1 Notation and terminology

- $\mathbb{N} = \{0, 1, 2, \ldots\}$, $\mathbb{N}_1 = \mathbb{N} \setminus \{0\}$, $\mathbb{R}_+ = [0, \infty)$.
- The symbol $\mathbb{K}$ is reserved to denote the field of real or complex numbers.
- All vector spaces are over the field $\mathbb{R}$ or $\mathbb{C}$. If the field of a vector space is not specified, it is denoted by $\mathbb{K}$.
- The terms *function* and *mapping* are treated as synonyms, whereas the term *operator* will be used only in reference to $\mathbb{R}$-linear mappings. A *functional* is a scalar-valued function.
- For any linear operator T between two vector spaces $\mathcal{N}(T)$ and $\mathcal{R}(T)$ stand for, respectively, the kernel and the range (the image) of T.
- Neighbourhoods do not need to be open. Subspaces do not need to be linear. Convex sets may be empty. Vector spaces can be trivial (that is, they can contain only their origins). Linear isometries do not need to be surjective.
- A *Banach* space is a normed vector space whose norm is complete.
- As a rule (unless otherwise stated), the norm on a vector space X is denoted by $\|\cdot\|_X$.
- *NVS* and TVS are abbreviations for, respectively, a *normed vector space* and a *topological vector space.*
- Open and closed *unit balls* in a NVS X, to be denoted by (resp.) B_X and $\bar{B}_X$, are the sets $\{x \in X\colon \ \|x\|_X < 1\}$ and $\{x \in X\colon \ \|x\|_X \le 1\}$. Similarly, balls in metric spaces are denoted by $B(a, r)$ and $\bar{B}(a, r)$. A *unit* vector in a NVS is a vector of norm 1.
- Finite Cartesian products of topological spaces are equipped with the product topologies.

2 Preliminaries

2.1 Definition.
Let X be a vector space over $\mathbb{K}$, and $A \subset \mathbb{K}$ and $B, C \subset X$ be arbitrary subsets. We define sets $A \cdot B$ and $B + C$ of X by the formulas:

$$A \cdot B \overset{\text{def}}{=} \{t \cdot u\colon \ t \in A, \ u \in B\}, \qquad B + C \overset{\text{def}}{=} \{b + c\colon \ b \in B, \ c \in C\}.$$

So, both $A \cdot B$ and $B + C = C + B$ are subsets of X.

Similarly, for any $t \in \mathbb{K}$, $v \in X$ and $B \subset X$, $t \cdot B \overset{\text{def}}{=} \{t\} \cdot B$ and $v + B = B + v \overset{\text{def}}{=} \{v\} + B$.

A subset W of X is said to be:

- *convex* if $tW + (1 - t)W \subset W$ for all $t \in [0, 1]$;
- *symmetric* if $-W = W$;
- *balanced* if $\bar{B}_{\mathbb{K}} \cdot W \subset W$;
- *absolutely convex* if W is convex and $\gamma W = W$ for any $\gamma \in \mathbb{K}$ with $|\gamma| = 1$.

In particular, W is absolutely convex iff it is both convex and balanced. Similarly, if $\mathbb{K} = \mathbb{R}$, then W is absolutely convex iff it is both convex and symmetric.

2.2 Proposition.
Let X be a normed vector space.

(A) *The mapping*

$$\mathbb{K} \times X \times X \ni (t, x, y) \mapsto tx + y \in X \tag{2:1}$$

is continuous.

(B) *For any $a \in X$ and $r > 0$, $B_X(a, r) = a + rB_X$, $\bar{B}_X(a, r) = a + r\bar{B}_X$; $\overline{B_X(a, r)} = \bar{B}_X(a, r)$ and $\operatorname{int} \bar{B}_X(a, r) = B_X(a, r)$.*

(proof—exercise)

Part (A) of the above basic result serves as a defining condition for a more general notion (than normed vectors spaces):

2.3 Definition.
A *topological vector space* is a pair (X, τ) where X is a vector space and τ is a topology on X such that the mapping (2:1) is continuous. T_2VS will stand for a Hausdorff TVS.
A TVS is *normable* if its topology is induced by a certain norm defined on the underlying vector space.
A *0-neighbourhood* in a TVS X is a neighbourhood of the origin of X.

Basic properties of TVS's are listed below.

2.4 Proposition.
Let X be a topological vector space.

(A) *For any $a \in X$ and $t \in \mathbb{K} \setminus \{0\}$, the function $X \ni x \mapsto tx + a \in X$ is a homeomorphism.*

(B) *For any 0-neighbourhood U in X there exists a 0-neighbourhood V that is both open and balanced and satisfies $V + V \subset U$.*

(C) *The space X is both contractible and locally contractible as a topological space.*

(D) *X is T_2VS iff it is T_0, iff X is T_3, iff the set $\{0\}$ is closed in X.*

Recall that a topological space M is

- *contractible* if there are $c \in M$ and a continuous mapping $H\colon M \times [0,1] \to M$ such that $H(x,0) = c$ and $H(x,1) = x$ for all $x \in M$;
- *locally contractible* if for any point $a \in M$ and its neighbourhood U in M there exists a neighbourhood $V \subset U$ of a and a continuous mapping $H\colon V \times [0,1] \to U$ such that $H(x,0) = a$ and $H(x,1) = x$ for all $x \in V$.

In particular, a (locally) contractible space is (locally) arcwise connected.

Proof of Proposition 2.4. Item (A) is left to the reader as a simple exercise.

To show (B), fix a 0-neighbourhood U. It follows from the continuity of the function $\mathbb{K} \times \mathbb{K} \times X \times X \ni (p,q,x,y) \mapsto px + qy \in X$ at $(0,0,0,0)$ that there exist $\varepsilon > 0$ and an open 0-neighbourhood W in X such that $(\varepsilon B_{\mathbb{K}}) \cdot W + (\varepsilon B_{\mathbb{K}}) \cdot W \subset U$. We define V as $(\varepsilon B_{\mathbb{K}}) \cdot W$. It is easily seen that V is balanced and satisfies $V + V \subset U$. Finally, V is open because $0 \in W$ and hence

$$V = \bigcup_{\substack{p \in \varepsilon B_{\mathbb{K}} \\ p \neq 0}} pW,$$

and each of the above sets pW is open, by (A).

We pass to (C). Observe that the (continuous) function $X \times [0,1] \ni (x,t) \mapsto tx \in X$ witnesses the contractibility of X. To show its local contractibility, fix a neighbourhood U of a point $a \in X$ and choose a balanced 0-neigbourhood W such that $W \subset U - a$ (we use here both items (A) and (B)). Then $V \stackrel{\text{def}}{=} W + a$ is a neighbourhood of a contained in U such that $(1-t)a + tx \in V$ for all $x \in V$ and $t \in [0,1]$. Consequently, the function $V \times [0,1] \ni (x,t) \mapsto (1-t)a + tx \in U$ is well defined and continuous, and we are done.

Now assume X is T_0 and choose arbitrary non-zero vector $v \in X$. It follows from T_0 axiom that there exists an open set U such that $U \cap \{0, v\}$ is a one-point set. If $v \notin U$, then $0 \notin U' \stackrel{\text{def}}{=} v - U$ and U' is an open neighbourhood of v. This shows that the set $X \setminus \{0\}$ is open in X. So, it follows from (A) that X is T_1. Now fix a closed set A in X and a vector $u \notin A$. Then $B \stackrel{\text{def}}{=} A - u$ is a closed set that does not contain the origin of X. We infer from (B) that there is an open symmetric 0-neighbourhood V such that $(V + V) \cap B = \varnothing$. Equivalently, $V \cap (B + V) = \varnothing$. Observe that $B + V$ is open (since $B + V = \bigcup_{b \in B}(b + V)$) and contains B (as $0 \in V$). So, the sets $u + V \ni u$ and $u + B + V \supset A$ are open and disjoint, as we wished. □

2.5 ***Remark.***
As it is well-known, each topological vector space is a topological group and every T_0-topological group is actually $T_{3\frac{1}{2}}$. On the other hand, there are known examples of locally convex TVS's that are not T_4. We will not use any of these properties in this textbook.

For the sake of completeness, we now list general properties of the class of all topological vector spaces.

2.6 Proposition.

(A) *Cartesian products of (families of arbitrary size of) TVS's equipped with the product topologies are TVS's as well.*

(B) *Cartesian products of a finite number of NVS's are normable.*

(C) *A linear subspace of a TVS equipped with the induced topology is a TVS.*

(D) *For any linear subspace F of an arbitrary TVS E, the quotient space E/F is a TVS when equipped with the quotient topology. Moreover, in that case the quotient map $\pi\colon E \to E/F$ is open; and E/F is T_2 iff F is closed.*

(E) *A linear operator $T\colon E \to F$ between two TVS's is continuous iff it is continuous at the origin of E.*

(F) *For any linear operator $T\colon X \to Y$ between two TVS's denote by $\pi_T\colon X \to X/\mathcal{N}(T)$ the quotient map and by $\tilde{T}\colon X/\mathcal{N}(T) \to Y$ the unique linear operator such that $T = \tilde{T} \circ \pi_T$. Then T is continuous (resp. open) iff so is $\tilde{T}$.*

Proof. We skip all boring technical details and here we prove only selected parts.

If E, F and π are as specified in (D) and U is an open subset of E, then $\pi^{-1}(\pi(U)) = U + F = \bigcup_{f \in F}(U + f)$. Consequently, $\pi^{-1}(\pi(U))$ is open in E and therefore $\pi(U)$ is open in E/F as well. Further, E/F is T_2VS iff its origin forms a closed set, iff $F = \pi^{-1}(\{0\})$ is closed in E.

Now assume that T, π_T and $\tilde{T}$ are as specified in (F). Since π_T is both continuous and open (by (D)), it is clear that T has any of these two properties provided so has $\tilde{T}$. Conversely, if T is continuous and U is an open subset of Y, then the set $\tilde{T}^{-1}(U) = \pi_T(T^{-1}(U))$ is also open. Similarly, if T is open and V is an open set in E/F, then the set $\tilde{T}(V) = T(\pi_T^{-1}(V))$ is open in Y as well. $\square$

2.7 ***Remark.***
It is less trivial to show that $\mathbb{K}^\omega$ is non-normable. (We leave it as a more difficult exercise.) In particular, the product of an infinite family of non-trivial NVS's is never normable.

Further properties of normed vector spaces are established in the two results stated below.

2.8 Theorem.
Let F be a closed linear subspace of a normed vector space E and let $\pi\colon E \to E/F$ denote the quotient map. Then the formula

$$\|b\|_{E/F} \stackrel{\text{def}}{=} \inf\{\|a\|_E\colon\ a \in E,\ \pi(a) = b\}$$

defines a norm on E/F that is compatible with the quotient topology. Moreover, if E is a Banach space, the above norm is complete.

Proof. A verification that $\|\cdot\|_{E/F}$ is indeed a norm is left to the reader. It follows from the very definition of $\|\cdot\|_{E/F}$ that $\pi(B_E) = B_{E/F}$ and $\|\pi(a)\|_{E/F} \leq \|a\|_E$ for each $a \in E$. In particular, π is both continuous and open when considered as a function from $(E, \|\cdot\|_E)$ into $(E/F, \|\cdot\|_{E/F})$. So, we infer from item (F) of Proposition 2.6 that the identity map on E/F, considered as a function from the quotient topology into the $\|\cdot\|_{E/F}$-norm topology is both continuous and open. So, it is a homeomorphism, which finishes the first part of the proof.

Now assume that, in addition, E is a Banach space. Consider an arbitrary sequence $(b_n)_{n=1}^\infty$ of vectors of E/F such that $\|b_n\|_{E/F} < 2^{-n}$ for all $n > 0$. To show that the norm $\|\cdot\|_{E/F}$ is complete, it is enough to prove that the series $\sum_{n=1}^\infty b_n$ converges in E/F. To this end, for each n we choose $a_n \in E$ such that $\pi(a_n) = b_n$ and $\|a_n\|_E < 2^{-n}$. Since E is Banach, the series $\sum_{n=1}^\infty a_n$ converges in E, say to $p \in E$. Then $\pi(p) = \pi(\lim_{n\to\infty} \sum_{k=1}^n a_k) = \lim_{n\to\infty} \sum_{k=1}^n \pi(a_n) = \lim_{n\to\infty} \sum_{k=1}^n b_k$, and we are done. $\square$

2.9 Proposition. (Banach)
For a linear operator $T\colon X \to Y$ between normed vector spaces X and Y the following conditions are equivalent:

(i) *T is continuous;*

(ii) *T is Lipschitz;*

(iii) *the set $T(\bar{B}_X)$ is bounded in $\|\cdot\|_Y$.*

Proof. If (iii) holds and $(a_n)_{n=1}^{\infty}$ is a sequence of non-zero vectors in X that converge to zero, then the vectors $b_n = T(a_n/\|a_n\|_X)$ form a bounded sequence in Y. Hence $\|T(a_n)\|_Y = \|a_n\|_X \cdot \|b_n\|_Y \to 0$ $(n \to \infty)$, which proves that (iii) is followed by (i). Conversely, if (iii) does not hold, then $\lim_{n\to\infty} \|T(a_n)\|_Y = \infty$ for a suitable sequence $(a_n)_{n=1}^{\infty} \subset B_X$. Without loss of generality, we may and do assume that $T(a_n) \neq 0$. Then the vectors $c_n \stackrel{\text{def}}{=} \frac{a_n}{\|T(a_n)\|_Y}$ converge to the origin of X, but $\|T(c_n)\|_Y = 1$ and thus T is discontinuous. So, conditions (i) and (iii) are equivalent. Since T is linear, it is easily seen that also (ii) and (iii) are equivalent, which finishes the proof. ☐

2.10 Definition.
Let X and Y be normed vector spaces over the same field $\mathbb{K}$. The set $\mathscr{L}(X,Y)$ of all continuous linear operators from X into Y becomes a vector space over $\mathbb{K}$ with pointwise operations. It is equipped with the *operator norm* defined as follows:

$$\|T\| = \|T\|_{\text{op}} = \sup\{\|T(x)\|_Y\colon\ x \in \bar{B}_X\} \qquad (T \in \mathscr{L}(X,Y)).$$

One proves in a standard manner that $\mathscr{L}(X,Y)$ is a Banach space provided so is Y. One writes $\mathscr{L}(X)$ in place of $\mathscr{L}(X,X)$.

When $Y = \mathbb{K}$, the space $\mathscr{L}(X,\mathbb{K})$ is denoted by X^* (or X') and called the *dual Banach space* of X (or simply the *dual*). Similarly, the Banach space $X^{**} \stackrel{\text{def}}{=} (X^*)^*$ is called the *bidual* of X.

Linear operators between normed vector spaces that are continuous are (equivalently) called *bounded*.

2.11 ***Remark.***
It is easy to check that for any continuous linear operator $T\colon X \to Y$ between normed vector spaces, $\|T\|$ is the least Lipschitz constant of T; and that $\|T\| = \sup\{\|T(x)\|_Y : x \in X,\ \|x\|_X = 1\}$, provided X is non-trivial.

2.12 Proposition.
Let X_0 be a dense linear subspace of a normed vector space X and Y be a Banach space. Every bounded linear operator $T_0\colon X_0 \to Y$ admits a unique extension $T\colon X \to Y$ that is bounded and linear as well. Moreover, $\|T\| = \|T_0\|$; and if T_0 is isometric, so is T.

Proof. Since T_0 is Lipschitz and Y is a complete metric space, T_0 is extendable to a Lipschitz map $T\colon X \to Y$ with the same Lipschitz constant. It is easy to check that T is linear, and isometric if T_0 is so (exercise). The uniqueness of T follows from the density of X_0 in X. ☐

It turns out that a finite-dimensional vector space admits a unique Hausdorff topology that makes it a TVS, as shown by:

2.13 Theorem.
Let E be a non-trivial finite-dimensional $\mathrm{T_2VS}$ *and $e_1, \ldots, e_n$ be a fixed basis of the vector space E. Then the topology of E coincides with the topology induced by the norm $\|\cdot\|_1$ on E given by*

$$\Big\|\sum_{k=1}^{n} \alpha_k e_k\Big\|_1 \stackrel{\text{def}}{=} \sum_{k=1}^{n} |\alpha_k| \qquad (\alpha_1, \ldots, \alpha_n \in \mathbb{K}).$$

Proof. Denote by τ_0 and τ_1, respectively, the given topology on E and the one induced by the norm $\|\cdot\|_1$. Since the mapping $\mathbb{K}^n \ni (\alpha_1, \dots, \alpha_n) \mapsto \sum_{k=1}^n \alpha_k e_k \in (E, \tau_0)$ is continuous, we infer that the identity map from (E, τ_1) into (E, τ_0) is continuous as well (equivalently, $\tau_0 \subset \tau_1$). So, it remains to check that the identity map in the reverse direction is also continuous (it is sufficient to check the continuity only at 0). To this end, we fix $\varepsilon > 0$ and consider the sphere $S(\varepsilon) \stackrel{\text{def}}{=} \{x \in E\colon\ \|x\|_1 = \varepsilon\}$. Since $K = S(\varepsilon)$ is compact w.r.t. τ_1, it is so w.r.t. τ_0. Thus, $E \setminus K$ is a 0-neighbourhood in τ_0 (here we make use of the Hausdorff separation axiom). So, it follows from Proposition 2.4 that there exists a balanced 0-neighbourhood U in (E, τ_0) that is disjoint from K. In particular, for any $x \in U$, $\{tx\colon t \in [0,1]\} \cap K = \varnothing$ and thus $\|x\|_1 < \varepsilon$. This shows that $U \subset B_{\|\cdot\|_1}(0, \varepsilon)$ and implies the continuity at the origin of the identity map, which finishes the proof. □

As immediate consequences of the above theorem, we obtain the following two results.

