leq h_f + h_g \leq f + g$. Thus, $\alpha(h_f) + \alpha(h_g) = \alpha(h_f + h_g) \leq \alpha^+(f+g)$, showing $\alpha^+(f) + \alpha^+(g) \leq \alpha^+(f+g)$. To show the opposite inequality, let $h \in C_c^+(X,\mathbb{R})$ with $0 \leq h \leq f + g$. Define $h_f := (h-g)^+$, $h_g := \min(h,g)$. Then $h_f, h_g \in C_c^+(X,\mathbb{R})$ with

$$0 \leq h_f \leq f \quad \wedge \quad 0 \leq h_g \leq g \quad \wedge \quad h_f + h_g = h.$$

Thus, $\alpha(h) = \alpha(h_f) + \alpha(h_g) \leq \alpha^+(f) + \alpha^+(g)$, showing $\alpha^+(f+g) \leq \alpha^+(f) + \alpha^+(g)$ and proving (4.75). Moreover, we also have

$$\underset{s \in \mathbb{R}_0^+}{\forall} \quad \underset{f \in C_c^+(X,\mathbb{R})}{\forall} \quad \alpha^+(sf) = s\,\alpha^+(f) : \tag{4.76}$$

For $s = 0$, there is nothing to prove. If $s \in \mathbb{R}^+$ and $f \in C_c^+(X,\mathbb{R})$, then

$$\begin{aligned}
\alpha^+(sf) &= \sup\{\alpha(h) : h \in C_c^+(X,\mathbb{R}),\, 0 \leq h \leq sf\} \\
&= \sup\{\alpha(h) : h \in C_c^+(X,\mathbb{R}),\, 0 \leq s^{-1}h \leq f\} \\
&= s\,\sup\{\alpha(s^{-1}h) : h \in C_c^+(X,\mathbb{R}),\, 0 \leq s^{-1}h \leq f\} \\
&= s\,\sup\{\alpha(h) : h \in C_c^+(X,\mathbb{R}),\, 0 \leq h \leq f\} = s\,\alpha^+(f),
\end{aligned}$$

proving (4.76). Now we can extend α^+ to $C_c(X,\mathbb{R})$ by defining

$$\alpha^+ : C_c(X,\mathbb{R}) \longrightarrow \mathbb{R}, \quad \alpha^+(f) := \alpha^+(f^+) - \alpha^+(f^-). \tag{4.77}$$

We claim α^+ to be positive, linear, and continuous: As positivity was built into the definition of α^+, we proceed to show linearity. Let $f,g \in C_c(X,\mathbb{R})$, $h := f+g$. Then

$$h^+ + f^- + g^- = f^+ + g^+ + h^-,$$

such that (4.75) implies

$$\alpha^+(h^+) + \alpha^+(f^-) + \alpha^+(g^-) = \alpha^+(f^+) + \alpha^+(g^+) + \alpha^+(h^-),$$

that means

$$\begin{aligned}\alpha^+(f+g) = \alpha^+(h) = \alpha^+(h^+) - \alpha^+(h^-) &= \alpha^+(f^+) + \alpha^+(g^+) - \alpha^+(f^-) - \alpha^+(g^-)\\ &= \alpha^+(f) + \alpha^+(g).\end{aligned}$$

If $s \in \mathbb{R}_0^+$, then

$$\alpha^+(sf) = \alpha^+(sf^+) - \alpha^+(sf^-) \overset{(4.76)}{=} s\,\alpha^+(f^+) - s\,\alpha^+(f^-) = s\,\alpha^+(f).$$

As also

$$\alpha^+(-f) = \alpha^+(f^-) - \alpha^+(f^+) = -\alpha^+(f),$$

the linearity proof for α^+ is complete. Still considering $f \in C_c(X,\mathbb{R})$, we now estimate

$$|\alpha^+(f)| \le \alpha^+(f^+) + \alpha^+(f^-) \overset{(4.73)}{\le} 2\,\|\alpha\|\,\|f\|_\infty,$$

showing α^+ to be bounded and, thus, continuous. If we now let

$$\alpha^- : C_c(X,\mathbb{R}) \longrightarrow \mathbb{R}, \quad \alpha^- := \alpha^+ - \alpha,$$

then the validity of (4.72) is clear as well as the linearity and continuity of α^-. If $f \in C_c^+(X,\mathbb{R})$, then, by (4.73), $\alpha^+(f) \ge \alpha(f)$, implying $\alpha^-(f) = \alpha^+(f) - \alpha(f) \ge 0$, verifying α^- to be positive. It remains to show the minimality of the decomposition: Let $\alpha = \beta - \gamma$ with positive linear maps $\beta,\gamma : C_c(X,\mathbb{R}) \longrightarrow \mathbb{R}$. Consider $f,h \in C_c(X,\mathbb{R})$ with $0 \le h \le f$. Then $\beta(f) \ge \beta(h) \ge \alpha(h)$. Thus, (4.73) implies $\beta(f) \ge \alpha^+(f)$, showing $\beta - \alpha^+ = \gamma - \alpha^-$ to be positive.

(b): By Lem. 3.2, α satisfies

$$\forall_{f \in C_c(X,\mathbb{C})} \quad \alpha(f) = \operatorname{Re}\alpha(f) - i\,\operatorname{Re}\alpha\big(if\big),$$

where $\operatorname{Re}\alpha$ is $\mathbb{R}$-linear and continuous. By (a), there exists a unique decomposition $(\operatorname{Re}\alpha)\restriction_{C_c(X,\mathbb{R})} = (\operatorname{Re}\alpha)^+ - (\operatorname{Re}\alpha)^-$ of $\operatorname{Re}\alpha$ into positive $\mathbb{R}$-linear maps $(\operatorname{Re}\alpha)^+, (\operatorname{Re}\alpha)^- : C_c(X,\mathbb{R}) \longrightarrow \mathbb{R}$ that is minimal in the sense of (a). If we define

$$\operatorname{Im}\alpha : C_c(X,\mathbb{C}) \longrightarrow \mathbb{R}, \quad \operatorname{Im}\alpha(f) := -\operatorname{Re}\alpha(if),$$

then there also exists a unique decomposition $(\operatorname{Im}\alpha)\restriction_{C_c(X,\mathbb{R})} = (\operatorname{Im}\alpha)^+ - (\operatorname{Im}\alpha)^-$ of $\operatorname{Im}\alpha$ into positive $\mathbb{R}$-linear maps $(\operatorname{Im}\alpha)^+, (\operatorname{Im}\alpha)^- : C_c(X,\mathbb{R}) \longrightarrow \mathbb{R}$ that is minimal in the sense of (a). Uniqueness: If we have a decomposition of α according to (4.74) that is minimal in the stated sense, then $\alpha^+ - \alpha^-$ is a minimal decomposition of $\operatorname{Re}\alpha$ and $\tilde{\alpha}^+ - \tilde{\alpha}^-$ is a minimal decomposition of $\operatorname{Im}\alpha$, such that the uniqueness follows from the uniqueness statement of (a). Existence: Define

$$\begin{aligned}
\alpha^+ &: C_c(X,\mathbb{C}) \longrightarrow \mathbb{C}, & \alpha^+(f) &:= (\operatorname{Re}\alpha)^+(\operatorname{Re} f) + i\,(\operatorname{Re}\alpha)^+(\operatorname{Im} f),\\
\alpha^- &: C_c(X,\mathbb{C}) \longrightarrow \mathbb{C}, & \alpha^-(f) &:= (\operatorname{Re}\alpha)^-(\operatorname{Re} f) + i\,(\operatorname{Re}\alpha)^-(\operatorname{Im} f),\\
\tilde{\alpha}^+ &: C_c(X,\mathbb{C}) \longrightarrow \mathbb{C}, & \tilde{\alpha}^+(f) &:= (\operatorname{Im}\alpha)^+(\operatorname{Re} f) + i\,(\operatorname{Im}\alpha)^+(\operatorname{Im} f),\\
\tilde{\alpha}^- &: C_c(X,\mathbb{C}) \longrightarrow \mathbb{C}, & \tilde{\alpha}^-(f) &:= (\operatorname{Im}\alpha)^-(\operatorname{Re} f) + i\,(\operatorname{Im}\alpha)^-(\operatorname{Im} f).
\end{aligned}$$

Then α^+, α^-, $\tilde{\alpha}^+$, $\tilde{\alpha}^-$ are $\mathbb{C}$-linear: Let $f, g \in C_c(X,\mathbb{C})$ and $r, s \in \mathbb{R}$. Then

$$\begin{aligned}
\alpha^+(f+g) &= (\operatorname{Re}\alpha)^+(\operatorname{Re} f + \operatorname{Re} g) + i\,(\operatorname{Re}\alpha)^+(\operatorname{Im} f + \operatorname{Im} g)\\
&= (\operatorname{Re}\alpha)^+(\operatorname{Re} f) + i\,(\operatorname{Re}\alpha)^+(\operatorname{Im} f) + (\operatorname{Re}\alpha)^+(\operatorname{Re} g) + i\,(\operatorname{Re}\alpha)^+(\operatorname{Im} g)\\
&= \alpha^+(f) + \alpha^+(g)
\end{aligned}$$

as well as

$$\begin{aligned}
\alpha^+((r+is)f) &= (\operatorname{Re}\alpha)^+(r\operatorname{Re} f - s\operatorname{Im} f) + i\,(\operatorname{Re}\alpha)^+(r\operatorname{Im} f + s\operatorname{Re} f)\\
&= r(\operatorname{Re}\alpha)^+(\operatorname{Re} f) - s(\operatorname{Re}\alpha)^+(\operatorname{Im} f)\\
&\quad + i\,\big(r(\operatorname{Re}\alpha)^+(\operatorname{Im} f) + s(\operatorname{Re}\alpha)^+(\operatorname{Re} f)\big)\\
&= (r+is)\alpha^+(f),
\end{aligned}$$

showing α^+ to be $\mathbb{C}$-linear. Analogously, the $\mathbb{C}$-linearity of α^-, $\tilde{\alpha}^+$, $\tilde{\alpha}^-$ follows. The positivity of α^+, α^-, $\tilde{\alpha}^+$, $\tilde{\alpha}^-$ is an immediate consequence of the positivity of $(\operatorname{Re}\alpha)^+$, $(\operatorname{Re}\alpha)^-$, $(\operatorname{Im}\alpha)^+$, $(\operatorname{Im}\alpha)^-$, respectively. Likewise, the continuity of α^+, α^-, $\tilde{\alpha}^+$, $\tilde{\alpha}^-$ is an immediate consequence of the continuity of $(\operatorname{Re}\alpha)^+$, $(\operatorname{Re}\alpha)^-$, $(\operatorname{Im}\alpha)^+$, $(\operatorname{Im}\alpha)^-$, respectively. For each $f \in C_c(X,\mathbb{R})$, we compute