2.14 Corollary.
A finite-dimensional linear subspace of a T_2VS *is closed.*

Proof. Let X be a T_2VS and F its finite-dimensional linear subspace. Take any vector $u \in X$ that belongs to the closure of F and consider the linear subspace $E \stackrel{\text{def}}{=} \mathbb{K}u + F$. Since E is finite-dimensional, it follows from Theorem 2.13 that E in normable (in the induced topology). In particular, F is closed in E, as a complete subset in the metric induced by a norm on E, and thus $u \in F$. □

2.15 Corollary.
Let $T\colon X \to Y$ be a linear operator between two T_2VS*'s. If the range of T is finite-dimensional, then T is continuous iff the kernel of T is closed.*

Proof. The necessity (of the closedness of the kernel) is clear. To see its sufficiency, we may and do assume (by replacing Y by the range of T) that Y has finite dimension. Then $T = \tilde{T} \circ \pi$ where $\pi\colon X \to X/\mathcal{N}(T)$ is the quotient map and $\tilde{T}\colon X/\mathcal{N}(T) \to Y$ is a (uniquely determined) linear operator. Since $E \stackrel{\text{def}}{=} X/\mathcal{N}(T)$ is finite-dimensional and Hausdorff, we infer from Theorem 2.13 that both E and Y are normable and, consequently, $\tilde{T}$ is continuous. Therefore T is continuous as well and we are done. □

Theorem 2.13 shows that, in particular, all finite-dimensional T_2VS's are locally compact. That there are no other locally compact T_2VS's is shown by

2.16 Theorem.
A locally compact T_2VS *is finite-dimensional.*

Proof. Fix a compact 0-neighbourhood K in a T_2VS E and choose an open 0-neighbourhood V such that $V + V \subset K$. Since all translations of V cover K, the compactness of K yields the existence of a finite non-empty set S such that $K \subset V + S$. Denote by F the linear span of S. We will show that $E = F$, which will finish the proof. To this end, observe that $K + K \subset (V + S) + (V + S) \subset (V + V) + F \subset K + F$. So, $K + K \subset K + F$. Now simple induction argument shows that

$$\underbrace{K + \ldots + K}_{n} \subset K + F \qquad \text{for any } n > 0. \tag{2:2}$$

Fix for a moment an arbitrary vector $x \in E$. It follows from the continuity at 0 of the function $\mathbb{R} \ni t \mapsto tx \in X$ that $\frac{1}{n}x \in K$ for some integer $n > 0$. But then x belongs to the left-hand side of (2:2). We conclude that $E = K + F$. This formula implies that the quotient space E/F is compact. Being T_2VS (by Corollary 2.14), E/F has to be trivial [why?—exercise!] and therefore $E = F$. □

*2.17 **Remark**.*
The property that an infinite-dimensional NVS is not locally compact may be proven by a more direct argument, due to F. Riesz. It is based on the following simple observation:

(∗) *If F is a closed proper linear subspace of a NVS E, then for any $\varepsilon \in (0,1)$ there exists a unit vector $u \in E$ such that* $\operatorname{dist}(u, E) \geq \varepsilon$.

Assuming (∗) holds, one defines inductively unit vectors $v_1, v_2, \ldots$ in E such that $\|v_j - v_k\|_E \geq \frac{1}{2}$ for all distinct j and k. (We get v_n from (∗) applied to $F_n = \operatorname{lin}\{v_j\colon\ j < n\}$ $[F_1 = \{0\}]$ and $\varepsilon = \frac{1}{2}$.) This sequence witnesses non-compactness of $\bar{B}_E$. Of course, then no other closed ball in E can be compact.

The property (⋆) simply follows from Theorem 2.8 (recall that the open unit ball in F/E is covered by the open unit ball of E via the quotient map).

2.18 Theorem. (Defining a TVS by a basis of [possibly non-open] 0-neighbourhoods)
Let E be a vector space and $\mathscr{O}$ a non-empty collection of subsets of E such that for any $W \in \mathscr{O}$ the following conditions are satisfied:

(TV0) $0 \in W$ *(equivalently: $W \neq \varnothing$);*

(TV1) *there exists $V \in \mathscr{O}$ such that $V + V \subset W$;*

(TV2) *there exists $U \in \mathscr{O}$ such that $\bar{B}_{\mathbb{K}} \cdot U \subset W$;*

(TV3) *W is* absorbing*; that is, $(0, \infty) \cdot W = E$.*

Then there exists a unique topology τ on E such that (E, τ) is a topological vector space and $\mathscr{O}$ is a basis of [possibly non-open] 0-neighbourhoods in (E, τ). Moreover,

- *$U \subset E$ belongs to τ iff for any $x \in U$ there is $V \in \mathscr{O}$ such that $x + V \subset U$;*
- *(E, τ) is* $\mathrm{T_2VS}$ *iff $\bigcap \mathscr{O} = \{0\}$.*

(without proof)

3 Classical Banach and normed spaces

3.1 Sequence spaces

- $\boldsymbol{\ell_p} = \boldsymbol{\ell_p^{\mathbb{K}}}$ where $p \in [1, \infty]$

(ℓ_p-P1) For any sequence $(x_n)_{n=1}^\infty \subset \mathbb{K}$,

$$\|\,(x_n)_{n=1}^\infty\,\|_p \overset{\text{def}}{=} \begin{cases} \left(\sum_{n=1}^\infty |x_n|^p\right)^{1/p} & \text{if } p < \infty \\ \sup_{n>0} |x_n| & \text{if } p = \infty \end{cases} \in [0, \infty].$$

(ℓ_p-P2) $\ell_p \overset{\text{def}}{=} \{(x_n)_{n=1}^\infty \in \mathbb{K}\colon\ \|\,(x_n)_{n=1}^\infty\,\|_p < \infty\}$.

(ℓ_p-P3) $(\ell_p, \|\cdot\|_p)$ is a Banach space; it is separable iff $p < \infty$.

(ℓ_p-P4) For finite p, ℓ_p^* is linearly isometric to ℓ_q where $q = \infty$ for $p = 1$ and $q = \frac{p}{p-1}$ otherwise.

(ℓ_p-P5) $\|\cdot\|_p$ is a *standard* or *classical* norm on ℓ_p (it is the default norm on that space). The triangle inequality for that norm with $1 < p < \infty$ is known as the *Minkowski's inequality*.

(ℓ_p-P6) For q defined as in (ℓ_p-P4) and any two sequences $(x_n)_{n=1}^\infty, (y_n)_{n=1}^\infty \subset \mathbb{K}$ the following inequality, known as the *Hölder's inequality*, holds:

$$\sum_{n=1}^\infty |x_n y_n| \leq \|\,(x_n)_{n=1}^\infty\,\|_p \cdot \|\,(y_n)_{n=1}^\infty\,\|_q.$$

- $\boldsymbol{c_0}$ and $\boldsymbol{c}$

(c-P1) c is the subspace of ℓ_∞ consisting of all convergent sequences.

(c-P2) c_0 is the subspace of c of all *null* sequences; that is, of all sequences that converge to 0.

(c-P3) Both c_0 and c are equipped with the norm $\|\cdot\|_\infty$. (This is a *standard*, *classical* and default norm on these spaces.)

(c-P4) Both $(c_0, \|\cdot\|_\infty)$ and $(c, \|\cdot\|)$ are separable Banach spaces.

(c-P5) Both the duals c_0^* and c^* are linearly isometric to ℓ_1.

(c-P6) c and c_0 are isomorphic Banach spaces, but they are not linearly isometric.

- c_{00}

(c_{00}-P1) It is the subspace of c_0 of all sequences that are eventually 0; that is, $(x_n)_{n=1}^\infty \subset \mathbb{K}$ belongs to c_{00} if $x_n = 0$ for all sufficiently large n.

(c_{00}-P2) c_{00} is a dense subspace of c_0 and all ℓ_p with finite p. It is dense neither in c nor ℓ_∞.

(c_{00}-P3) There is **no** standard or classical or default norm on c_{00}.

(c_{00}-P4) There does not exist a complete norm on c_{00}. This follows from a more general statement which asserts that no vector space over $\mathbb{R}$ or $\mathbb{C}$ whose linear dimension is $\aleph_0$ (that is, whose Hamel basis is countably infinite) admits a complete norm and is a direct consequence of the Baire category theorem. (Indeed, each such a space E can be written as $\bigcup_{n=1}^\infty F_n$ where each F_n is finite-dimensional. In any norm on E all these subspaces F_n are closed. So, if there were a complete norm on E, we would infer from the Baire theorem that $\mathrm{int}(F_n) \neq \varnothing$ for some n; this would easily imply that $F_n = E$.)

(c_{00}-P5) c_{00} is separable in every norm (exercise).

3.2 Function spaces

- $C([0,1])$ and $C(K)$-spaces

($C(K)$-P1) For any non-empty compact Hausdorff space K (e.g., for $K = [0,1]$), $C(K) = C(K, \mathbb{K})$ consists of all continuous scalar-valued functions defined on K.

($C(K)$-P2) $C(K)$ is equipped with the sup-norm, commonly denoted by one of the symbols: $\|\cdot\|_\infty$, $\|\cdot\|_{\sup}$, $\|\cdot\|_{\max}$, $\|\cdot\|_K$ ($\|f\|_\infty = \sup_{x\in K} |f(x)|$). It is a *standard* / *classical* / default norm of that space.

($C(K)$-P3) $(C(K), \|\cdot\|_\infty)$ is a Banach space. It is separable iff K is metrisable.

($C(K)$-P4) [Riesz(-Markov-Kakutani) representation theorem] The dual $C(K)^*$ is linearly isometric to the Banach space of all regular scalar-valued Borel measures on K, equipped with the total variation norm. More precisely, for any continuous linear functional ϕ on $C(K)$ there exists a unique regular scalar-valued Borel measure μ such that

$$\phi(f) = \int_K f \, \mathrm{d}\mu \qquad (f \in C(K)).$$

What is more, $\|\mu\| = \|\phi\|$.

($C(K)$-P5) [Banach-Mazur theorem] Every separable normed vector space over $\mathbb{K}$ is linearly isometric to a linear subspace of $C([0,1], \mathbb{K})$ (see Theorem 8.19 (p. 41) in Chapter 8).

- $C_0(X)$

($C_0(X)$-P1) For any locally compact Hausdorff space X, the space $C_0(X) = C_0(X, \mathbb{K})$ consists of all continuous scalar-valued functions defined on X that vanish at infinity; that is, a continuous mapping $f\colon X \to \mathbb{K}$ belongs to $C_0(X)$ iff for any $\varepsilon > 0$ there is a compact set $K \subset X$ such that $|f(x)| \leq \varepsilon$ for any $x \notin K$.

($C_0(X)$-P2) By default, $C_0(X)$ is equipped with its *standard* / *classical* norm $\|\cdot\|_\infty$ (cf. ($C(K)$-P2)).

($C_0(X)$-P3) $(C_0(X), \|\cdot\|_\infty)$ is a Banach space. It is separable iff X is second countable.

($C_0(X)$-P4) The dual of $C_0(X)$ is characterised in the same way as for compact X (cf. ($C(K)$-P4)).

- $\ell_\infty(X) = B(X)$ for arbitrary non-empty set X

($B(X)$-P1) For any non-empty set X, the space $\ell_\infty(X) = \ell_\infty^{\mathbb{K}}(X) = B(X) = B(X, \mathbb{K})$ consists of all bounded scalar-valued functions defined on X.

($B(X)$-P2) By default, $\ell_\infty(X)$ is equipped with its *standard* / *classical* norm $\|\cdot\|_\infty$ (cf. ($C(K)$-P2)).

($B(X)$-P3) $(\ell_\infty(X), \|\cdot\|_\infty)$ is a Banach space. It is separable iff the set X is finite (in that case this Banach space is finite-dimensional).

($B(X)$-P4) $\ell_\infty(X)^*$ is linearly isometric to the Banach space of all finitely additive scalar-valued measures (defined on the whole power set of X) with finite total variation (as a norm). (There is also other classical description of this dual: since $\ell_\infty(X)$ is linearly isometric to a certain space of the form $C(K)$, the space $\ell_\infty(X)^*$ is characterised by ($C(K)$-P4). Here K is the so-called Čech-Stone compactification of the discrete topological space X.)

- $C^k([0,1])$ with finite $k > 0$

(C^k-P1) The space $C^k([0,1]) = C^k([0,1],\mathbb{K})$ of all scalar-valued function of class C^k (where k is finite!) on $[0,1]$ is equipped with the topology of uniform convergence of all derivatives up to k. More precisely, functions $f_1, f_2, \ldots \in C^k([0,1])$ converge (in the *standard* topology of that space) to $g \in C^k([0,1])$ iff the functions $f_1^{(j)}, f_2^{(j)}, \ldots$ converge uniformly to $g^{(j)}$ for $j = 0, \ldots, k$.

(C^k-P2) Although $C^k([0,1])$ has a standard topology (defined above), there is no single standard or classical norm on that space—we can speak about a "collection of standard norms" instead. Each of the norms of the form $\|f\|_p^{(k)} = (\sum_{j=0}^k \|f^{(j)}\|_\infty^p)^{1/p}$ (where $p \in [1,\infty)$) as well as $\|f\|_\infty^{(k)} = \max(\|f\|_\infty, \ldots, \|f^{(k)}\|_\infty)$ belong to this collection and any norm from this collection is compatible with the standard topology of that space. (There is no standard notation for the norms introduced here.)

(C^k-P3) Equipped with any of the norms described in the above item, $C^k([0,1])$ is a separable Banach space. It is (quite naturally) isomorphic to $\mathbb{K}^k \times C([0,1])$ (exercise), so its dual is well-known.

(C^k-P4) It is a more difficult exercise to show that there does not exist a norm on $C^\infty([0,1])$ that induces the topology of uniform convergence of all derivatives. Although $C^\infty([0,1])$ has a standard topology (just aforementioned), it is not normable. Also, there is no standard or classical norm topology on $C^\infty([0,1])$.

3.3 "Pseudofunction" spaces = L^p spaces

- For any non-negative measure μ on a measurable space $(\Omega, \mathfrak{M})$, one considers a natural equivalence relation "$\sim_\mu$" of μ-almost everywhere equality on the set $\mathscr{M}(\Omega,\mathfrak{M})$ of all scalar-valued $\mathfrak{M}$-measurable functions defined on Ω. (So, $f \sim_\mu g$ or, as is commonly written, $f = g$ μ-a.e., if $\mu(\{\omega \in \Omega\colon\ f(\omega) \neq g(\omega)\}) = 0$.) The quotient space $L(\Omega,\mu) \stackrel{\text{def}}{=} \mathscr{M}(\Omega,\mathfrak{M})/\sim_\mu$ has a natural structure of a vector space over $\mathbb{K}$ and the assignemt $\mathscr{M}(\Omega,\mathfrak{M}) \ni f \mapsto \int_\Omega |f|\,\mathrm{d}\mu$ (as well as $f \mapsto \int_\Omega f\,\mathrm{d}\mu$ for μ-integrable f) is constant on equivalence classes from $L(\Omega,\mu)$. So, to simplify the notation, we write $\int_\Omega |f|\,\mathrm{d}\mu$ (and $\int_\Omega f\,\mathrm{d}\mu$ for f being the equivalence class of a μ-integrable function) for any $f \in L(\Omega,\mu)$ to denote this common value of the respective integrals corresponding to functions from f.
- For $f \in L(\Omega,\mu)$ and $p \in [1,\infty]$ the following quantity is also well defined:

$$\|f\|_p \stackrel{\text{def}}{=} \begin{cases} \left(\int_\Omega |f|^p\,\mathrm{d}\mu\right)^{1/p} & \text{if } p < \infty, \\ \operatorname{ess\,sup}|f| & \text{if } p = \infty \end{cases}$$

  where, for real-valued g, $\operatorname{ess\,sup} g \stackrel{\text{def}}{=} \inf\{t \in \mathbb{R}\colon\ \mu(g^{-1}((t,\infty))) = 0\}$ (with the convention that $\inf(\varnothing) = \infty$).
- $L^p(\mu) = L^p(\Omega,\mu) = L^p(\Omega,\mathbb{K})$ consists of all $f \in L(\Omega,\mu)$ such that $\|f\|_p < \infty$.
- $(L^p(\mu), \|\cdot\|_p)$ is a Banach space and $\|\cdot\|_p$ is its *standard* / *classical* / default norm. For any σ-finite Borel measure μ on a Polish space (that is, on a completely metrisable separable space), $L^p(\mu)$ is separable for finite p. $L^\infty(\mu)$ is hardly ever separable.
- When μ is the one-dimensional Lebesgue measure on an interval I, we will write $L^p(I)$ instead of $L^p(\mu)$.
- In most classical cases (that is, when μ vanishes at all points of Ω), $L^p(\mu)$ is not a function space. The last statement means that the assignment $f \mapsto f(\omega)$ for arbitrarily fixed $\omega \in \Omega$ does not define a function on $L^p(\mu)$.
- For $1 < p < \infty$, the dual $L^p(\mu)^*$ is linearly isometric to $L^q(\mu)$ where $q = \frac{p}{p-1}$. (For σ-finite μ this is a classical result; for other measures the proof is more subtle but is based on the σ-finite case.)
- Very often $L^1(\mu)^*$ is linearly isometric to $L^\infty(\mu)$—it is the case for σ-finite measures. However, in its full generality this statement is false.
- For all L^p spaces, analogs of the Minkowski's and Hölder's inequalities hold. Both of them are fundamental tools in studies of L^p spaces. Another useful tool is the classical Radon-Nikodym theorem.