$$\begin{aligned}
\alpha(f) &= \operatorname{Re}\alpha(f) - i\ \operatorname{Re}\alpha\big(if\big)\\
&= (\operatorname{Re}\alpha)^+(f) - (\operatorname{Re}\alpha)^-(f) + i\,\big((\operatorname{Im}\alpha)^+(f) - (\operatorname{Im}\alpha)^-(f)\big)\\
&= \alpha^+(f) - \alpha^-(f) + i\,\big(\tilde{\alpha}^+(f) - \tilde{\alpha}^-(f)\big),
\end{aligned}$$

proving (4.74) on $C_c(X,\mathbb{R})$. However, as both sides of the above equality are $\mathbb{C}$-linear, (4.74) then also holds on $C_c(X,\mathbb{C})$. Finally, the claimed minimality of the decomposition (4.74) follows from $(\operatorname{Re}\alpha)^+ - (\operatorname{Re}\alpha)^-$ being a minimal decomposition of $\operatorname{Re}\alpha$ and $(\operatorname{Im}\alpha)^+ - (\operatorname{Im}\alpha)^-$ being a minimal decomposition of $\operatorname{Im}\alpha$. ∎

Theorem 4.76 (Riesz Representation Theorem)**.** *Let the topological space $(X,\mathcal{T})$ be locally compact and T_2, $\mathcal{B} := \sigma(\mathcal{T})$. Consider the map*

$$\varphi : \mathcal{M}_{\mathbb{K},\mathrm{r}}(X,\mathcal{B}) \longrightarrow (C_{\mathrm{c}}(X))', \quad \varphi(\mu) := \alpha_\mu, \tag{4.78}$$

where α_μ is as in (4.70)*. Then φ is a (linear) isometric isomorphism (in particular, $\mathcal{M}_{\mathbb{K},\mathrm{r}}(X,\mathcal{B})$ is a Banach space).*

Proof. From Ex. 4.74, we know φ to be well-defined; its linearity is a consequence of Prop. 4.70. We verify surjectivity next: Let $\alpha \in (C_{\mathrm{c}}(X))'$. According to Prop. 4.75, α can be written as a linear combination of positive continuous $\mathbb{K}$-valued linear functionals $\alpha_j \in (C_{\mathrm{c}}(X))'$ (in the form (4.72) or in the form (4.74)). For each α_j, the Riesz representation Th. 4.63 yields a positive Borel measure μ_j such that $\alpha_j = \alpha_{\mu_j}$. Moreover, the μ_j are finite by Cor. 4.64 and, thus, regular, by Th. 4.63(ii),(iii). In consequence, $\mu_j \in \mathcal{M}_{\mathbb{K},\mathrm{r}}(X,\mathcal{B})$ with $\varphi(\mu_j) = \alpha_{\mu_j} = \alpha_j$. As φ is linear, this proves surjectivity. Thus, it merely remains to prove that φ is isometric (which, as usual, then also yields φ to be injective as well as the continuity of φ and φ^{-1}). Let $\mu \in \mathcal{M}_{\mathbb{K},\mathrm{r}}(X,\mathcal{B})$. While Ex. 4.74 already yields $\|\varphi(\mu)\| \leq \|\mu\|$, it remains to show the opposite inequality. To this end, fix $\epsilon \in \mathbb{R}^+$. By the definition of $|\mu|$, there exist disjoint sets $A_1,\dots,A_n \in \mathcal{B}$, $n \in \mathbb{N}$, such that

$$\sum_{k=1}^{n} |\mu(A_k)| > |\mu|(X) - \epsilon = \|\mu\| - \epsilon.$$

As μ is regular, $|\mu|$ is regular and there exist compact sets $K_1,\dots,K_n$ such that $K_k \subseteq A_k$ and

$$\sum_{k=1}^{n} |\mu(K_k)| > \|\mu\| - 2\epsilon.$$

Moreover, as a consequence of Prop. 2.5(a), there exist disjoint open sets $O_1,\dots,O_n \subseteq X$ with $K_k \subseteq O_k$. Due to Prop. 4.73(a), we see we can even choose the disjoint open sets $O_1,\dots,O_n \subseteq X$ such that they also satisfy

$$\forall_{k\in\{1,\dots,n\}} \quad |\mu|(O_k \setminus C_k) < \frac{\epsilon}{n}$$

with compact sets $C_1,\dots,C_n$ such that $K_k \subseteq C_k \subseteq A_k$ and still $\sum_{k=1}^{n} |\mu(C_k)| > \|\mu\| - 2\epsilon$ (by possibly making the original $O_1,\dots,O_n$ smaller). We now use Prop. 4.62 to conclude

$$\forall_{k\in\{1,\dots,n\}} \quad \exists_{f_k \in C_{\mathrm{c}}(X)} \quad \Big(0 \leq f_k \leq 1 \ \wedge \ f_k{\restriction}_{C_k} \equiv 1 \ \wedge \ \operatorname{supp} f_k \subseteq O_k\Big).$$

Define

$$f := \sum_{k\in I_\mu} \frac{\overline{\mu(C_k)}}{|\mu(C_k)|}\, f_k, \quad I_\mu := \big\{k \in \{1,\dots,n\} : \mu(C_k) \neq 0\big\}.$$