4 Hilbert spaces

4.1 Definition.
A *sesquilinear form* is any function $\phi\colon V \times W \to \mathbb{K}$ (where V and W are two vector spaces over the same field) that satisfies the following two conditions:

(SF1) $\phi(\cdot, w)$ is linear for any $w \in W$;

(SF2) $\phi(v, \cdot)$ is antilinear for any $v \in V$; that is, it is additive and $\phi(v, cu) = \bar{c}\phi(v,u)$ for all $c \in \mathbb{K}$ and $u \in W$.

If, in addition, $W = V$ and ϕ satisfies

$$\phi(u,v) = \overline{\phi(v,u)} \qquad (u, v \in V),$$

it is called a *symmetric sesquilinear form* or a *Hermitian form* on V.

A Hermitian form ϕ on V such that for all non-zero $u \in V$,

$$\phi(u,u) > 0$$

(resp. $\phi(u,u) \geq 0$) is said to be an *inner product* or a *scalar product* (resp. a *semi-inner product*) on V and V is called an *inner product space* or a *scalar product space*.

Inner products on a vector space V are commonly denoted by $\langle \cdot, - \rangle_V$ or simply $\langle \cdot, - \rangle$.

4.2 ***Remark.***
In the realm of real vector spaces, "sesquilinear" is a synonym of "bilinear"; and Hermitian forms coincide with symmetric bilinear ones.

4.3 Theorem. (Schwarz(-Cauchy-Bunyakovsky-. . .) inequality)
Let ϕ be a semi-inner product on a vector space. Then for all $u, v \in V$:

$$|\phi(u,v)|^2 \leq \phi(u,u)\phi(v,v). \tag{4:1}$$

Moreover, if ϕ is an inner product, then the above inequality becomes an equation iff u and v are linearly dependent.

Proof. Fix u and v and choose unit scalar γ such that $|\phi(u,v)| = \gamma\phi(u,v)$. Everywhere in this proof t is a real number.

It follows that $\phi(\gamma u + tv, \gamma u + tv) \geq 0$. Equivalently (since $\phi(a+b, a+b) = \phi(a,a) + 2\operatorname{Re}\phi(a,b) + \phi(b,b)$),

$$\phi(u,u) + 2|\phi(u,v)|t + \phi(v,v)t^2 \geq 0 \qquad (\forall t \in \mathbb{R}). \tag{4:2}$$

Since all the coefficients in the above inequality are non-negative, we conclude that either:

- $\phi(v,v) = \phi(u,v) = 0$; or
- $\phi(v,v) > 0$ and the discriminant of the above quadratic form is non-positive; that is, $4|\phi(u,v)|^2 - 4\phi(u,u)\phi(v,v) \leq 0$.

It is clear that in both the above cases (4:1) holds.

Now assume that ϕ is inner and that (4:1) becomes an equation. Then either $\phi(v,v) = 0$ and, consequently, $v = 0$; or else the discriminant of (4:2) vanishes, which implies that this left-hand side of this inequality has a root, say $s \in \mathbb{R}$. But then $\phi(\gamma u + sv, \gamma u + sv) = 0$ and hence $u = -\bar{\gamma}sv$. Anyway, u and v are linearly dependent. The reverse implication is left to the reader. □

4.4 Corollary.
If ϕ is a [semi-]inner product on V, then the assignment $v \mapsto \sqrt{\phi(v,v)}$ defines a [semi-]norm $\|\cdot\|_V$ on V. If ϕ is an inner product and $u, v \in V$ are such that $\|u+v\|_V = \|u\|_V + \|v\|_V$, then $u = tv$ or $v = tu$ for some $t \geq 0$.

Proof. We only need to check the triangle inequality. To this end, we fix $u, v \in V$ and square both sides of the inequality $\sqrt{\phi(u+v, u+v)} \leq \sqrt{\phi(u,u)} + \sqrt{\phi(v,v)}$ to reduce it to $2\operatorname{Re}\phi(u,v) \leq 2\sqrt{\phi(u,u)\phi(v,v)}$, which is implied by the Schwarz inequality (4:1).

Now assume that ϕ is an inner product and $\|u+v\|_V = \|u\|_V + \|v\|_V$. Then, continuing the above argument, we get that

$$\operatorname{Re}\phi(u,v) = \sqrt{\phi(u,u)\phi(v,v)}. \tag{4:3}$$

In particular, $|\phi(u,v)|^2 = \phi(u,u)\phi(v,v)$. So, u and v are linearly dependent. We may and do assume that $v \neq 0$. Then $u = cv$ for some $c \in \mathbb{K}$. Substituting this formula to (4:3), we obtain $\operatorname{Re} c \cdot \phi(v,v) = |c|\phi(v,v)$ and, consequently, $c \geq 0$. □

4.5 Definition.
Each inner product space $(V, \langle\cdot,-\rangle_V)$ is equipped with the norm $\|\cdot\|_V$ given by $\|v\|_V \overset{\text{def}}{=} \sqrt{\langle v,v\rangle_V}$. If this norm is complete, we call $(V, \|\cdot\|_V)$ a *Hilbert space*. In other words, a Hilbert space is a Banach space whose norm comes from a certain inner product.

4.6 Proposition.
Let u and v be two vectors in an inner product space $(V, \langle\cdot,-\rangle_V)$.

(Par) [***Parallelogram identity***] $\|u+v\|_V^2 + \|u-v\|_V^2 = 2\|u\|_V^2 + 2\|v\|_V^2$.

($\mathbb{R}$P) [***Polarization identity for*** $\mathbb{K}=\mathbb{R}$] $\langle u,v\rangle_V = \frac{1}{4}(\|u+v\|_V^2 - \|u-v\|_V^2)$ *provided that* $\mathbb{K}=\mathbb{R}$.

($\mathbb{C}$P) [***Polarization identity for*** $\mathbb{K}=\mathbb{C}$] $\langle u,v\rangle_V = \frac{1}{4}(\|u+v\|_V^2 - \|u-v\|_V^2 + i\|u+iv\|_V^2 - i\|u-iv\|_V^2)$ *provided that* $\mathbb{K}=\mathbb{C}$.

(*proof—exercise*)

It turns out that the Parallelogram identity, introduced in the above result, characterises norms that come from inner products, as shown by

4.7 Theorem. (Jordan–von Neumann)
Let $\|\cdot\|_V$ be a norm on a vector space V that satisfies the Parallelogram identity (see (Par) *in Proposition* 4.6*). Then the formula* ($\mathbb{K}$P) *defines an inner product on V.*

(*without proof*)

4.8 Proposition.

(A) *For any semi-norm p on a vector space E the set $p^{-1}(\{0\})$ is a linear subspace of E and the formula $\|\pi(x)\|_F \overset{\text{def}}{=} p(x)$ $(x \in E)$ correctly defines a norm $\|\cdot\|_F$ on $F \overset{\text{def}}{=} E/p^{-1}(\{0\})$ where $\pi\colon E \to F$ is the canonical projection.*

(B) *The metric completion of a normed vector space admits a natural Banach space structure that extends the structure of the given space.*

(C) *For any semi-inner product space (E,ϕ) the formula $\langle\pi(x),\pi(y)\rangle_F \overset{\text{def}}{=} \phi(x,y)$ correctly defines an inner product on $F \overset{\text{def}}{=} E/\{x\in E\colon\ \phi(x,x)=0\}$ where $\pi\colon E\to F$ is the canonical projection.*

(D) *The metric completion of an inner product space admits a natural Hilbert space structure that extends the structure of the given space.*

(*proof—exercise*)

4.9 Example.

(A) ℓ_2 is a Hilbert space, as well as all spaces of the form $\ell_2(X)$. Indeed, the following formula defines a (standard) inner product on that space:

$$\langle (x_n)_{n=1}^\infty, (y_n)_{n=1}^\infty\rangle_{\ell_2} \overset{\text{def}}{=} \sum_{n=1}^\infty x_n\bar{y}_n.$$

(B) More generally, for any non-negative measure μ (on a set Ω), $L^2(\mu)$ is a Hilbert space. A standard inner product on $L^2(\mu)$ is given by

$$\langle f,g\rangle_{L^2(\mu)} \overset{\text{def}}{=} \int_\Omega f(\omega)\overline{g(\omega)}\,\mathrm{d}\mu(\omega).$$

The following result is a starting point in the Hilbert space theory.

4.10 Theorem. (Best approximation in Hilbert spaces)
Let $(H, \|\cdot\|_H)$ be a Hilbert space. Every non-empty closed convex set W in H contains a unique point c that minimizes the norm; that is:

$$\|c\|_H = \inf\{\|w\|_H\colon\ w \in W\}.$$

Proof. Let $(v_n)_{n=1}^{\infty} \subset W$ be a sequence such that $\lim_{n\to\infty} \|v_n\|_H = d \stackrel{\text{def}}{=} \inf\{\|w\|_H\colon\ w \in W\}$. We claim that this sequence converges in H. Indeed, it is a Cauchy sequence (below we apply the Parallelogram identity and use the fact that $\frac{v_n+v_m}{2} \in W$, thus $\|\frac{v_n+v_m}{2}\|_H \geq d$):

$$\|v_n - v_m\|_H^2 = 2\|v_n\|_H^2 + 2\|v_m\|_H^2 - \|v_n + v_m\|_H^2 = 2\left(\|v_n\|_H^2 + \|v_m\|_H^2 - 2\left\|\frac{v_n+v_m}{2}\right\|_H^2\right)$$
$$\leq 2(\|v_n\|_H^2 + \|v_m\|_H^2 - 2d^2) \to 0 \quad (\min(n,m) \to \infty).$$

Denote by g the limit of $v_1, v_2, \ldots$ and note that $g \in W$ (as W is closed) and $\|g\|_H = \lim_{n\to\infty} \|v_n\|_H = d$. To show the uniqueness of g, take arbitrary $h \in W$ such that $\|h\|_H = d$. Then $\frac{g+h}{2} \in W$ as well and $d \leq \left\|\frac{g+h}{2}\right\|_H \leq \frac{\|g\|_H + \|h\|_H}{2} = d$. Consequently, $\|g + h\|_H = \|g\|_H + \|h\|_H$. So, we infer from Corollary 4.4 that $g = th$ or $h = tg$ for some $t \geq 0$. But $\|g\|_H = \|h\|_H$ and thus $h = g$. □

4.11 Corollary.
Let W be a non-empty closed convex set in a Hilbert space $(H, \|\cdot\|_H)$. For any vector $a \in H$ there exists a unique vector $b \in W$ that realizes the distance of a from W; that is:

$$\|a - b\|_H = \inf\{\|a - w\|_H\colon\ w \in W\}, \qquad b \in W.$$

The above vector b is called the *best approximation* of a (in W).

Proof of Corollary 4.11. It is enough to observe that b is the best approximation of a in W iff the vector $b - a$ has minimal norm among all the vectors from $W - a$, and to apply Theorem 4.10. □

4.12 Definition.
Two vectors a and b in an inner product space $(E, \langle\cdot, -\rangle_E)$ are *orthogonal* if $\langle a, b\rangle_E = 0$. Orthogonality of these vectors shall be denoted by $a \perp b$. More generally, for two subsets A and B of E we will write $A \perp B$ to express that $a \perp b$ for any $a \in A$ and $b \in B$. If A consists of a single vector x, we will also write $x \perp B$ in place of $A \perp B$.

For any set $A \subset E$, $A^{\perp}$ stands for the set of all vectors $u \in E$ such that $u \perp A$. In the case when E is a Hilbert space and F is a closed linear subspace, $F^{\perp}$ is called the *orthogonal complement* of F.

4.13 Proposition.
Let a, b and c be three vectors in a scalar product space $(E, \langle\cdot, -\rangle_E)$. Then:

(a) $a \perp a \iff a = 0$;

(b) $a \perp b \iff b \perp a$;

(c) *if* $a \perp b$ *and* $a \perp c$, *then* $a \perp xb + yc$ *for any scalars* x *and* y;

(d) [*Pythagorean equation*] *if* $a \perp b$, *then*

(4:4) $$\|a + b\|_E^2 = \|a\|_E^2 + \|b\|_E^2;$$

(e) *if* $\mathbb{K} = \mathbb{R}$ *and* (4:4) *holds, then* $a \perp b$;

(f) *if* $\mathbb{K} = \mathbb{C}$, *then* $a \perp b$ *iff* $\|a + z\|_E^2 = \|a\|_E^2 + \|z\|_E^2$ *for* $z \in \{b, ib\}$.

(proof—exercise)

4.14 Theorem. (Best approximation in a linear subspace)
Let E *ba a closed linear subspace of a Hilbert space* $(H, \langle \cdot, - \rangle_H)$. *For* $a \in H$ *and* $b \in E$ *the following conditions are equivalent:*

- *b is the best approximation of a in E;*
- *$a - b \perp E$.*

Proof. First assume $b \in E$ is such that $a - b \perp E$. Then, for any $w \in E$ (thanks to (4:4)),

$$\|a - w\|_H^2 = \|(a - b) + (b - w)\|_H^2 = \|a - b\|_H^2 + \|b - w\|_H^2 \geq \|a - b\|_H^2,$$

which shows that b is the best approximation of a in E.

Conversely, assume that $b \in E$ is such that $\eta \stackrel{\text{def}}{=} \langle a - b, c \rangle_E$ is non-zero for some unit vector $c \in E$. Then $f \stackrel{\text{def}}{=} b + \eta c$ also belongs to E and satisfies $a - f \perp c$ [exercise]. In particular, $a - f \perp \eta c$ $(= f - b)$ and therefore (again by (4:4)) $\|a - b\|_H^2 = \|(a - f) + (f - b)\|_H^2 = \|a - f\|_H^2 + \|f - b\|_H^2 > \|a - f\|_H^2$, which shows that b is not the best approximation of a in E. □

4.15 Definition.
Let $(H, \|\cdot\|_H)$ be a Hilbert space and E be its closed linear subspace. For any $x \in H$ denote by $P_E(x)$ the best approximation of x in E. Equivalently (cf. Theorem 4.14), $P_E(x)$ is a unique vector $v \in E$ such that $a - v \perp E$. $P_E \colon H \to E$ is called the *orthogonal projection* onto the subspace E.

4.16 Theorem.
For each closed linear subspace E *of a Hilbert space* $(H, \langle \cdot, - \rangle_H)$ *the orthogonal projection* P_E *is linear and bounded. Moreover,*

(a) $\mathcal{R}(P_E) = E$ *and* $P^2 = P$; *in particular,* H *is the direct sum of* E *and* $\mathcal{N}(P_E)$, *and both these subspaces are closed and linear;*

(b) $\mathcal{N}(P_E) = E^\perp$;

(c) $\|P_E\| = 1$ *iff* $E \neq \{0\}$;

(d) $\langle P_E x, y \rangle_H = \langle x, P_E y \rangle$ *for all* $x, y \in H$;

(e) $\langle P_E x, x \rangle_H \geq 0$ *for any* $x \in H$.

Proof. Fix two vectors $x, y \in H$ and, for clarity, set $u = P_E(x)$ and $v = P_E(y)$. Then both $x - u$ and $y - v$ are orthogonal to E. So, $\alpha(x - u) + (y - v) \perp E$ as well for each $\alpha \in \mathbb{K}$. Since $\alpha u + v \in E$, we get $P_E(\alpha x + y) = \alpha u + v$. In this way we have shown that P_E is linear. Moreover, since $x - u \perp u$, we infer from (4:4) that $\|x\|_H^2 = \|(x - u) + u\|_H^2 = \|x - P_E(x)\|^2 + \|P_E(x)\|_H^2 \geq \|P_E(x)\|_H^2$. Consequently, P_E is bounded and $\|P_E\| \leq 1$. Further, it is clear that $P_E(H) \subset E$ and $P_E(v) = v$ for any $v \in E$. These two properties imply that $P_E(H) = E$ and $P^2 = P$. In particular, the assertion of (c) and the last property of (a) easily follow. Moreover, $P_E(x) = 0$ iff $x \perp E$ (thanks to Theorem 4.14), which yields (b).

Finally, continuing notation introduced at the very beginning of this proof, we obtain $\{x - u, y - v\} \perp \{u, v\}$. So, $\langle u, v \rangle_H = \langle u, y \rangle_H$ as well as $\langle u, v \rangle_H = \langle x, v \rangle_H$. Combining these two equations we get (d), and substituting $y = x$ one obtains $\langle u, x \rangle_H = \|u\|_H^2$, which is followed by (e). □

4.17 ***Remark.***
Theorem 4.16 asserts, among other things, that each closed linear subspace E of a Hilbert space H admits a closed (linear) supplement; that is, for each such E there exists a closed linear subspace F of H such that $E \cap F = \{0\}$ and $E + F = H$. A deep result due to Lindenstrauss and Tzafriri states that such a property characterises Hilbert spaces among Banach ones (up to isomorphism). More precisely, if a Banach space has the property that each its closed linear subspace admits a closed supplement, then this space is isomorphic to a certain Hilbert space.

4.18 Theorem. (Orthogonal decomposition)
Let A be an arbitrary subset of a Hilbert space H. Then:

- *the set $F \stackrel{\text{def}}{=} A^\perp$ is a closed linear subspace of H;*
- *the set $E \stackrel{\text{def}}{=} (A^\perp)^\perp$ coincides with the closed linear span $\overline{\text{lin}} A$ of A;*
- *H is the direct sum of E and F;*
- *$E = \mathcal{N}(P_F) = \mathcal{R}(P_E)$ and $F = \mathcal{R}(P_F) = \mathcal{N}(P_E)$.*

In particular, for any closed linear subspace V of H, $V = (V^\perp)^\perp$.

Proof. It is easy to see that F is a closed linear subspace. In particular, so is E. Moreover, $A \subset E$ and, consequently, $\overline{\text{lin}} A \subset E$ as well. Denote by P the orthogonal projection onto $E_0 \stackrel{\text{def}}{=} \overline{\text{lin}} A$. Then $\mathcal{N}(P) = F$ [why?] and H is the direct sum of E_0 and F (see Theorem 4.14). By a similar argument, H is the direct sum of F and E. Since $E_0 \subset E$, we conclude that $E_0 = E$. In particular, the last statement of the theorem is proved. The remaining parts are left to the reader as (trivial) exercises. □

4.19 Theorem. (F. Riesz representation theorem for Hilbert spaces)
For any bounded linear functional ϕ on a Hilbert space $(H, \langle \cdot, - \rangle_H)$ there exists a unique vector $a \in H$ such that:

$$\phi(x) = \langle x, a \rangle_H \qquad (x \in H). \tag{4:5}$$

Moreover, $\|\phi\| = \|a\|_H$.

Proof. Denote by E the kernel of ϕ (which is a closed linear subspace of H). If $E = H$, we set $a = 0$ (in particular, $\|\phi\| = \|a\|$). Below we assume $E \neq \{0\}$. It follows from Theorem 4.18 that $E^\perp \neq \{0\}$ (since $(E^\perp)^\perp \neq H$). Let p be a unit vector orthogonal to E. Since $p \notin E$, we have $\phi(p) \neq 0$. Set $a \stackrel{\text{def}}{=} \overline{\phi(p)} p$ and observe that $\|a\|_H = |\phi(b)|$ and $\phi(a) = |\phi(p)|^2$. Hence,

$$a \perp E, \qquad \phi(a) = \|a\|_H^2 > 0. \tag{4:6}$$

Using (4:6), we will show that

$$\phi(x) = \langle x, a \rangle_H \qquad (x \in H). \tag{4:7}$$

To this end, fix $x \in H$ and set $y \stackrel{\text{def}}{=} x - \frac{\phi(x)}{\|a\|_H^2} a$. A direct calculation shows that $\phi(y) = 0$. So, $y \in E$ and therefore $a \perp y$. This yields $0 = \langle y, a \rangle_H = \langle x, a \rangle_H - \phi(x)$, which proves (4:7). To convince onself that $\|\phi\| = \|a\|_H$, first one applies the Schwarz inequality (to get $\|\phi\| \leq \|a\|_H$) and finally one substitutes $x = \frac{a}{\|a\|_H}$.

To see the uniqueness of a, assume $b \in H$ is such that $\langle x, b \rangle_H = \langle x, a \rangle_H$ for any $x \in H$. Equivalently, $a - b \perp H$. In particular, $a - b \perp a - b$ and hence $a - b = 0$. □

Basic consequences of the above result follow.

4.20 Corollary.
Let $(H, \langle \cdot, - \rangle_H)$ and $(K, \langle \cdot, - \rangle_K)$ be two Hilbert spaces.

(A) *For any* $T \in \mathscr{L}(H,K)$ *the formula* $\Phi(x,y) \overset{\text{def}}{=} \langle Tx, y\rangle_K$ *correctly defines a sesquilinear form* $\Phi\colon H \times K \to \mathbb{K}$ *such that*

$$\sup\{|\Phi(x,y)|\colon\ x \in \bar{B}_H,\ y \in \bar{B}_K\} < \infty. \tag{4:8}$$

Actually, the above quantity equals $\|T\|$.

(B) *Conversely, if* $\Phi\colon H \times K \to \mathbb{K}$ *is a sesquilinear form for which* (4:8) *holds, then there exists a unique* $T \in \mathscr{L}(H,K)$ *such that* $\Phi(x,y) = \langle Tx, y\rangle_K$ *for all* $x \in H$ *and* $y \in K$.

Proof. Part (A) is left as an exercise. Here we focus only on (B). For any $x \in H$ the formula $\phi_x(y) \overset{\text{def}}{=} \overline{\Phi(x,y)}$ correctly defines a bounded linear functional on K. So, we infer from Theorem 4.19 that there is a unique vector from K, to be denoted by Tx, such that $\phi_x(y) = \langle y, Tx\rangle_K$. In other words, $\Phi(x,y) = \overline{\langle y, Tx\rangle}_K = \langle Tx, y\rangle_K$. It readily follows from the uniqueness part of the Riesz' theorem that $T\colon H \to K$ is linear. Its continuity is left to the reader. □

4.21 Corollary.
For any bounded linear operator $T\colon H \to K$ *between Hilbert spaces* $(H, \langle\cdot,-\rangle_H)$ *and* $(K, \langle\cdot,-\rangle_K)$ *there exists a unique bounded linear operator* $T^*\colon K \to H$ *such that*

$$\langle Tx, y\rangle_K = \langle x, T^*y\rangle_H \qquad (x \in H,\ y \in K). \tag{4:9}$$

Moreover, $\|T\| = \|T^*\|$.

Proof. It is sufficient to apply the previous result to $\Phi\colon K \times H \ni (y,x) \mapsto \langle y, Tx\rangle_K \in \mathbb{K}$. The details are left to the reader. □

4.22 Definition.
The operator T^*, defined in Corollary 4.21 for $T \in \mathscr{L}(H,K)$, is called the *adjoint* operator (or simply the *adjoint*) of T.

An operator $T \in \mathscr{L}(H)$ is said to be:

- *selfadjoint* if $T = T^*$;
- *unitary* if $T^*T = TT^* = I_H$ where I_H is the identity operator on H;
- *normal* if $TT^* = T^*T$.

4.23 Proposition.
Let $(H, \langle\cdot,-\rangle_H)$, $(K, \langle\cdot,-\rangle_K)$ *and* $(W, \langle\cdot,-\rangle_W)$ *be Hilbert spaces.*

(A) *The function* $\mathscr{L}(H,K) \ni T \mapsto T^* \in \mathscr{L}(K,H)$ *is an antilinear isometric* involution *(that is,* $(T^*)^* = T$*).*

(B) $(TS)^* = S^*T^*$ *for any* $T \in \mathscr{L}(H,K)$ *and* $S \in \mathscr{L}(W,H)$*; and* $I_H^* = I_H$.

(C) *A bounded linear operator* $A\colon H \to H$ *is selfadjoint iff* $\langle Ax, y\rangle_H = \langle x, Ay\rangle_H$ *for all* $x, y \in H$. *If* $\mathbb{K} = \mathbb{C}$, $A = A^*$ *iff* $\langle Ax, x\rangle_H \in \mathbb{R}$ *for all* $x \in H$.

(D) *A bounded linear operator* $U\colon H \to K$ *is unitary iff it is a bijective isometry, iff* U *is surjective and* $\langle Ux, Uy\rangle_K = \langle x, y\rangle_H$ *for all* $x, y \in H$. *If* U *is a unitary operator, so is* U^* *and* $U^* = U^{-1}$.

(E) *An operator* $V \in \mathscr{L}(H,K)$ *is an isometry iff* $V^*V = I_H$.

(F) *The orthogonal projection onto any closed linear subspace of* H *is selfadjoint.*

(proof—exercise)

4.24 ***Remark.***
In the realm of complex Hilbert spaces, each bounded linear operator acting on a single Hilbert spaces (that is, not between two different spaces) is a linear combination of two selfadjoint operators. Indeed, if $T \in \mathscr{L}(H)$, then $\mathrm{Re}(T) \overset{\text{def}}{=} \frac{T+T^*}{2}$ and $\mathrm{Im}(T) \overset{\text{def}}{=} \frac{T-T^*}{2i}$ are selfadjoint and $T = \mathrm{Re}(T) + i\,\mathrm{Im}(T)$. Since all the bounded selfadjoint operators on a given Hilbert space always form a vector space over $\mathbb{R}$, such a phenomenon does not occur in the realm of real Hilbert spaces.

4.25 Definition.
An *orthogonal system* in a Hilbert space $(H, \langle \cdot, - \rangle_H)$ is any collection of pairwise orthogonal vectors. More precisely, a family $\{u_s\}_{s\in S} \subset H$ is orthogonal iff $u_s \perp u_t$ for all <u>distinct</u> indices $s, t \in S$. If, in addition, $\|u_s\|_H = 1$ for any $s \in S$, the system $\{u_s\}_{s\in S}$ is called *orthonormal.*

An *orthonormal* (resp. *orthogonal*) *basis* of H is a maximal orthonormal system (resp. maximal orthogonal system consisting of non-zero vectors).

4.26 Example.

(A) Consider the Hilbert space $\ell_2(X)$ (where X is an arbitrary non-empty set; e.g. $X = \mathbb{N}_1$ or $X = \{1, \dots, n\}$) with its standard inner product:

$$\langle u, v\rangle_{\ell_2(X)} = \sum_{x\in X} u(x)\overline{v(x)}.$$

The following functions form an orthonormal basis of $\ell_2(X)$, called *canonical*:

$$f_s(x) = \begin{cases} 1 & x = s \\ 0 & x \neq s \end{cases} \qquad (s \in X).$$

(B) Consider $L^2([-\pi, \pi])$ with the following inner product:

$$\langle f, g\rangle_{L^2([-\pi,\pi])} \overset{\text{def}}{=} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\overline{g(t)}\,\mathrm{d}t.$$

All the trigonometric functions:

$$\begin{array}{llll} f_0(x) = 1, & f_1(x) = \cos(x), & \dots & f_n(x) = \cos(nx), \quad \dots \\ & g_1(x) = \sin(x), & \dots & g_n(x) = \sin(nx), \quad \dots \end{array}$$

form an orthogonal system in $L^2([-\pi, \pi])$. Actually, this is an orthogonal basis of that space (but this property is less trivial).

(C) Now consider $L^2([0, \pi])$ with the following inner product:

$$\langle f, g\rangle_{L^2([0,\pi])} \overset{\text{def}}{=} \frac{1}{\pi}\int_0^{\pi} f(t)\overline{g(t)}\,\mathrm{d}t.$$

The functions $f_0, f_1, \dots$ (restricted to $[0, \pi]$) defined in the previous example form an orthogonal basis of $L^2([0, \pi])$.

(D) Equip the circle group $\mathbb{T} \overset{\text{def}}{=} \{z \in \mathbb{C}\colon\ |z| = 1\}$ with its probabilistic *Haar* measure λ; that is,

$$\lambda(A) \overset{\text{def}}{=} \frac{1}{2\pi}|\{t \in [0, 2\pi)\colon\ e^{it} \in A\}| \qquad (A \subset \mathbb{T} \text{ Borel})$$

where $|B|$ (for a Borel set $B \subset \mathbb{R}$) stands for the (one-dimensional) Lebesgue measure of B. One proves that

$$\chi_k(z) = z^k \qquad (k \in \mathbb{Z},\ z \in \mathbb{T})$$

are all continuous homomorphisms from $\mathbb{T}$ into $\mathbb{T}$ and they form an orthonormal basis of $L^2(\lambda)$. (More generally, for an arbitrary compact Abelian group G, all the continuous homomorphisms from G into $\mathbb{T}$ form an orthonormal basis of $L^2(\mu)$ where μ is the probabilistic Haar measure of G.)

4.27 Proposition.
Let $(H, \|\cdot\|_H)$ be a Hilbert space.

(A) *For any orthogonal system (resp. basis) $\{u_s\}_{s\in S} \subset H$ consisting of non-zero vectors, the vectors $v_s \overset{\text{def}}{=} \frac{u_s}{\|u_s\|_H}$ form an orthonormal system (resp. basis).*

(B) *An orthogonal system consisting of non-zero vectors is linearly independent.*

(C) *Let $\mathscr{B} = \{u_s\}_{s\in S}$ be an orthonormal system (resp. orthogonal system consisting of non-zero vectors) in H. Then $\mathscr{B}$ is an orthonormal (resp. orthogonal) basis iff the following condition is fulfilled:*

(4:10) $$(v \in H,\ \forall s \in S\colon\ v \perp u_s) \implies v = 0.$$

(proof—exercise)

4.28 Theorem.
Let $\{u_s\}_{s\in S}$ be an orthonormal system in a Hilbert space $(H, \langle\cdot,-\rangle_H)$. Then, for any vector $g \in H$:

(OS1) *The series $\sum_{s\in S}\langle g, u_s\rangle_H u_s$ converges unconditionally in H (in the norm topology) to a certain vector $h \in H$.*

(OS2) *$g - h \perp u_s$ for any $s \in S$.*

(OS3) *$\|h\|_H^2 = \sum_{s\in S} |\langle g, u_s\rangle_H|^2$ and $\|g\|_H^2 = \|g-h\|_H^2 + \|h\|_H^2$.*

(OS4) ***Bessel's inequality**:*
$$\sum_{s\in S} |\langle g, u_s\rangle_H|^2 \le \|g\|_H^2.$$
In particular, there are only countably many indices $s \in S$ for which $\langle g, u_s\rangle_H \ne 0$.

Proof. Fix a finite set $F \subset S$, say $F = \{s_1, \dots, s_n\}$ (where n is the size of F). Set $h_F \overset{\text{def}}{=} \sum_{j=1}^n \langle g, u_{s_j}\rangle_H u_{s_j}$. A direct calculation shows that $g - h_F \perp \{u_s\colon\ s \in F\}$. So, it follows from (4:4) that

(4:11) $$\|h_F\|_H^2 = \sum_{s\in F} |\langle g, u_s\rangle_H|^2 \qquad \text{and} \qquad \|g\|_H^2 = \|g-h_F\|_H^2 + \sum_{s\in F} |\langle g, u_s\rangle_H|^2.$$

In particular,

(4:12) $$\sum_{s\in F} |\langle g, u_s\rangle_H|^2 \le \|g\|_H^2 < \infty.$$

Consequently, the set $J \overset{\text{def}}{=} \{s \in S\colon\ \langle g, u_s\rangle_H \ne 0\}$ is at most countable (why?). Denoting by $N \in \{0, 1, 2, \dots, \infty\}$ its size, we may write $J = \{t_n\}_{n=1}^N$ (where $t_n \in S$ are all different; we assume here that $\{t_n\}_{n=1}^0$ is the empty set). Note that $\sum_{s\in S}\langle g, u_s\rangle_H u_s = \sum_{n=1}^N \langle g, u_{t_n}\rangle_H u_{t_n}$ (provided the left-hand side expression of this equation is well defined) and $\sum_{s\in S} |\langle g, u_s\rangle_H|^2 = \sum_{n=1}^N |\langle g, u_{t_n}\rangle_H|^2$. Observe that if J is finite, all the claims of the theorem easily follow. Below we assume that $N = \infty$.