As the O_k are disjoint, we have $\|f\|_\infty \leq 1$. Due to

$$\left|\sum_{k=1}^{n}\int_{C_k} f \,\mathrm{d}\mu\right| = \left|\int_X f\,\mathrm{d}\mu - \sum_{k=1}^{n}\int_{O_k\setminus C_k} f\,\mathrm{d}\mu\right| \stackrel{\text{Prop. 4.69(b)}}{\leq} \left|\int_X f\,\mathrm{d}\mu\right| + \sum_{k=1}^{n}\int_{O_k\setminus C_k}|f|\,\mathrm{d}|\mu|,$$

one estimates

$$\begin{aligned}|\varphi(\alpha)(f)| = \left|\int_X f\,\mathrm{d}\mu\right| &\geq \left|\sum_{k=1}^{n}\int_{C_k} f\,\mathrm{d}\mu\right| - \sum_{k=1}^{n}\int_{O_k\setminus C_k}|f|\,\mathrm{d}|\mu| \\ &\geq \left|\sum_{k\in I_\mu}\int_{C_k}\frac{\overline{\mu(C_k)}}{|\mu(C_k)|}\,\mathrm{d}\mu\right| - \epsilon = \sum_{k=1}^{n}|\mu(C_k)| - \epsilon > \|\mu\| - 3\epsilon.\end{aligned}$$

As $\epsilon \in \mathbb{R}^+$ was arbitrary, this shows $\|\varphi(\mu)\| \geq \|\mu\|$ and concludes the proof. ■

Corollary 4.77. *Let the topological space* $(K,\mathcal{T})$ *be compact and* T_2, $\mathcal{B} := \sigma(\mathcal{T})$. *Then the map*

$$\varphi: \mathcal{M}_{\mathbb{K},\mathrm{r}}(K,\mathcal{B}) \longrightarrow (C(K))', \quad \varphi(\mu) := \alpha_\mu,$$

where α_μ *is as in* (4.70)*, constitutes a (linear) isometric isomorphism (in particular,* $\mathcal{M}_{\mathbb{K},\mathrm{r}}(K,\mathcal{B})$ *is a Banach space).*

Proof. Since $(K,\mathcal{T})$ is compact, $C_c(K) = C(K)$, and the corollary is merely a special case of Th. 4.76. ■

Corollary 4.78. *Let* $a,b \in \mathbb{R}$, $a<b$. *Let* $(f_k)_{k\in\mathbb{N}}$ *be a bounded sequence in* $(C[a,b],\|\cdot\|)$. *Then the following statements are equivalent:*

(i) $(f_k)_{k\in\mathbb{N}}$ *converges weakly to* 0.

(ii) $(f_k)_{k\in\mathbb{N}}$ *converges pointwise to* 0.

Proof. "(i)⇒(ii)": If $f_k \rightharpoonup 0$, then

$$\forall_{\alpha\in(C[a,b])'} \quad \lim_{k\to\infty}\alpha(f_k) = 0.$$

Thus, as $[a,b]$ is a compact T_2 space, Cor. 4.77 implies

$$\forall_{\mu\in\mathcal{M}_{\mathbb{K},\mathrm{r}}([a,b],\mathcal{B}^1)} \quad \lim_{k\to\infty}\int_{[a,b]} f_k\,\mathrm{d}\mu = 0. \tag{4.79}$$

According to Ex. 4.72(b), $\delta_t \in \mathcal{M}_{\mathbb{K},\mathrm{r}}([a,b],\mathcal{B}^1)$ for each $t \in [a,b]$, implying

$$\underset{t\in[a,b]}{\forall} \quad \lim_{k\to\infty} f_k(t) = \lim_{k\to\infty} \int_{[a,b]} f_k \,\mathrm{d}\delta_t = 0,$$

i.e. $f_k \to 0$ pointwise (note that we did not actually used the boundedness of the sequence for this direction).

"(ii)⇒(i)": As $[a,b]$ is a compact T_2 space, $f_k \rightharpoonup 0$ is actually equivalent to (4.79). Thus, it remains to show that $f_k \to 0$ pointwise implies (4.79). Let $\mu \in \mathcal{M}_{\mathbb{K},\mathrm{r}}([a,b],\mathcal{B}^1)$ Since $(f_k)_{k\in\mathbb{N}}$ is bounded, there exists $M \in \mathbb{R}_0^+$ such that, for each $k \in \mathbb{N}$, $|f_k| \leq g \cong M$, where g is $|\mu|$-integrable, since $|\mu|$ is finite. Thus, if $f_k \to 0$ pointwise, then (4.79) holds by the dominated convergence theorem. ■

A Exhaustion by Compact Sets

Definition A.1. Let $(X,\mathcal{T})$ be a topological space, $A \subseteq X$. Let $(K_i)_{i\in\mathbb{N}}$ be a sequence of compact subsets of A. Then this sequence is called an *exhaustion by compact sets* of A if, and only if, it satisfies the following two conditions:

$$A = \bigcup_{i\in\mathbb{N}} K_i, \tag{A.1a}$$

$$\underset{i\in\mathbb{N}}{\forall} \quad K_i \subseteq K_{i+1}^\circ \tag{A.1b}$$

(where (A.1b) says that K_i lies in the interior of K_{i+1}).