Substituting $F = \{t_1, \dots, t_n\}$ in (4:12) and letting $n \to \infty$, we get $\sum_{n=1}^\infty |\langle g, u_{t_n}\rangle_H|^2 \le \|g\|_H^2$, which is equivalent to (OS4). In particular, for any $\varepsilon > 0$ there is finite m such that $\sum_{n=m+1}^\infty |\langle g, u_{t_n}\rangle_H|^2 \le \varepsilon^2$, which implies that for any finite set $F \subset S$ disjoint from $\{t_1, \dots, t_m\}$ one has $\|h_F\|_H \le \varepsilon$ (cf. (4:11)). This implies that the series specified in (OS1) is unconditionally convergent to $h \overset{\text{def}}{=} \sum_{n=1}^\infty \langle g, u_{t_n}\rangle_H u_{t_n}$. For simplicity, set $h_n \overset{\text{def}}{=} \sum_{k=1}^n \langle g, u_{t_k}\rangle_H u_{t_k}$. Since both g and h are orthogonal to u_s for $s \in S \setminus J$, and $g - h_n \perp u_{t_k}$ for $k \le n$, we get (OS2). Finally, (OS3) follows from (4:11) applied to $F = \{t_1, \dots, t_n\}$ by letting $n \to \infty$. □

As a consequence of the above result, we obtain the following important

4.29 Theorem.
Let $\mathscr{B} = \{e_s\}_{s\in S}$ be an orthonoemal basis in a Hilbert space $(H, \langle\cdot,-\rangle_H)$. Then, for any vector $v \in H$:

(OB1) $v = \sum_{s\in S}\langle v, e_s\rangle_H e_s$ *and this series converges unconditionally in the norm topology.*

(OB2) ***Parseval's identity:***

$$\|f\|_H^2 = \sum_{s\in S} |\langle v, e_s\rangle_H|^2.$$

(OB3) $\overline{\operatorname{lin}}\mathscr{B} = H$.

Proof. It follows from (OS1) that the series specified in (OB1) is unconditionally convergent to some vector h such that $g - h \perp \mathscr{B}$. We infer from (4:10) that $h = g$, which implies (OB1) and (OB3). Finally, (OB2) is implied by the first part of (OS3). □

4.30 Definition.
For any metrisable space X let $\operatorname{dens}(X)$ denote the least cardinal number among sizes of dense subsets of X. (In particular, $\operatorname{dens}(X) \le \aleph_0$ iff X is separable.)

4.31 Proposition.
Let H be a Hilbert space.

(A) *There exists an orthonormal basis in H.*

(B) *All the orthonormal bases of H have the same cardinality. More precisely, if $\mathscr{B}$ is an orthonormal basis of H, then:*

$$\operatorname{card}(\mathscr{B}) = \begin{cases} \dim(H) & \text{if } H \text{ is finite-dimensional} \\ \operatorname{dens}(H) & \text{if } H \text{ is infinite-dimensional} \end{cases}.$$

Proof. Part (A) follows from Zorn's lemma. To show (B), we consider two cases. If H is finite-dimensional, then any of its orthonormal bases has to be a 'Hamel' basis, thanks to part (B) of Proposition 4.27 and (OB3). So, its size equals $\dim(H)$. On the other hand, if H is infinite-dimensional, its orthonormal basis $\mathscr{B}$ has to be infinite (again by (OB3)). Moreover, since $\|e - f\|_H = \sqrt{2}$ for any distinct $e, f \in \mathscr{B}$, we conclude that $\operatorname{dens}(\mathscr{B}) = \operatorname{card}(\mathscr{B})$. But $\operatorname{dens}(\mathscr{B}) \le \operatorname{dens}(H)$ (inequality valid in metrisable spaces), which yields $\operatorname{card}(\mathscr{B}) \le \operatorname{dens}(H)$. To see the reverse inequality, consider the set Q of all finite linear combinations of vectors from $\mathscr{B}$ with scalars from a countable dense subset of $\mathbb{K}$ (e.g., from $\mathbb{Q}$ if $\mathbb{K} = \mathbb{R}$ or from $\mathbb{Q} + i\mathbb{Q}$ if $\mathbb{K} = \mathbb{C}$). Since $\mathscr{B}$ is infinite, it follows that $\operatorname{card}(Q) = \operatorname{card}(\mathscr{B})$. However, Q is dense in H (thanks to (OB3)) and therefore $\operatorname{dens}(H) \le \operatorname{card}(Q)$, which finishes the proof. □

4.32 Definition.
For any Hilbert space H, the cardinality of any of its orthonormal basis is called *Hilbert space dimension* or briefly *dimension* of H and denoted by $\dim(H)$.

4.33 Example.
It follows from properties exhibited in Example 4.26 that:

- $\dim(L^2([0,1])) = \aleph_0$;
- $\dim(\ell_2(X)) = \operatorname{card}(X)$.

In particular, for any cardinal number α there exists a Hilbert space of dimension α.

4.34 Theorem. (Uniqueness of Hilbert spaces)
Let H be a non-trivial Hilbert space of dimension α and X be a set of cardinality α. Then there exists a unitary (linear) operator $U\colon H \to \ell_2(X)$. In particular:

- *Any two Hilbert spaces (over the same field) of the same dimension are linearly isometric.*
- *All infinite-dimenional separable Hilbert spaces over the same field are linearly isometric.*

Proof. Let $\{f_x\}_{x\in X}$ stand for the canonical basis of $\ell_2(X)$, defined in Example 4.26, and denote by $c_{00}(X)$ the subspace of $\ell_2(X)$ consisting of all the functions $f \in \ell_2(X)$ for which the set $\{x \in X\colon\ f(x) \neq 0\}$ is finite. Take any orthonormal basis $\{e_x\}_{x\in X}$ of H and consider a well defined linear operator $U_0\colon c_{00}(X) \to H$ given by $U_0(f) = \sum_{x\in X} f(x)e_x$. It follows from orthogonality of all the vectors e_x that $\|U_0(f)\|_H = \|f\|_{\ell_2(X)}\|$. In particular, U_0 is isometric and extends uniquely to a linear isometric operator $U\colon \ell_2(X) \to H$ (note that $c_{00}(X)$ is dense in $\ell_2(X)$). Then $\mathcal{R}(U)$ is linear and closed (why?), which implies (thanks to (OB3)) that U is surjective. So, it is a unitary operator, and the proof is finished. □

4.35 Corollary.
Let $\mathscr{B} = \{e_n\}_{n=1}^{\infty}$ be an orthonormal system in a Hilbert space $(H, \langle\cdot, -\rangle_H)$. Then:

(a) *$\mathscr{B}$ is an orthonormal basis of H iff $\overline{\mathrm{lin}}\mathscr{B} = H$, iff*

(4:13)
$$\|h\|_H^2 \le \sum_{n=1}^{\infty} |\langle h, e_n\rangle_H|^2$$

for all $h \in H$.

(b) *For any sequence $(a_n)_{n=1}^{\infty} \subset \mathbb{K}$ the series $\sum_{n=1}^{\infty} a_n e_n$ is convergent in H (in the norm topology) iff $(a_n)_{n=1}^{\infty} \in \ell_2$.*

(c) *For any $(a_n)_{n=1}^{\infty} \in \ell_2$ and $h \overset{\text{def}}{=} \sum_{n=1}^{\infty} a_n e_n$ $(\in H)$, one has $a_n = \langle h, e_n\rangle_H$ for all n.*

In particular, if $\mathscr{B}$ is an orthonormal basis of H, then for any $h \in H$ there exists a unique sequence $(a_n)_{n=1}^{\infty} \subset \mathbb{K}$ such that the series $\sum_{n=1}^{\infty} a_n e_n$ converges in the norm topology to h.

Proof. If $\mathscr{B}$ is an orthonormal basis of H, then (OB2)–(OB3) hold. Conversely, if (OB3) holds and $u \in H$ is orthogonal to each e_n, then $u \perp \overline{\mathrm{lin}}\mathscr{B} = H$ and therefore $u = 0$, which shows that then $\mathscr{B}$ is an orthonormal basis of H. Finally, if (4:13) is satisfied for any $h \in H$, then $h = 0$ is the only vector from H that is orthogonal to $\mathscr{B}$ (so, again, $\mathscr{B}$ is the orthonormal basis of H.

Now take any sequence $(a_n)_{n=1}^{\infty}$ of scalars. If the series specified in (b) converges, then the sequence of all $\|\sum_{k=1}^{n} a_k e_k\|_H^2$ $(= \sum_{k=1}^{n} |a_k|^2)$ is bounded and therefore $(a_n)_{n=1}^{\infty} \in \ell_2$. Conversely, if $(a_n)_{n=1}^{\infty} \in \ell_2$, then the partial sums of the series under the consideration form a Cauchy sequence (exercise) and hence the series converges.

The remaining parts are left to the reader. □

4.36 Theorem.
Let E be a closed linear subspace of a Hilbert space $(H, \langle\cdot, -\rangle_H)$ and let $\{e_s\}_{s\in S}$ be an orthonormal basis of E. Then:

$$P_E(x) = \sum_{s\in S} \langle x, e_s\rangle_H e_s \qquad (x \in H).$$

Proof. We know from (OS1)–(OS2) that the series $\langle x, e_s\rangle_H e_s$ converges unconditionally (in the norm topology) to some $h \in H$ such that $x - h \perp \mathscr{B} \overset{\text{def}}{=} \{e_s\colon\ s \in S\}$. In particular, $\mathscr{B} \subset (x-h)^{\perp}$ and, consequently (thanks to (OB3) and Theorem 4.18), $E = \overline{\mathrm{lin}}\mathscr{B} \subset (x-h)^{\perp}$. We also have $h \in E$ and thus $P_E(x) = h$, by Theorem 4.14. □

4.37 Example.
The Parseval's identity may be used to find explicit values of certain series. Below we give an illustrative example of this method.

Consider $H = L^2([0,\pi])$ with the inner product $\langle \cdot, - \rangle_H$ and orthogonal basis $\{f_n\}_{n=0}^{\infty}$ defined in part (C) of Example 4.26. Setting $e_n \stackrel{\text{def}}{=} \frac{f_n}{\|f_n\|_H}$, we obtain an orthonormal basis of H. It is easy to check that:

$$e_n(x) = \begin{cases} 1 & n = 0 \\ \sqrt{2}\cos(nx) & n > 0 \end{cases} \qquad (x \in [0,\pi]).$$

Now let $g(x) = x$. A direct calculations show that:

- $\|g\|_H^2 = \frac{\pi^2}{3}$;
- $\langle g, e_0 \rangle_H = \frac{\pi}{2}$;
- $\langle g, e_n \rangle_H = \frac{(-1)^n - 1}{\pi n^2}\sqrt{2}$ for $n > 0$.

It follows from the Parseval's identity that $\|g\|_H^2 = |\langle g, e_0 \rangle_H|^2 + \sum_{n=1}^{\infty} |\langle g, e_n \rangle_H|^2$, which yields

$$\sum_{k=0}^{\infty} \frac{8}{(2k+1)^4} = \frac{\pi^4}{12}. \tag{4:14}$$

Further, observe that $\sum_{n=1}^{\infty} \frac{1}{n^4} = \sum_{k=1}^{\infty} \frac{1}{(2k)^4} + \sum_{k=0}^{\infty} \frac{1}{(2k+1)^4}$, which implies that $\sum_{n=1}^{\infty} = \frac{16}{15} \sum_{k=0}^{\infty} \frac{1}{(2k+1)^4}$. The last formula, combined with (4:14), finally gives:

$$\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90},$$

which is a well-known formula due to Euler.

Digression: Euler found the explicit formulas for all series of the form $\zeta(k) \stackrel{\text{def}}{=} \sum_{n=1}^{\infty} \frac{1}{n^k}$ where k is a posivite even integer. In particular, $\frac{\zeta(2p)}{\pi^{2p}}$ is always rational (for $p \in \mathbb{N}_1$) and its integer denominator is precisely described (in terms of p). Explicit value of $\zeta(n)$ is known for no odd $n > 1$. (However, Apéry proved that $\zeta(3)$ is irrational.)

4.38 Example. (Gram-Schmidt process)
Let $g_1, g_2, \ldots$ be a finite or infinite sequence of vectors in a scalar product space $(E, \langle \cdot, - \rangle_E)$. The algorithm described below is known as the *Gram-Schmidt process* (or *Gram-Schmidt algorithm*) and it allows us to achieve two effects:

- to determine if this sequence consists of linearly independent vectors; and
- if they are linearly independent, to construct an orthonormal system $e_1, e_2, \ldots$ such that $\operatorname{lin}\{e_1, \ldots, e_k\} = \operatorname{lin}\{g_1, \ldots, g_k\}$ for all possible k.

The algorithm goes as follows:

(Step 0) start from $n = 1$ and define $f_1 = g_1$;

(Step 1) whenever f_k has been defined, compute $\alpha_k \stackrel{\text{def}}{=} \langle f_k, f_k \rangle_E$;

(Step 2) if $\alpha_k \neq 0$, define $e_n = \frac{f_k}{\sqrt{\alpha_k}}$;

(Step 3) if $\alpha_k = 0$ for some k, then the vectors $g_1, \ldots, g_k$ are linearly dependent and the algorithm finishes;

(Step 4) if all $\alpha_1, \ldots, \alpha_k$ are non-zero, pass to $n = k+1$ and set $f_n = g_n - \sum_{j=1}^{n-1} \frac{\langle g_n, f_j \rangle_E}{\alpha_j} f_j$, and return to Step 1.

It is left as an exercise to verify that this procedure has all the properties listed above.

The above scheme enables us, e.g., to construct an orthonormal basis of $L^2([a,b])$ consisting of polynomials such that the nth polynomial (with $n \geq 0$) has degree equal to n. Such sequences are known as *orthogonal polynomials*.

4.39 ***Remark.***
The property of orthonormal bases in separable Hilbert spaces formulated in the last sentence of Corollary 4.35 may be seen as foundations for a more general notion, namely of Schauder bases in Banach spaces: a sequence $(e_n)_{n=1}^{\infty}$ of vectors in a separable Banach space E is said to be a *Schauder basis* of E if for any vector $x \in E$ there is a unique sequence $(a_n)_{n=1}^{\infty}$ of scalars such that the series $\sum_{n=1}^{\infty} a_n e_n$ converges in the norm topology to x. All classical separable Banach spaces have Schauder bases. However, there are known examples of separable Banach spaces that admit no such basis. More on this notion the reader can find in any classical book on Banach spaces.

4.40 ***Remark.***
A deep result due to Toruńczyk (from the 80's of the 20th century) asserts that each infinite-dimensional Banach space E is homeomorphic to $\ell_2(X)$ where X is a set of cardinality $\operatorname{dens}(E)$. It was a solution of a Banach's question from the Scottish Book. A similar result is false if one searches for homeomorphisms that are uniformly continuous in both directions.

5 Central tools of functional analysis

In this chapter we will present results which have made functional analysis an important branch of mathematics. Almost all of them are due to Banach and his collaborators.

5.1 Theorem. (Uniform Boundedness Principle)
Let X be a Banach space and for each $s \in S$ (where S is a non-empty set) let $T_s\colon X \to Y_s$ be a bounded linear operator from X into a normed vector space Y_s. If

$$\sup_{s\in S} \|T_s(x)\|_{Y_s} < \infty \tag{5:1}$$

for all $x \in X$, then $\sup_{s\in S} \|T_s\| < \infty$.

Proof. For $n > 0$ let F_n consist of all $x \in X$ such that $\|T_s(x)\|_{Y_s} \leq n$ for each $s \in S$. It follows from (5:1) that $X = \bigcup_{n=1}^{\infty} F_n$, and since $F_n = \bigcap_{s\in S} T_s^{-1}(n\bar{B}_{Y_s})$, we infer that all these sets are closed. So, the Baire category theorem implies that F_N has non-empty interior for some N. Fix $a \in X$ and $r > 0$ so that $a + r\bar{B}_X \subset F_N$. Then, for arbitrary $x \in \bar{B}_X$ and $s \in S$ we have $a, a+rx \in F_N$, thus $\|T_s(a)\|_{Y_s} \leq N$ and, similarly, $\|T_s(a+rx)\|_{Y_s} \leq N$. We conclude that $\|T_s(x)\|_{Y_s} \leq \frac{1}{r}(\|T_s(a+rx)\|_{Y_s} + \|T_s(a)\|_{Y_s}) \leq \frac{2N}{r}$, and therefore $\|T_s\| \leq \frac{2N}{r}$ for any $s \in S$. □

5.2 Corollary. (Banach-Steinhaus Theorem)
If $T_1, T_2, T_3, \ldots$ are bounded linear operators from a Banach space X into a normed vector space Y such that $T_n(x) \xrightarrow{Y} L(x)$ $(n \to \infty)$ for any $x \in X$ and some $L\colon X \to Y$, then L is a bounded linear operator as well.

Proof. It is easily seen that L is a linear operator. The pointwise convergence implies that (5:1) is satisfied for all $x \in X$. So, $M \stackrel{\text{def}}{=} \sup_{n>0} \|T_n\| < \infty$, by Theorem 5.1. Consequently, $\|L(x)\|_Y = \lim_{n\to\infty} \|T_n(x)\|_Y \leq M\|x\|_X$, which finishes the proof. □

5.3 Theorem. ([Banach] Isomorphism Theorem)
Let $T\colon X \to Y$ be a bounded bijective linear operator between two Banach spaces. Then the inverse T^{-1} of T is bounded as well.

The above result shall be generalised (to the context of complete metric TVS's) in the next chapter. As we will see, the proof of a general case is more subtle than the one presented below (although quite similar).

Proof of Theorem 5.3. Since $Y = \bigcup_{n=1}^{\infty} T(n\bar{B}_X)$, it follows from the Baire category theorem that $b + c\bar{B}_Y \subset F \stackrel{\text{def}}{=} \overline{T(N\bar{B}_X)}$ for some $b \in Y$, $c > 0$ and $N \in \mathbb{N}_1$. As $T(N\bar{B}_X)$ is a symmetric convex set, so is its closure F. Hence

$-b + c\bar{B}_Y \subset F$. So, for any $y \in \bar{B}_Y$ we have that $b + cy, -b + cy \in F$ and therefore also $cy = \frac{1}{2}(b + cy) + \frac{1}{2}(-b + cy)$ belongs to F. We conclude that $\bar{B}_Y \subset \frac{1}{c}F = \overline{T(\frac{N}{c}\bar{B}_X)}$. In this way have shown that $\bar{B}_Y \subset \overline{T(r\bar{B}_X)}$ for some $r > 0$. In particular,

$$\forall y \in \bar{B}_Y \ \exists x \in r\bar{B}_X \colon \ \|y - T(x)\|_Y \leq \frac{1}{2}. \tag{5:2}$$

Fix $v \in \bar{B}_Y$. We will now construct inductively a sequence $(x_n)_{n=1}^{\infty} \subset r\bar{B}_X$ such that

$$\Big\| v - \sum_{k=1}^{n} \frac{1}{2^{k-1}} T(x_k) \Big\|_Y \leq 2^{-n} \tag{5:3}$$

for all $n > 0$. The case $n = 1$ immediately follows from (5:2) applied to $y = v$. Now assume that for some $m > 0$ the vectors $x_1, \ldots, x_m$ has already been defined so that (5:3) holds for $n = 1, \ldots, m$. Then $y \stackrel{\text{def}}{=} 2^m(v - \sum_{k=1}^{m} \frac{1}{2^{k-1}} T(x_k))$ belongs to $\bar{B}_Y$ and it follows from (5:2) that there is a vector $x_{m+1} \in r\bar{B}_X$ such that $\|2^m(v - \sum_{k=1}^{m} \frac{1}{2^{k-1}} T(x_k)) - T(x_{m+1})\|_Y \leq \frac{1}{2}$. Dividing both sides of these inequality by 2^m, we obtain (5:3) for $n = m + 1$, as we wished.

Now since $\|x_k\|_X \leq r$ for each $k > 0$, we infer that the series $\sum_{n+1}^{\infty} \frac{x_n}{2^{n-1}}$ converges in X, say to u. Then $\|u\|_X \leq 2r$ and $T(u) = v$, by (5:3). Consequently, $T^{-1}(\bar{B}_Y) \subset 2r\bar{B}_X$ and hence $\|T^{-1}\| \leq 2r$. □

As immediate consequences of the above result, we obtain the next two results.

5.4 Theorem. ([Banach] Open Mapping Theorem)
If a bounded linear operator $T \colon X \to Y$ between Banach spaces is surjective, then it is an open map.

Proof. It follows from part (F) of Proposition 2.6 (p. 3) that there exists a continuous linear operator $\tilde{T} \colon X/\mathcal{N}(T) \to Y$ such that $T = \tilde{T} \circ \pi_T$ where $\pi_T \colon X \to X/\mathcal{N}(T)$ is the quotient map. Item (D) therein shows that π_T is an open map. Further, we conclude from Theorem 2.8 (p. 3) that $X/\mathcal{N}(T)$ is a Banach space. Finally, since T is surjective, $\tilde{T}$ is a bijection. Thus, Theorem 5.3 yields that $\tilde{T}$ is a homeomorphism. Consequently, T is open as the composition of two such maps. □

5.5 Theorem. ([Banach] Closed Graph Theorem)
Let $T \colon X \to Y$ be a linear operator between Banach spaces X and Y. Then T is bounded iff the following condition is satisfied:

$$(x_n \in X, \ \lim_{n\to\infty} x_n = 0, \ \lim_{n\to\infty} T(x_n) = y \in Y) \implies y = 0. \tag{$*$}$$

Proof. We only need to show the sufficiency of $(*)$. To this end, consider the graph $\Gamma \stackrel{\text{def}}{=} \{(x, T(x)) \colon \ x \in X\}$ of T and a linear bijection $L \colon \Gamma \ni (x, y) \mapsto x \in X$. The condition $(*)$ (combined with linearity of T) implies that Γ is a closed subspace of $X \times Y$, and thus it is a Banach space. Since L is continuous, we conclude from Theorem 5.3 that L^{-1} is a bounded operator, and the assertion follows. □

The following result is a prime example illustrating the power of the Closed Graph Theorem.

5.6 Corollary.
If a Banach space X is the direct sum of its two closed linear subspaces E and F, then the projection $P_E \colon X \to E$ onto E along F is bounded.

Proof. Assume $x_1, x_2, \ldots \in X$ converge to 0 and $\lim_{n\to\infty} P_E(x_n) = e \in E$. Then the vectors $x_n - P_E(x_n)$ belong to F and tend to $-e$. It follows from the closedness of F that $-e \in E \cap F = \{0\}$ and therefore $e = 0$. The Closed Graph Theorem finishes the proof. □

5.7 Definition.
A function $p \colon X \to \mathbb{R}$ (defined on a real vector space X) is said to be *sublinear* if:

- $f(tx) = tf(x)$ for all $x \in X$ and $t > 0$;
- $f(x+y) \leq f(x) + f(y)$ for any $x, y \in X$.

It is worth underlying that sublinear functionals can take negative values.

5.8 Example.
The following functions are classical sublinear functionals take are not semi-norms:

- $p(f) = \sup f(X), f \in \ell_\infty(X)$ (where X is an arbitrary non-empty set);
- $q(f) = \limsup_{x\to\infty} f(x), f \in \ell_\infty(Z)$ where Z is one the sets: $\mathbb{N}$, $\mathbb{R}_+$, $\mathbb{Z}$, $\mathbb{R}$.

It is also readily seen that linear functionals are sublinear and that the pointwise supremum of a non-empty collection of sublinear functionals (defined on a common vector space) is sublinear as well provided it is real-valued. The next theorem (due to Hahn and Banach) implies that all sublinear functionals can be expressed as pointwise suprema of certain non-empty families of linear functionals (defined on the underlying vector space).

5.9 Theorem. (Hahn-Banach Theorem)
Let $p\colon X \to \mathbb{R}$ be a sublinear functional on a real vector space X and X_o be a linear subspace of X. Then every linear functional $\phi_o\colon X_o \to \mathbb{R}$ such that

$$\phi_o(x) \leq p(x) \qquad (x \in X_o)$$

extends to a linear functional $\phi\colon X \to \mathbb{R}$ satisfying

$$\phi(x) \leq p(x) \qquad (x \in X). \tag{5:4}$$

Proof. It follows from the Zorn's lemma that among all possible linear extensions $\psi\colon V \to \mathbb{R}$ of ϕ_o (where V is a linear subspace of X containing X_o) satisfying appropriate inequality (5:4) (for all $x \in V$) there exists a maximal functional, say $\phi\colon W \to \mathbb{R}$. It remains to show that $W = X$. To this end, we assume that $W \neq X$ and take any $u \in X\setminus W$. To obtain a contradiction, it is enough to find a linear extension $\psi\colon V \to \mathbb{R}$ of ϕ fulfilling (5:4) (for all $x \in V$) where $V = W + \mathbb{R}u$. So, we are looking for a real number c (which we will assign to u) such that $(\psi(x + tu) =)\ \phi(x) + tc \leq p(x + tu)$ for any $x \in W$ and $t \in \mathbb{R}$. Note that c has to satisfy the following inequalities:

- $c \leq \frac{p(x+tu)-\phi(x)}{t}$ and
- $c \geq \frac{\phi(x)-p(x-tu)}{t}$

for all $t > 0$ and $x \in W$. Equivalently:

$$\sup\Big\{\frac{\phi(x) - p(x - tu)}{t}\colon\ x \in W,\ t > 0\Big\} \leq c \leq \inf\Big\{\frac{p(x + tu) - \phi(x)}{t}\colon\ x \in X,\ t > 0\Big\}. \tag{5:5}$$

Observe that there exists c satisfying the above inequalities (and, consequently, ϕ can be extended to $\psi\colon V \to \mathbb{R}$) iff

$$\sup\Big\{\frac{\phi(x) - p(x - tu)}{t}\colon\ x \in W,\ t > 0\Big\} \leq \inf\Big\{\frac{p(x + tu) - \phi(x)}{t}\colon\ x \in X,\ t > 0\Big\}.$$

To show that the above inequality holds, it is enough to verify that $\frac{\phi(x)-p(x-tu)}{t} \leq \frac{p(y+su)-\phi(y)}{s}$ for all $x, y \in W$ and any positive scalars s and t. But the last inequality is equivalent to $\phi(sx + ty) \leq sp(x - tu) + tp(y + su)$, which is satisfied, since

$$\phi(sx + ty) \leq p(sx + ty) \leq p(sx - stu) + p(ty + stu) = sp(x - tu) + tp(y + su).$$

Consequently, there exists c satisfying (5:5) and therefore ϕ is not a maximal functional—and we are done. □

Most important consequences of Theorem 5.9 follow.

5.10 Theorem. (Classical Hahn-Banach Theorem)
For any continuous linear functional $\phi_o\colon E_o \to \mathbb{K}$ defined on a linear subspace E_o of a normed vector space $(E, \|\cdot\|_E)$ there exists $\phi \in E^*$ such that $\phi \restriction E_o = \phi_o$ and $\|\phi\| = \|\phi_o\|$.

Proof. First we consider the real case (that is, when $\mathbb{K} = \mathbb{R}$). Set $M \stackrel{\text{def}}{=} \|\phi_o\|$ and $p = M\|\cdot\|_E$, and note that p is a sublinear functional and $\phi_o \leq p \restriction E_o$. So, we infer from Theorem 5.9 that ϕ extends to a linear functional $\phi\colon E \to \mathbb{R}$ such that $\phi \leq p$. In particular, $\phi(x) \leq M\|x\|_E$ and $-\phi(x) \leq M\|-x\|_E$, hence $|\phi(x)| \leq M\|x\|_E$. This shows that $\phi \in X^*$ and $\|\phi\| \leq M$. But also $\|\phi\| \geq \|\phi_o\| = M$, which finishes the proof in the real case.

Now we pass to the complex case. Since any complex normed vector space is real as well, it follows from the first part of the proof that there exists a continuous $\mathbb{R}$-linear functional $\psi\colon E \to \mathbb{R}$ that extends $\operatorname{Re}\phi_o$ and has the same norm (as $\operatorname{Re}\phi_o$). We define $\phi\colon E \to \mathbb{C}$ by the rule $\phi(x) \stackrel{\text{def}}{=} \psi(x) - i\psi(ix)$. Observe that ϕ is $\mathbb{R}$-linear and continuous, and, since ψ is real-valued, $\operatorname{Re}\phi = \psi$. It is less obvious that $\phi(ix) = i\phi(x)$ (because $\phi(ix) = \psi(ix) - i\psi(-x) = i(\psi(x) - i\psi(ix))$). So, $\phi \in X^*$. Further, since both $\phi \restriction E_o$ and ϕ_o are $\mathbb{C}$-linear and their real parts coincide, these two functionals coincide as well (why?). In particular, $\|\phi\| \geq \|\phi_o\|$. Finally, for any $x \in E$ take a unit scalar γ such that $|\phi(x)| = \gamma\phi(x)$ and note that $|\phi(x)| = \phi(\gamma x) = \operatorname{Re}\phi(\gamma x) = \psi(\gamma x) \leq \|\operatorname{Re}\phi_o\| \cdot \|\gamma x\|_E \leq \|\phi_o\| \cdot \|x\|_E$, and we are done. □

5.11 Corollary. ("Norm extraction theorem")
For any non-zero vector x in a normed vector space E there exists $\phi \in E^*$ such that $\phi(x) = \|x\|_E$ and $\|\phi\| = 1$.

Proof. Consider $E_o \stackrel{\text{def}}{=} \mathbb{K}x$ and $\phi_o\colon E_o \ni wx \mapsto w\|x\|_E \in \mathbb{K}$, and apply Theorem 5.10. □

5.12 ***Remark.***
There are plenty of Banach spaces whose each non-zero vector admits a unique bounded linear functional satisfying the assertion of the above result. (In particular, all $\ell_p(X)$ and $L^p(\mu)$ spaces with $1 < p < \infty$.) Each such a Banach space is called *smooth*.

Corollary 5.11 implies

5.13 Corollary.
For any vector x in a normed vector space E, $\|x\|_E = \sup_{\phi \in \bar{B}_{E^*}} |\phi(x)|$.

5.14 Definition.
For any normed vector space E define $\kappa_E\colon E \to E^{**}$ by the rule:

$$\kappa_E(x) = \mathfrak{e}_x \quad \text{where} \quad \mathfrak{e}_x\colon X^* \ni \phi \mapsto \phi(x) \in \mathbb{K}.$$

The mapping κ_E is called the *canonical embedding* of E into its bidual.

5.15 Theorem.
For any normed vector space E, κ_E is linear and isometric.

Proof. Linearity of κ_E is left as an easy exercise. Its remaining property follows from Corollary 5.13, as: $\|\kappa_E(x)\| = \sup_{\phi \in \bar{B}_{E^*}} |\mathfrak{e}_x(\phi)| = \|x\|_E$. □

5.16 Corollary.
A subset A of a normed vector space E is bounded iff $\phi(A)$ is a bounded subset of $\mathbb{K}$ for any $\phi \in E^*$.

Proof. We may and do assume that A is non-empty. The 'only if' part of the result is trivial, whereas the 'if' part follows from the Uniform Boundedness Principle. Indeed, under the assumption formulated in the result, we get $\sup_{a \in A} |\kappa_E(a)(\phi)| < \infty$ for any $\phi \in E^*$. Consequently, since E^* is Banach, we infer from Theorem 5.1 (combined with Theorem 5.15) that $(\sup_{a\in A} \|a\| =) \ \sup_{a\in A} \|\kappa_E(a)\| < \infty$. □

5.17 Definition.
A Banach space is *reflexive* if its canonical embedding is surjective.

5.18 Proposition.
Each Hilbert space is reflexive.

Proof. Fix a Hilbert space $(H, \langle \cdot, - \rangle_H)$ and for each vector $h \in H$ set $\phi_h \stackrel{\text{def}}{=} \langle \cdot, h\rangle_H$. It follows from Theorem 4.19 (p. 13) that the operator $\Phi\colon H \ni h \mapsto \varphi_h \in H^*$ is bijective and isometric. It is easily seen that it is antilinear. In particular, the formula

$$\langle \phi_p, \phi_q \rangle_{H^*} \stackrel{\text{def}}{=} \langle q, p \rangle_H$$

correctly defines an inner product on H^* that induces the norm of that space. So, H^* is a Hilbert space and hence for any $\psi \in H^{**}$ there exists $\xi \in H^*$ such that $\psi = \langle \cdot, \xi \rangle_{H^*}$ (again, thanks to the Riesz representation theorem). But ξ is of the form $\xi = \phi_h$ for some $h \in H$. So, for any $x \in H$, $\psi(\phi_x) = \langle \phi_x, \phi_h \rangle_{H^*} = \langle h, x\rangle_H = \phi_x(h)$. Consequently, $\kappa_H(h) = \psi$ and we are done. □

5.19 ***Remark.***
One proves that all spaces $\ell_p(X)$ and $L^p(\mu)$ with $1 < p < \infty$ are reflexive. In Theorem 8.22, p. 42 (see Chapter 8) we will give an intrinsic characterisation of reflexive Banach spaces.

Note also that if κ_E is surjective for some normed vector space E, then E is automatically Banach (and hence it is a reflexive Banach space).