Theorem A.2. *Let $n \in \mathbb{N}$, and let $O \subseteq \mathbb{K}^n$ be open. Then there exists an exhaustion by compact sets of O.*

Proof. If $O = \mathbb{K}^n$, then the closed balls $(\overline{B}_i(0))_{i\in\mathbb{N}}$ (with respect to some fixed norm, e.g., $\|\cdot\|_2$), clearly, form an exhaustion by compact sets of O. If $O \neq \mathbb{K}^n$, then the function

$$d: O \longrightarrow \mathbb{R}_0^+, \quad d(x) := \operatorname{dist}(x, O^{\mathrm{c}}), \tag{A.2}$$

is well-defined and continuous (cf. [Phi16b, Ex. 2.6(b)]). Define

$$\underset{i\in\mathbb{N}}{\forall} \quad K_i := \overline{B}_i(0) \cap d^{-1}\big([\tfrac{1}{i},\infty[\big). \tag{A.3}$$

Then $(K_i)_{i\in\mathbb{N}}$ is a sequence of compact subsets of O. We show the sequence to be an exhaustion by compact sets of O: If $x \in O$, then it is immediate that there exists $i_1 \in \mathbb{N}$ such that $x \in \overline{B}_{i_1}(0)$. Since O is open, there also exists $i_2 \in \mathbb{N}$ such that $x \in d^{-1}\big([\frac{1}{i_2},\infty[\big)$. Thus $x \in K_{\max\{i_1,i_2\}}$, showing (A.1a) to hold (with A replaced by O). If $x \in K_i$, $i \in \mathbb{N}$, then $\operatorname{dist}(x,0) < i+1$ and $\operatorname{dist}(x,O^{\mathrm{c}}) > \frac{1}{i+1}$, showing $x \in K_{i+1}^\circ$. Thus, (A.1b) holds, completing the proof. ■

B Interchanging Derivatives with Pointwise Limits

While interchanging derivatives with pointwise limits is not always admissible, it does work if the derivatives converge uniformly:

Theorem B.1. *Let $a, b \in \mathbb{R}$, $a < b$. For each $n \in \mathbb{N}$, let $f_n : [a,b] \longrightarrow \mathbb{K}$ be continuously differentiable. Let $f : [a,b] \longrightarrow \mathbb{K}$ and assume $f_n \to f$ pointwise in $[a,b]$. If there exists $g : [a,b] \longrightarrow \mathbb{K}$ such that $f_n' \to g$* uniformly *in $[a,b]$, then f is continuously differentiable with $f' = g$, i.e.*

$$\forall_{x \in [a,b]} \quad f'(x) = \Big(\lim_{n\to\infty} f_n\Big)'(x) = \lim_{n\to\infty} f_n'(x) = g(x). \tag{B.1}$$

Proof. Since the f_n' are continuous, the uniform convergence to g implies g to be continuous, and we can compute, for each $x \in [a,b]$,

$$\begin{aligned} \int_a^x g(t)\,\mathrm{d}t &\overset{\text{[Phi17, Prop. G.4]}}{=} \lim_{n\to\infty} \int_a^x f_n'(t)\,\mathrm{d}t \overset{\text{[Phi16a, Th. 10.20(b)]}}{=} \lim_{n\to\infty} \big(f_n(x) - f_n(a)\big) \\ &= f(x) - f(a). \end{aligned} \tag{B.2}$$

Since g is continuous, [Phi16a, Th. 10.20(a)] implies the left-hand side of (B.2) to be differentiable with respect to x with derivative g. In consequence, (B.2) implies f to be differentiable with $f' = g$ as desired. ■

C Weierstrass Approximation Theorem

Theorem C.1 (Weierstrass Approximation Theorem)**.** *Let $a, b \in \mathbb{R}$ with $a < b$. For each continuous function $f \in C[a,b]$ and each $\epsilon > 0$, there exists a polynomial $p : \mathbb{R} \longrightarrow \mathbb{R}$ such that $\|f - p\restriction_{[a,b]}\|_\infty < \epsilon$, where $p\restriction_{[a,b]}$ denotes the restriction of p to $[a,b]$.*

Theorem C.1 will be a corollary of the fact that the Bernstein polynomials corresponding to $f \in C[0,1]$ (see Def. C.2) converge uniformly to f on $[0,1]$ (see Th. C.3 below).

Definition C.2. Given $f : [0,1] \longrightarrow \mathbb{R}$, define the *Bernstein polynomials* $B_n f$ corresponding to f by

$$B_n f : \mathbb{R} \longrightarrow \mathbb{R}, \quad (B_n f)(x) := \sum_{\nu=0}^{n} f\left(\frac{\nu}{n}\right) \binom{n}{\nu} x^\nu (1-x)^{n-\nu} \quad \text{for each } n \in \mathbb{N}. \tag{C.1}$$

Theorem C.3. *For each $f \in C[0,1]$, the sequence of Bernstein polynomials $(B_n f)_{n\in\mathbb{N}}$ corresponding to f according to Def.* C.2 *converges uniformly to f on $[0,1]$, i.e.*

$$\lim_{n\to\infty} \|f - (B_n f)\restriction_{[0,1]}\|_\infty = 0. \tag{C.2}$$

Proof. We begin by noting

$$(B_n f)(0) = f(0) \quad \text{and} \quad (B_n f)(1) = f(1) \quad \text{for each } n \in \mathbb{N}. \tag{C.3}$$

For each $n \in \mathbb{N}$ and $\nu \in \{0,\dots,n\}$, we introduce the abbreviation

$$q_{n\nu}(x) := \binom{n}{\nu} x^\nu (1-x)^{n-\nu}. \tag{C.4}$$

Then

$$1 = \big(x + (1-x)\big)^n = \sum_{\nu=0}^{n} q_{n\nu}(x) \quad \text{for each } n \in \mathbb{N} \tag{C.5}$$

implies

$$f(x) - (B_n f)(x) = \sum_{\nu=0}^{n} \left(f(x) - f\left(\frac{\nu}{n}\right)\right) q_{n\nu}(x) \quad \text{for each } x \in [0,1],\ n \in \mathbb{N},$$

and

$$\big|f(x) - (B_n f)(x)\big| \le \sum_{\nu=0}^{n} \left|f(x) - f\left(\frac{\nu}{n}\right)\right| q_{n\nu}(x) \quad \text{for each } x \in [0,1],\ n \in \mathbb{N}. \tag{C.6}$$