We end the chapter with an interesting application of the Hahn-Banach theorem, devoted to the so-called *means* or *Banach limits.*

5.20 Definition.
A linear functional $L\colon \ell_\infty^{\mathbb{R}} \to \mathbb{R}$ is called a *Banach limit* if for any sequence $(a_n)_{n=1}^\infty \in \ell_\infty^{\mathbb{R}}$:

- $\liminf_{n\to\infty} a_n \leq L((a_n)_{n=1}^\infty) \leq \limsup_{n\to\infty} a_n$, and
- $L((a_{n+1})_{n=1}^\infty) = L((a_n)_{n=1}^\infty)$.

5.21 ***Remark.***
It follows from the axioms of Banach limits that they are unit vectors of the dual of ℓ_∞^*.

Banach limits are special cases of more general *means* that are defined for arbitrary (non-empty) semi-groups. For any semi-group $(S, \cdot)$ a linear functional $M\colon \ell_\infty^{\mathbb{R}}(S) \to \mathbb{R}$ is said to be a (*left*) *mean* if for all $f \in \ell_\infty^{\mathbb{R}}(S)$ and each $s \in S$:

- $\inf f(S) \leq M(f) \leq \sup f(S)$, and
- $M(f_s) = M(f)$ where $f_s(x) = f(sx)$ $(x \in S)$.

In general, means are non-unique (on a fixed semi-group) and may not exist. The semi-group S is called (*left*)-*amenable* if it admits a left mean. There exist countable groups that are not amenable. On the other hand, one proves that all Abelian semi-groups are amenable. Amenability of $(\mathbb{N}, +)$ is established in the next result.

5.22 Theorem.
There exists a Banach limit.

Proof. Recall that a real-valued sequence $(a_n)_{n=1}^\infty$ is called *Cesàro summable* if there exists a real number A such that $\lim_{n\to\infty} \frac{a_1+\ldots+a_n}{n} = A$. The above number A is called the *Cesàro sum* of $(a_n)_{n=1}^\infty$. For the purpose of this proof we will denote this term by $\text{C-lim}_{n\to\infty} a_n$.

Denote by V and $C\colon V \to \mathbb{R}$ and $p\colon \ell_\infty^{\mathbb{R}} \to \mathbb{R}$, respectively, the space consisting of all Cesàro summable bounded real-valued sequences, and functionals given by the formulas:

$$C((a_n)_{n=1}^\infty) = \text{C-}\lim_{n\to\infty} a_n \quad ((a_n)_{n=1}^\infty \in V),$$
$$p((a_n)_{n=1}^\infty) = \limsup_{n\to\infty} a_n \quad ((a_n)_{n=1}^\infty \in \ell_\infty^{\mathbb{R}}).$$

Observe that V is a linear subspace of $\ell_\infty^{\mathbb{R}}$, C is linear and p is sublinear. Moreover, $C \leq p \restriction V$ (exercise). So, it follows from Theorem 5.9 that C extends to a linear functional $L\colon \ell_\infty^{\mathbb{R}} \to \mathbb{R}$ such that $L \leq p$. We will now show that L is a Banach limit. To this end, fix a bounded real-valued sequence $(a_n)_{n=1}^\infty$ and note that $L(-(a_n)_{n=1}^\infty) \leq p(-(a_n)_{n=1}^\infty)$, which yields $L((a_n)_{n=1}^\infty) \geq -\limsup_{n\to\infty}(-a_n) = \liminf_{n\to\infty} a_n$. Of course, $L((a_n)_{n=1}^\infty) \leq \limsup_{n\to\infty} a_n$. So, it remains to check that $L((a_n)_{n=1}^\infty - (a_{n+1})_{n=1}^\infty) = 0$. To obtain this equation, it is sufficient to verify that $(a_n - a_{n+1})_{n=1}^\infty$ is Cesàro convergent and its Cesàro sum is equal to 0 (because L extends C). But both these properties easily follow from the boundedness of $(a_n)_{n=1}^\infty$: setting $b_n \stackrel{\text{def}}{=} a_n - a_{n+1}$, we get $|\frac{b_1+\ldots+b_n}{n}| = \frac{|a_1-a_{n+1}|}{n} \leq \frac{2\|(a_k)_{k=1}^\infty\|_\infty}{n}$, which finishes the proof. $\square$

5.23 Definition.
For any subset A of a normed vector space E we define the *annihilator* $A^\perp$ of A as the set of all $\phi \in E^*$ that vanish at all points of A; that is, $\phi \in E^*$ belongs to $A^\perp$ if $\phi(a) = 0$ for all $a \in A$.

Similarly, for any set $B \subset E^*$, the *preannihilator* ${}^\perp B$ of B is the set of all points $x \in E$ at which all functionals from B vanish; in other words, if $B \subset E^*$ is non-empty, then ${}^\perp B = \bigcap_{\phi\in B} \mathcal{N}(\phi)$.

5.24 ***Remark****.*
For any Hilbert space H the assignment $H \ni h \mapsto \langle \cdot, h\rangle \in H^*$ establishes a **natural** / **canonical** one-to-one correspondence between bounded linear functionals on H and vectors from H. Under this identification, for any set A, the annihilator $A^\perp$ of A (introduced in Definition 5.23), coincides with the set, denoted by the same symbol ($A^\perp$), of all vectors that are orthogonal to A (cf. Definition 4.12, p. 11). So, no essential confusion occurs.

It is also worth remembering that for any set $B \subset E^*$ (where E is a normed vector space), ${}^\perp B$ is a subset of E, whereas $B^\perp$ is a subset of E^{**}. It is an easy exercise to check that $\kappa_E({}^\perp B) = \kappa_E(E) \cap B^\perp$. However, $\kappa_E({}^\perp B)$ differs from $B^\perp$ in general.

5.25 Proposition.
Let E be a normed vector spaceand $A \subset E$ and $B \subset E^*$ be arbitrary sets.

(a) *$A^\perp$ is a closed linear subspace of E^* and ${}^\perp B$ is a closed linear subspace of E.*

(b) *$A^\perp = (\overline{\text{lin}}(A))^\perp$ and ${}^\perp B = {}^\perp(\overline{\text{lin}}(B))$.*

(c) *${}^\perp(A^\perp) = \overline{\text{lin}}(A)$.*

Proof. Items (a) and (b) are left to the reader. Here we focus only on (c). Since A is contained in ${}^\perp(A^\perp)$, it follows from (a) that also $F \stackrel{\text{def}}{=} \overline{\text{lin}}(A) \subset {}^\perp(A^\perp)$. To see the reverse implication, fix an arbitrary vector $a \in E \setminus F$ and define $\phi_o\colon F + \mathbb{K}a \to \mathbb{K}$ by $\phi_o(f + ta) = t$ (where $f \in F$ and $t \in \mathbb{K}$). Since $\mathcal{N}(\phi_o) = F$, we infer from Corollary 2.15 (p. 5) that ϕ_o is bounded. Now it follows from the Hahn-Banach theorem that ϕ_o extends to $\phi \in E^*$. Then $\phi \in A^\perp$, but $\phi(a) \neq 0$ and therefore $a \notin {}^\perp(A^\perp)$, which finishes the proof. $\square$

It is worth underlying here that a similar result as item (b) above does not hold for the annihilator of the pre-annihilator (that is, in general $({}^\perp B)^\perp$ differs from $\overline{\text{lin}}(B)$ for a subset B of the dual E^* of a Banach space E; cf. Proposition 8.16, p. 39, in Secion 8).

Annihilators play important role in describing the dual spaces of a subspace and of a quotient space, as shown by

5.26 Theorem.
Let E be a normed vector space.

(A) *For a closed linear subspace F of E, the assignment*

$$(E/F)^* \ni \phi \mapsto \phi \circ \pi \in F^\perp$$

correctly defines a bijective linear isometry Φ where $\pi\colon E \to E/F$ is the quotient map and E/F is endowed with the quotient norm (introduced in Theorem 2.8, p. 3).

(B) *For any linear subspace V of E the rule*

$$V^* \ni \phi \restriction V \mapsto \Pi(\phi) \in E^*/V^\perp \qquad (\phi \in E^*)$$

correctly defines a bijective linear isometry Ψ where $\Pi\colon E^* \to E^*/V^\perp$ is the quotient map and $E^*/V^\perp$ is endowed with the quotient norm.

Proof. We start from (A). Observe that Φ is a well defined linear operator. Moreover, since $\pi(B_E) = B_{E/F}$, we infer that Φ is isometric. Finally, if $\psi \in E^*$ vanishes at all points of F, we may correctly define a linear functional $\phi\colon E/F \to \mathbb{K}$ by the rule $\phi \circ \pi = \psi$. Then $\phi(B_{E/F}) = \psi(B_E)$ and hence ϕ is bounded. So, $\Phi(\phi) = \psi$ and we are done.

We pass to (B), which is much less elementary than (A). Firstly, it follows from the Hahn-Banach theorem that each $\alpha \in V^*$ can be written in the form $\phi \restriction V$ for certain $\phi \in E^*$. Secondly, if ϕ_1 and ϕ_2 are two functionals from E^* that coincide on V, then $\phi_1 - \phi_2 \in V^\perp$ and, consequently, $\Pi(\phi_1) = \Pi(\phi_2)$. These two remarks explain why Ψ is well (and fully) defined. Observe also that Ψ is a linear surjection. So, it remains to show it is isometric. To this end, note that for each $\alpha \in V^*$ there is $\phi \in E^*$ such that $\alpha = \phi \restriction V$ and $\|\alpha\| = \|\phi\|$ (thanks to Theorem 5.10). So, $\Psi(\alpha) = \Pi(\phi)$ and hence $\|\Psi(\alpha)\| \leq \|\phi\| = \|\alpha\|$. On the other hand, $\|\Psi(\alpha)\| = \inf\{\|\phi - \beta\|\colon\ \beta \in V^\perp\} \geq \inf\{\|(\phi - \beta) \restriction V\|\colon\ \beta \in V^\perp\} = \|\alpha\|$, which finishes the proof. □

6 Metrisable topological vector spaces

6.1 Definition.
Let E be a vector space. A semi-metric $d\colon E \times E \to \mathbb{R}_+$ is *invariant* if $d(x+z, y+z) = d(x, y)$ for all $x, y, z \in E$.

A *value* on E is a function $p\colon E \to \mathbb{R}_+$ that satisfies all the following conditions (for all $x, y \in E$):

(v0) $p(x) = 0 \iff x = 0$;

(v1) $p(-x) = p(x)$;

(v2) $p(x + y) \leq p(x) + p(y)$.

If p fulfills only (v1)–(v2) and vanishes at the origin of E, it is called a *semi-value*. A semi-value $q\colon E \to \mathbb{R}_+$ is said to be

- *balanced* if $q(\gamma x) = q(x)$ for all $x \in E$ and $\gamma \in \mathbb{K}$ with $|\gamma| = 1$;
- *monotone* if for any vector $x \in E$ the function $\mathbb{R}_+ \ni t \mapsto q(tx) \in \mathbb{R}_+$ is monotone increasing.

Each invariant [semi-]metric $\rho\colon E \times E \to \mathbb{R}_+$ induces a [semi-]value $q_\rho\colon E \ni x \mapsto \rho(x, 0) \in \mathbb{R}_+$ on E. Conversely, each [semi-]value $p\colon E \to \mathbb{R}_+$ induces an invariant [semi-]metric $d_p\colon E \times E \ni (x, y) \mapsto p(x - y) \in \mathbb{R}_+$ on E. Actually, there is a one-to-one correspondence between [semi-]values and invariant [semi-]metrics.

6.2 Proposition.
Let p be a value on a vector space E that is both balanced and monotone. Then (E, d_p) is a topological vector space iff $\inf_{t>0} p(tx) = 0$ for any $x \in E$.

Proof. Necessity is clear. To show sufficiency, we fix three seqeunces $(\alpha_n)_{n=1}^\infty \subset \mathbb{K}$, $(x_n)_{n=1}^\infty \subset E$ and $(y_n)_{n=1}^\infty \subset E$ that converge, respectively, to $\beta \in \mathbb{K}$, $a \in E$ and $b \in E$ (the last two convergences are w.r.t. d_p). We need to check that then the numbers $d_p(\alpha_n x_n + y_n, \beta a + b)$ converge to 0. Since p is balanced, we obtain:

$$d_p(\alpha_n x_n + y_n, \beta a + b) = p(\alpha_n x_n - \beta a + y_n - b) \le p(\alpha_n(x_n - a)) + p((\alpha_n - \beta)a) + p(y_n - b)$$
$$= p(|\alpha_n|(x_n - a)) + p(|\alpha_n - \beta|a) + p(y_n - b).$$

Further, it follows from the monotonicity of p that:

- $p(|\alpha_n|(x_n - a)) \le p(N(x_n - a))$ where $N > 0$ is an integer such that $|\alpha_n| \le N$ for all $n > 0$;
- there exists $\lim_{s\to 0^+} p(sa)$.

The latter property, combined with our assumption about p (in the statement of the result), implies that $\lim_{s\to 0^+} p(sa) = 0$. Therefore, setting $s_n \stackrel{\text{def}}{=} |\alpha_n - \beta|$ and continuing our previous estimations, we get:

$$d_p(\alpha_n x_n + y_n, \beta a + b) \le p(|\alpha_n|(x_n - a)) + p(s_n a) + p(y_n - b) \le p(N(x_n - a)) + p(s_n a) + p(y_n - b)$$
$$\le N p(x_n - a) + p(s_n a) + p(y_n - b) = N d_p(x_n, a) + p(s_n a) + d_p(y_n, b) \to 0 \quad (n \to \infty)$$

and we are done. □

The following result is a special case of a more general Kakutani-Birkhoff theorem (on metrisability of topological groups).

6.3 Theorem. (Metrisability of TVS's)
A TVS is metrisable iff it is first countable and T_0. Moreover, for any T_0-topological vector space E that has a countable basis of 0-neighbourhoods there exists a value p that is both balanced and monotone, and for which d_p is compatible with the given topology of E.

Proof. It is sufficient to prove the second claim of the theorem. For clarity, denote by τ the topology of E. It follows from the first countability of E that there exist 0-neighbourhoods $U_0 = E, U_1, U_2, \ldots$ that form a basis of neighbourhoods of the origin and have the following properties:

- $U_n + U_n \subset U_{n-1}$ for each $n > 0$;
- each U_n is both open and balanced.

Since E is T_0, we have

$$\bigcap_{n=1}^{\infty} U_n = \{0\}. \tag{6:1}$$

Denote by I the set of all rationals of the form $\frac{k}{2^n}$ where k and n are positive integers. We define the sets $\{V_w\}_{w\in I}$ recursively as follows:

- $V_w = E$ for each $w \ge 1$ from I;
- $V_{2^{-n}} = U_n$ for any $n > 0$;
- $V_{\frac{k}{2^n}} = V_{\frac{k-1}{2^n}} + U_n$ for any odd $k > 1$ and $n > 0$.

Equivalently, if $w < 1$ has the form $w = \sum_{j=1}^p 2^{-k_j}$ where $k_1, \ldots, k_p$ are all positive and distinct, then

$$V_w = U_{k_1} + \ldots + U_{k_p}. \tag{6:2}$$

To simplify further arguments, we will write $\sum_{j=1}^p U_{k_j}$ to denote the set $U_{k_1} + \ldots + U_{k_p}$.

Since the algebraic sum of two open (resp. balanced) sets is open (resp. balanced) as well, we infer that

($\star$) for any $w \in I$, V_w is an open and balanced 0-neighbourhood.

Our nearest aim is to show the following important property:

$$V_s + V_t \subset V_{s+t} \qquad (s, t \in I). \tag{6:3}$$

We start from the proof of a special case of (6:3):

$$V_s + U_n \subset V_{s+2^{-n}} \qquad (s \in I \cap (0,1),\ n \geq 0). \tag{6:4}$$

To this end, express s in the form $s = \sum_{j=1}^{p} 2^{-k_j}$ where $0 < k_1 < \ldots < k_p$. We will show (6:4) by induction on n. If $n = 0$, (6:4) is trivial. Now assume $n > 0$ and that (6:4) holds for $n-1$. If $k_q = n$ for some $q \in \{1,\ldots,p\}$, then (since $U_n + U_n \subset U_{n-1}$):

$$V_s + U_n = \Big(\sum_{j=1}^{p} U_{k_j}\Big) + U_n = \Big(\sum_{j\neq q} U_{k_j}\Big) + U_n + U_n \subset \Big(\sum_{j\neq q} U_{k_j}\Big) + U_{n-1} = V_{\sum_{j\neq q} 2^{-k_j}} + U_{n-1}$$

and it follows from the induction hypothesis that $V_{\sum_{j\neq q} 2^{-k_j}} + U_{n-1} \subset V_w$ where $w = (\sum_{j\neq q} 2^{-k_j}) + 2^{1-n} = s + 2^{-n}$ which yields (6:4). Finally, if $k_j \neq n$ for any j, then $V_s + U_n = (\sum_{j=1}^{p} U_{k_j}) + U_n = V_{s+2^{-n}}$ (thanks to (6:2)). So, (6:4) has been proved.