As f is continuous on the compact interval $[0,1]$, it is uniformly continuous, i.e. for each $\epsilon > 0$, there exists $\delta > 0$ such that, for each $x \in [0,1]$, $n \in \mathbb{N}$, $\nu \in \{0,\dots,n\}$:

$$\left|x - \frac{\nu}{n}\right| < \delta \quad \Rightarrow \quad \left|f(x) - f\left(\frac{\nu}{n}\right)\right| < \frac{\epsilon}{2}. \tag{C.7}$$

For the moment, we fix $x \in [0,1]$ and $n \in \mathbb{N}$ and define

$$N_1 := \left\{\nu \in \{0,\dots,n\} : \left|x - \frac{\nu}{n}\right| < \delta\right\},$$
$$N_2 := \left\{\nu \in \{0,\dots,n\} : \left|x - \frac{\nu}{n}\right| \ge \delta\right\}.$$

Then

$$\sum_{\nu\in N_1} \left|f(x) - f\left(\frac{\nu}{n}\right)\right| q_{n\nu}(x) \le \frac{\epsilon}{2} \sum_{\nu\in N_1} q_{n\nu}(x) \le \frac{\epsilon}{2} \sum_{\nu=0}^{n} q_{n\nu}(x) = \frac{\epsilon}{2}, \tag{C.8}$$

and with $M := \|f\|_\infty$,

$$\sum_{\nu\in N_2}\left|f(x)-f\left(\frac{\nu}{n}\right)\right|q_{n\nu}(x) \leq \sum_{\nu\in N_2}\left|f(x)-f\left(\frac{\nu}{n}\right)\right|q_{n\nu}(x)\frac{\left(x-\frac{\nu}{n}\right)^2}{\delta^2}$$
$$\leq \frac{2M}{\delta^2}\sum_{\nu=0}^{n} q_{n\nu}(x)\left(x-\frac{\nu}{n}\right)^2. \tag{C.9}$$

To compute the sum on the right-hand side of (C.9), observe

$$\left(x-\frac{\nu}{n}\right)^2 = x^2 - 2x\,\frac{\nu}{n} + \left(\frac{\nu}{n}\right)^2 \tag{C.10}$$

and

$$\sum_{\nu=0}^{n}\binom{n}{\nu}x^\nu(1-x)^{n-\nu}\frac{\nu}{n} = x\sum_{\nu=1}^{n}\binom{n-1}{\nu-1}x^{\nu-1}(1-x)^{(n-1)-(\nu-1)} \overset{\text{(C.5)}}{=} x \tag{C.11}$$

as well as

$$\sum_{\nu=0}^{n}\binom{n}{\nu}x^\nu(1-x)^{n-\nu}\left(\frac{\nu}{n}\right)^2 = \frac{x}{n}\sum_{\nu=1}^{n}(\nu-1)\binom{n-1}{\nu-1}x^{\nu-1}(1-x)^{(n-1)-(\nu-1)} + \frac{x}{n}$$
$$= \frac{x^2}{n}(n-1)\sum_{\nu=2}^{n}\binom{n-2}{\nu-2}x^{\nu-2}(1-x)^{(n-2)-(\nu-2)} + \frac{x}{n}$$
$$= x^2\left(1-\frac{1}{n}\right) + \frac{x}{n} = x^2 + \frac{x}{n}(1-x). \tag{C.12}$$

Thus, we obtain

$$\sum_{\nu=0}^{n} q_{n\nu}(x)\left(x-\frac{\nu}{n}\right)^2 \overset{\text{(C.10),(C.5),(C.11),(C.12)}}{=} x^2\cdot 1 - 2x\cdot x + x^2 + \frac{x}{n}(1-x)$$
$$\leq \frac{1}{4n} \quad \text{for each } x\in[0,1],\ n\in\mathbb{N},$$

and together with (C.9):

$$\sum_{\nu\in N_2}\left|f(x)-f\left(\frac{\nu}{n}\right)\right|q_{n\nu}(x) \leq \frac{2M}{\delta^2}\,\frac{1}{4n} < \frac{\epsilon}{2} \quad \text{for each } x\in[0,1],\ n > \frac{M}{\delta^2\epsilon}. \tag{C.13}$$

Combining (C.6), (C.8), and (C.13) yields

$$\left|f(x)-(B_nf)(x)\right| < \frac{\epsilon}{2}+\frac{\epsilon}{2} = \epsilon \quad \text{for each } x\in[0,1],\ n > \frac{M}{\delta^2\epsilon},$$

proving the claimed uniform convergence. ∎

Proof of Th. C.1. Define

$$\phi: [a,b] \longrightarrow [0,1], \quad \phi(x) := \frac{x-a}{b-a},$$
$$\phi^{-1}: [0,1] \longrightarrow [a,b], \quad \phi^{-1}(x) := (b-a)x + a.$$

Given $\epsilon > 0$, and letting

$$g: [0,1] \longrightarrow \mathbb{R}, \quad g(x) := f\big(\phi^{-1}(x)\big),$$

Th. C.3 provides a polynomial $q: \mathbb{R} \longrightarrow \mathbb{R}$ such that $\|g - q{\restriction_{[0,1]}}\|_\infty < \epsilon$. Defining

$$p: \mathbb{R} \longrightarrow \mathbb{R}, \quad p(x) := q\big(\phi(x)\big) = q\left(\frac{x-a}{b-a}\right)$$

(having extended ϕ in the obvious way) yields a polynomial p such that

$$|f(x) - p(x)| = \big|g(\phi(x)) - q(\phi(x))\big| < \epsilon \quad \text{for each } x \in [a,b]$$

as needed. ■

D Topological Invariants

[Phi16b, Prop. D.7] already provides several topological invariants (i.e. properties preserved under homeomorphisms). The following Prop. D.1 provides additional topological invariants relevant to the present class:

Proposition D.1. *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces, let $f: X \longrightarrow Y$ be a homeomorphism, $A \subseteq X$.*

(a) *$(X, \mathcal{T}_X)$ is locally compact if, and only if, $(Y, \mathcal{T}_Y)$ is locally compact.*

(b) *A is nowhere dense (resp. of the first category, resp. of the second category) in X if, and only if, $f(A)$ is nowhere dense (resp. of the first category, resp. of the second category) in Y.*

Proof. We will make use of the topological invariants already proved in [Phi16b, Prop. D.7]. Since f is a homeomorphism if, and only if, f^{-1} is a homeomorphism, it always suffices to prove one direction of the claimed equivalences.

(a): If $(X, \mathcal{T}_X)$ is locally compact and $x \in X$, then there exists a compact $C \in \mathcal{U}(x)$. Then $f(C)$ is a compact neighborhood of $f(x)$, showing $(Y, \mathcal{T}_Y)$ to be locally compact.

(b): If A is nowhere dense, then $(\overline{A})^\circ = \emptyset$, implying

$$\left(\overline{f(A)}\right)^\circ = \left(f\left(\overline{A}\right)\right)^\circ = f\left((\overline{A})^\circ\right) = f(\emptyset) = \emptyset,$$

showing $f(A)$ to be nowhere dense. If A is of the first category, then $A = \bigcup_{k=1}^\infty A_k$ with nowhere dense sets A_k, $k \in \mathbb{N}$. Then $f(A) = \bigcup_{k=1}^\infty f(A_k)$ with $f(A_k)$ nowhere dense, showing $f(A)$ to be of the first category. If A is of the second category, then A is not of the first category, i.e. $f(A)$ is not of the first category, i.e. $f(A)$ is of the second category. ■

E Orthogonalization

Theorem E.1 (Gram-Schmidt Orthogonalization)**.** *Let $\big(X, \langle\cdot,\cdot\rangle\big)$ be an inner product space with induced norm $\|\cdot\|$. Let $x_0, x_1, \dots$ be a finite or infinite sequence of vectors in X. Define $v_0, v_1, \dots$ recursively as follows:*

$$v_0 := x_0, \quad v_n := x_n - \sum_{\substack{k=0,\\ v_k \neq 0}}^{n-1} \frac{\langle x_n, v_k\rangle}{\|v_k\|^2}\, v_k \tag{E.1}$$

for each $n \in \mathbb{N}$, additionally assuming that n is less than or equal to the max index of the sequence $x_0, x_1, \dots$ if the sequence is finite. Then the sequence $v_0, v_1, \dots$ constitutes an orthogonal system. Of course, by omitting the $v_k = 0$ and by dividing each $v_k \neq 0$ by its norm, one can also obtain an orthonormal system (nonempty if at least one $v_k \neq 0$). Moreover, $v_n = 0$ if, and only if, $x_n \in \operatorname{span}\{x_0, \dots, x_{n-1}\}$. In particular, if the $x_0, x_1, \dots$ are all linearly independent, then so are the $v_0, v_1, \dots$.

Proof. We show by induction on n, that, for each $0 \leq m < n$, $v_n \perp v_m$. For $n = 0$, there is nothing to show. Thus, let $n > 0$ and $0 \leq m < n$. By induction, $\langle v_k, v_m\rangle = 0$ for each $0 \leq k, m < n$ such that $k \neq m$. For $v_m = 0$, $\langle v_n, v_m\rangle = 0$ is clear. Otherwise,

$$\langle v_n, v_m\rangle = \left\langle x_n - \sum_{\substack{k=0,\\ v_k \neq 0}}^{n-1} \frac{\langle x_n, v_k\rangle}{\|v_k\|^2}\, v_k,\ v_m \right\rangle = \langle x_n, v_m\rangle - \frac{\langle x_n, v_m\rangle}{\|v_m\|^2}\, \langle v_m, v_m\rangle = 0,$$

thereby establishing the case. So we know that $v_0, v_1, \dots$ constitutes an orthogonal system. Next, by induction, for each n, we obtain $v_n \in \operatorname{span}\{x_0, \dots, x_n\}$ directly from (E.1). Thus, $v_n = 0$ implies $x_n = \sum_{\substack{k=0,\\ v_k \neq 0}}^{n-1} \frac{\langle x_n, v_k\rangle}{\|v_k\|^2}\, v_k \in \operatorname{span}\{x_0, \dots, x_{n-1}\}$. Conversely, if

$x_n \in \operatorname{span}\{x_0, \dots, x_{n-1}\}$, then

$$\begin{aligned}\dim \operatorname{span}\{v_0, \dots, v_{n-1}, v_n\} &= \dim \operatorname{span}\{x_0, \dots, x_{n-1}, x_n\} = \dim \operatorname{span}\{x_0, \dots, x_{n-1}\} \\ &= \dim \operatorname{span}\{v_0, \dots, v_{n-1}\},\end{aligned}$$

which implies $v_n = 0$. Finally, if all $x_0, x_1, \dots$ are linearly independent, then all $v_k \neq 0$, $k = 0, 1, \dots$, such that the $v_0, v_1, \dots$ are linearly independent. ■

F Partition of Unity

Theorem F.1. *Let the topological space $(X, \mathcal{T})$ be locally compact and T_2. Moreover, let $O_1, \dots, O_N \subseteq X$ be open, $N \in \mathbb{N}$, and let $K \subseteq X$ be compact. If $(O_1, \dots, O_N)$ forms an open cover of K, then there exists a corresponding* partition of unity*, i.e. there exist functions $\varphi_1, \dots, \varphi_N \in C_c(X)$ such that*

$$\underset{i \in \{1,\dots,N\}}{\forall} \Big(0 \leq \varphi_i \leq 1 \ \wedge \ \operatorname{supp} \varphi_i \subseteq O_i\Big) \tag{F.1}$$

and

$$\underset{x \in K}{\forall} \quad \sum_{i=1}^{N} \varphi_i(x) = 1. \tag{F.2}$$

Proof. For each $x \in K$, there exists $i(x) \in \{1, \dots, N\}$ such that $x \in O_{i(x)}$. Then, by Prop. 2.5(a), there exists an open set $V_x \subseteq X$ such that $\overline{V}_x$ is compact and

$$x \in V_x \subseteq \overline{V}_x \subseteq O_{i(x)}.$$

As K is compact, there exist $x_1, \dots, x_M \in K$ such that $K \subseteq V_{x_1} \cup \dots \cup V_{x_M}$, $M \in \mathbb{N}$. Define

$$\underset{i \in \{1,\dots,N\}}{\forall} \left(J(i) := \big\{k \in \{1, \dots, M\} : \overline{V}_{x_k} \subseteq O_i\big\}, \quad H_i := \bigcup_{k \in J(i)} \overline{V}_{x_k} \subseteq O_i \right).$$

Then each H_i is compact and Prop. 4.62 provides $f_i \in C_c(X)$, satisfying

$$0 \leq f_i \leq 1 \ \wedge \ f_i\restriction_{H_i} \equiv 1 \ \wedge \ \operatorname{supp} f_i \subseteq O_i.$$

From the f_i, we can now define

$$\underset{i \in \{1,\dots,N\}}{\forall} \quad \varphi_i := f_i \prod_{j=1}^{i-1} (1 - f_j)$$

(recall the convention that empty products are defined to be 1). Then, clearly, $\varphi_i \in C_c(X)$ for each $i \in \{1, \dots, N\}$. Since $\operatorname{supp} \varphi_i \subseteq \operatorname{supp} f_i \subseteq O_i$ and since, as a product of $[0,1]$-valued functions, φ_i is $[0,1]$-valued, we see (F.1) to be satisfied. To verify (F.2), we prove

$$\forall_{j \in \{1,\dots,N\}} \quad \sum_{i=1}^{j} \varphi_i = 1 - \prod_{i=1}^{j} (1 - f_i) \tag{F.3}$$

via induction on j: For $j = 1$, we have $1-(1-f_1) = f_1 = \varphi_1$ as required. For $1 \leq j < N$, we compute

$$\sum_{i=1}^{j+1} \varphi_i \stackrel{\text{ind.hyp.}}{=} \varphi_{j+1} + 1 - \prod_{i=1}^{j} (1 - f_i) = f_{j+1} \prod_{i=1}^{j} (1 - f_i) + 1 - \prod_{i=1}^{j} (1 - f_i) = 1 - \prod_{i=1}^{j+1} (1 - f_i),$$

proving (F.3). Finally, if $x \in K$, then there exists $k \in \{1, \dots, M\}$ such that $x \in V_{x_k} \subseteq \overline{V}_{x_k} \subseteq O_{i(x_k)}$, implying $x \in H_{i(x_k)}$. Thus, $f_{i(x_k)}(x) = 1$ and (F.3) yields $\sum_{i=1}^{N} \varphi_i(x) = 1$, proving (F.2) and the theorem. ■

References

[Els07] Jürgen Elstrodt. *Maß- und Integrationstheorie*, 5th ed. Grundwissen Mathematik, Springer-Verlag, Berlin, 2007 (German).

[Phi16a] P. Philip. *Analysis I: Calculus of One Real Variable.* Lecture Notes, LMU Munich, 2015/2016, AMS Open Math Notes Ref. # OMN:202109.111306, available in PDF format at
`https://www.ams.org/open-math-notes/omn-view-listing?listingId=111306`.

[Phi16b] P. Philip. *Analysis II: Topology and Differential Calculus of Several Variables.* Lecture Notes, LMU Munich, 2016, AMS Open Math Notes Ref. # OMN:202109.111307, available in PDF format at
`https://www.ams.org/open-math-notes/omn-view-listing?listingId=111307`.

[Phi17] P. Philip. *Analysis III: Measure and Integration Theory of Several Variables.* Lecture Notes, LMU Munich, 2016/2017, AMS Open Math Notes Ref. # OMN:202109.111308, available in PDF format at
`https://www.ams.org/open-math-notes/omn-view-listing?listingId=111308`.

[Rud73] W. Rudin. *Functional Analyis.* McGraw-Hill Book Company, New York, 1973.

[Rud87] W. Rudin. *Real and Complex Analysis*, 3rd ed. McGraw-Hill Book Company, New York, 1987.

[Wer11] D. Werner. *Funktionalanalysis*, 7th ed. Springer-Verlag, Berlin, 2011 (German).