Now we turn to the proof of (6:3). To this end, we fix $s, t \in I$. We may and do assume that $s + t < 1$. Express t in the form $t = \frac{k}{2^n}$ where k is odd. We will show (6:3) by induction on k. The case $k = 1$ is covered by (6:4). Hence, we assume $k > 1$ (is odd) and that (6:3) holds whenever the odd numerator of the second index therein is less than k. But then $V_t = V_w + U_n$ where $w = \frac{k-1}{2^n}$. Note that the odd numerator of w is less than k and therefore $V_s + V_w \subset V_{s+w}$ (by the induction hypothesis). Finally, an application of (6:4) yields $V_{s+w} + U_n \subset V_{s+w+2^{-n}}$ and thus:

$$V_s + V_t = V_s + V_w + U_n \subset V_{s+w} + U_n \subset V_{s+w+2^{-n}} = V_{s+t},$$

which finishes the proof of (6:3).

Now we define $p\colon E \to [0,1]$ by:

$$p(x) \stackrel{\text{def}}{=} \inf\{w \in I\colon\ x \in V_w\}$$

(recall that $V_1 = E$). It readily follows from (6:1), ($\star$) and (6:3) that p is a value on E that is both balanced and monotone. Moreover, it is easy to check that $\inf_{t>0} p(tx) = 0$ for all $x \in E$ (since each V_w is a balanced 0-neighbourhood). We infer from Proposition 6.2 that (E, d_p) is a TVS. Thus, it remains to check that the identity map between (E,τ) and (E, d_p) is continuous at the origin in both directions. But both these properties follow from the following two inclusions (whose simple proof is left to the reader):

$$B_{d_p}(0,r) \subset V_r \subset \bar{B}_{d_p}(0,r) \qquad (r \in I).$$

□

6.4 Definition.
A net $(x_\sigma)_{\sigma\in\Sigma}$ in a TVS E is called *Cauchy* (or *fundamental*) if for any 0-neighbourhood U in E there exists an index $\omega \in \Omega$ such that $x_\sigma - x_\tau \in U$ for all indices $\sigma, \tau \in \Sigma$ satisfying $\sigma \geq \omega$ and $\tau \geq \omega$.

A T_2VS is said to be *complete* if any its Cauchy net is convergent.

6.5 Proposition.

(A) *A convergent net in a TVS is Cauchy.*

(B) *Let E be a TVS whose topology is induced by an invariant metric d. Then:*

- *The metric d is complete iff E is a complete TVS.*
- *The metric completion of (E,d) admits a natural TVS structure that extends the structure of the space E and with respect to which its metric is invariant.*

(*proof—exercise*)

The following is a counterpart of Theorem 2.8 (p. 3) for metrisable TVS's. Since the proof goes almost the same manner, we skip it.

6.6 Theorem.
Let F be a closed linear subspace of a TVS E whose topology is induced by an invariant metric D, and let $\pi\colon E \to E/F$ denote the quotient map. Then the formula

$$q(b) \stackrel{\text{def}}{=} \inf\{D(a,0)\colon\ a \in E,\ \pi(a) = b\}$$

defines a value on E/F such that d_q is compatible with the quotient topology. Moreover, if D is complete, then d_q is complete as well.

(*proof—exercise*)

6.7 Definition.
An *F-space* is a pair (E, d) where E is a metrisable TVS and d is a compatible metric that is both invariant and complete.

It turns out that Isomorphism Theorem, Open Mapping Theorem and Closed Graph Theorem hold also for F-spaces, as shown by the next three results.

6.8 Theorem. (Generalised Open Mapping Theorem)
Let (E, d_E) and (F, d_F) be two F-spaces. For a continuous linear operator $T\colon E \to F$ the following conditions are equivalent:

(i) *T is an open mapping;*

(ii) *T is surjective;*

(iii) *$\mathcal{R}(T)$ is of second Baire category (that is, it cannot be expressed as a countable union of nowhere dense sets).*

Proof. Implications (i) $\implies$ (ii) $\implies$ (iii) are left as simple exercises. Here we focus only on the hardest part; that is, we will show that (i) is followed by (iii). To this end, it is sufficient (why?) to show that for any $r > 0$, the set $T(B_{d_E}(0, r))$ is a 0-neighbourhood.

Observe that $E = \bigcup_{n=1}^{\infty} nB_{d_E}(0, r)$ (because 0-neighbourhoods are absorbing; cf. the proof of Theorem 2.16, p. 5). So, $\mathcal{R}(T) = \bigcup_{n=1}^{\infty} nT(B_{d_E}(0, r))$ and it follows from our assumption in (iii) that the closure of $nT(B_{d_E}(0, r))$ has non-empty interior for some $n > 0$. Since the multiplication by n is a homeomorphism of F, we infer that the closure of $T(B_{d_E}(0, r))$ has non-empty interior. This property is valid for all $r > 0$. In particular, there is a non-empty open set $V \subset F$ that is contained in the closure C of $T(B_{d_E}(0, r/2))$. But then $U \stackrel{\text{def}}{=} V - V$ is contained in $C - C$ and the latter set is contained in the closure of $T(B_{d_E}(0, r/2) - B_{d_E}(0, r/2))$ (which follows from the continuity of $(x, y) \mapsto x - y$ in F). However, U is a 0-neighbourhood in F and $B_{d_E}(0, r/2) - B_{d_E}(0, r/2) \subset B_{d_E}(0, r)$, which yields:

$$0 \in \operatorname{int} \overline{T(B_{d_E}(0, r))}. \tag{6:5}$$

Again, the above relation is valid for all positive r.

Using (6:5) and starting from $\varepsilon_0 \stackrel{\text{def}}{=} r$, we inductively construct positive numbers $\varepsilon_1, \varepsilon_2, \ldots$ such that for all $n > 0$:

- $\varepsilon_n \leq 2^{-n} r$; and
- $\bar{B}_{d_F}(0, \varepsilon_n) \subset \overline{T(B_{d_E}(0, \varepsilon_{n-1}))}$.

In particular, for all $n > 0$:

$$\forall y \in \bar{B}_{d_F}(0, \varepsilon_n)\ \exists x \in E\colon\ d_E(x, 0) < \varepsilon_{n-1}\ \wedge\ d_F(y - T(x), 0) \leq \varepsilon_{n+1}. \tag{6:6}$$

We will show that $B_{d_F}(0, \varepsilon_1) \subset T(B_{d_E}(0, 2r))$, which will finish the proof (because of the arbitrarity of r).

Take an arbitrary $y \in B_{d_F}(0, \varepsilon_1)$. It follows from (6:6) that there exists $x_1 \in B_{d_E}(0, \varepsilon_0)$ such that $d_F(y - T(x_1), 0) \leq \varepsilon_2$. Now assume that we have already defined $x_1, \ldots, x_k$ (for some $k > 0$) such that $x_j \in B_{d_E}(0, \varepsilon_{j-1})$ (for $j = 1, \ldots, k$) and $d_F(y - \sum_{j=1}^{k} T(x_j), 0) \leq \varepsilon_{k+1}$. Then $y - \sum_{j=1}^{k} T(x_j) \in B_{d_F}(0, \varepsilon_{k+1})$ and we conclude from (6:6) that there exists $x_{k+1} \in B_{d_E}(0, \varepsilon_k)$ with $d_F((y - \sum_{j=1}^{k} T(x_j)) - T(x_{k+1})) \leq \varepsilon_{k+2}$.

In this way we have constructed a sequence $(x_n)_{n=1}^{\infty} \subset E$ such that for all $n > 0$:

- $d_E(x_n, 0) < \varepsilon_{n-1}$; and
- $d_F(y, T(\sum_{k=1}^n x_k)) \leq \varepsilon_{n+1}$ $(\to 0)$.

Since d_E is complete and $\sum_{n=0}^\infty \varepsilon_n < \infty$, we infer from the triangle inequality that the series $\sum_{n=1}^\infty x_n$ is convergent, say to a. Then $d(a, 0) < \sum_{n=0}^\infty \varepsilon_n \leq 2r$ and $T(a) = \lim_{n\to\infty} T(\sum_{k=1}^n x_k) = y$, and we are done. □

As immediate consequences of the above theorem, we obtain the following two result (for a proof of the latter, consult the proof of Theorem 5.5).

6.9 Theorem. (Isomorphism Theorem)
Let $T\colon X \to Y$ be a continuous bijective linear operator between two F-spaces. Then the inverse T^{-1} of T is continuous as well.

6.10 Theorem. (Closed Graph Theorem)
Let $T\colon X \to Y$ be a linear operator between F-spaces X and Y. Then T is continuous iff the following condition is satisfied:

$$(x_n \in X,\ \lim_{n\to\infty} x_n = 0,\ \lim_{n\to\infty} T(x_n) = y \in Y) \implies y = 0.$$

*6.11 **Remark**.*
One proves that a metrisable TVS is completely metrisable iff it is complete as a TVS. A most common proof of this result uses Theorem 6.3, Proposition 6.5, the Alexandrov-Hausdorff theorem (that characterises completely metrisable spaces among metrisable ones as absolute $\mathscr{G}_\delta$ sets) and the Baire category theorem. We leave the details to interested readers as a (little) challenge.

7 Locally convex spaces

7.1 Definition.
A *locally convex* space is a TVS that possesses a basis of 0-neighbourhoods consisting of convex sets.

7.2 Proposition.

(A) *Both the interior and the closure of a convex set in a TVS is convex as well.*

(B) *Each convex 0-neighbourhood in a TVS contains a 0-neighbourhood that is both open and absolutely convex.*

(*proof—exercise*)

7.3 Definition.
Let K be a convex 0-neighbourhood in a TVS E. The *Minkowski functional* of K is a function $p_K\colon E \to \mathbb{R}_+$ defined as follows:

$$p_K(x) = \inf\{r > 0\colon\ x \in rK\}$$

(the set appearing on the right-hand side of the above formula is non-empty because each 0-neighbourhood is absorbing).

Basic properties of Minkowski functionals are listed below.

7.4 Proposition.
Let K and L be two convex 0-neighbourhoods in a TVS E.

(a) *The function p_K is sublinear and continuous.*

(b) *If K is absolutely convex, then p_K is a semi-norm.*

(c) $p_K^{-1}([0,1)) \subset K \subset p_K^{-1}([0,1])$.

(d) *If $K \subset L$, then $p_L \leq p_K$.*

Proof. Since both (c) and (d) easily follow from the defining formulas of p_K and p_L, we will show only (a) and (b). Positive homogeneity of p_K (that is, the equation $p_K(tx) = tp_K(x)$ for a scalar $t > 0$) is also a direct consequence of the defining formula. To show the triangle inequality, fix arbitrary $x, y \in E$ and real numbers s and t such that $p_K(x) < s$ and $p_K(y) < t$. There are positive reals a and b such that $a < s$, $b < t$ and $x \in aK$ and $y \in bK$. Then both $\frac{1}{a}x$ and $\frac{1}{b}y$ belong to K and it follows from the convexity of K that $\frac{a}{a+b} \cdot \frac{1}{a}x + \frac{b}{a+b} \cdot \frac{1}{b}y = \frac{x+y}{a+b}$ lies in K as well. Consequently, $p_K(x+y) \leq a+b < s+t$ and hence $p_K(x+y) \leq p_K(x) + p_K(y)$. If, in addition, K is absolutely convex, then $p_K(\gamma x) = p_K(x)$ for any unit scalar $\gamma \in \mathbb{K}$ (because $\gamma K = K$ for such γ) which implies that p_K is a semi-norm. It remains to check that p_K is continuous. Since $|p_K(x) - p_K(y)| \leq \max(p_K(x-y), p_K(y-x))$, it is enough to show that p_K is continuous at the origin of E. And the last property is a consequence of the relation $\varepsilon K \subset p_K^{-1}([0,\varepsilon])$ (for any $\varepsilon > 0$). □

7.5 Definition.
Let $\mathcal{P} = \{p_s\}_{s\in S}$ be a collection of semi-norms defined on a common vector space E. Topology $\tau_{\mathcal{P}}$ on E *induced by* $\mathcal{P}$ is given by a basis of (open) 0-neighbourhoods of the form:

$$U(P,\varepsilon) \stackrel{\text{def}}{=} \{x \in E |\ \forall p \in P\colon\ p(x) < \varepsilon\} \qquad (\varepsilon > 0,\ P \subset \mathcal{P} \text{ finite}). \tag{7:1}$$

In other words, a set $V \subset E$ belongs to $\tau_{\mathcal{P}}$ iff for any $x \in V$ there are a finite set $P \subset \mathcal{P}$ and a real number $\varepsilon > 0$ such that $x + U(P,\varepsilon) \subset V$. It is not difficult to verify that $(E, \tau_{\mathcal{P}})$ is a topological vector space. Since the sets $U(P,\varepsilon)$ are convex, we conclude that this space is locally convex.

$\mathcal{P}$ is said to be *separating* if for any non-zero vector $u \in E$ there exists an index $s \in S$ such that $p_s(u) > 0$.

7.6 Proposition.
Let $\mathcal{P}$ be a collection of semi-norms on a vector space E.

- *A net $(x_\sigma)_{\sigma\in\Sigma} \subset E$ converges to $w \in E$ in the topology induced by $\mathcal{P}$ iff $\lim_{\sigma\in\Sigma} p(x_\sigma - w) = 0$ for all $p \in \mathcal{P}$.*
- *$(E, \tau_{\mathcal{P}})$ is Hausdorff iff $\mathcal{P}$ is separating.*
- *Each semi-norm from $\mathcal{P}$ is continuous w.r.t. $\tau_{\mathcal{P}}$.*

(*proof—exercise*)

7.7 Theorem. (Defining locally convex topology by semi-norms)
For every locally convex space (E,τ) there is a collection $\mathcal{P}$ of semi-norms on E such that $\tau = \tau_{\mathcal{P}}$.

Proof. Let $\mathscr{B}$ be any basis of 0-neighbourhoods that consists of absolutely convex sets. (Such a basis exists thanks to Proposition 7.2 and local convexity of E.) Let $\mathcal{P}$ consist of all semi-norms of the form p_K with $K \in \mathscr{B}$ (cf. item (b) of Proposition 7.4). Since $\mathcal{P}$ consists of continuous functions, we infer that $U(P,\varepsilon) \in \tau$ and, consequently, $\tau_{\mathcal{P}} \subset \tau$. On the other hand, if V is a 0-neighbourhood in (E,τ), then $K \subset V$ for some $K \in \mathscr{B}$. But $U(P,1) \subset K$ for $P \stackrel{\text{def}}{=} \{p_K\}$ (by item (c) of Proposition 7.4) and hence $U(P,1) \subset V$, which shows that the identity map from $(E,\tau_{\mathcal{P}})$ to (E,τ) is continuous at the origin. So, $\tau \subset \tau_{\mathcal{P}}$, and we are done. □

7.8 Corollary. (Metrisability of locally convex spaces)
A locally convex space is metrisable iff its topology is induced by a separating countable collection of semi-norms. Moreover, for a separating collection $\mathcal{P} = \{p_1, p_2, p_3, \ldots\}$ of semi-norms on a vector space E, the formula

$$q_{\mathcal{P}}(x) = \sum_{n=1}^{\infty} \frac{p_n(x)}{2^n(1+p_n(x))} \qquad (x \in E)$$

correctly defines a balanced monotone value on E such that the metric d_q is compatible with the topology $\tau_{\mathcal{P}}$.

Proof. The 'only if' part follows from the previous proof (since $\mathscr{B}$ specified therein may be countable). On the other hand, the 'if' part follows from the additional claim, whose proof is left as an exercise. □

7.9 Proposition.
Let the topology of a locally convex space E be induced by a family $\mathcal{P}$ of semi-norms. A linear functional $f\colon E \to \mathbb{K}$ is continuous iff there exist a finite set $P \subset \mathcal{P}$ and a real number $M > 0$ such that

$$|f(x)| \leq M \sum_{p \in P} p(x) \qquad (x \in E). \tag{7:2}$$

Proof. Sufficiency of (7:2) is left to the reader. We will show only the main part of this result—namely, necessity of (7:2). So, we assume f is continuous and take a finite set $P \subset \mathcal{P}$ and $\varepsilon > 0$ such that $f(U(P,\varepsilon)) \subset B_{\mathbb{K}}$. Set $M \stackrel{\text{def}}{=} \frac{2}{\varepsilon}$ and fix arbitrary $x \in E$. If $S \stackrel{\text{def}}{=} \sum_{p\in P} p(x)$ equals zero, then $tx \in U(P,\varepsilon)$ for each $t > 0$ and hence $|f(tx)| < 1$ (for all such t). Consequently, $f(x) = 0$ and (7:2) holds for x. In the other case, that is, when $S > 0$, then $\frac{\varepsilon}{2S}x \in U(P,\varepsilon)$ and therefore $|f(\frac{\varepsilon x}{2S})| < 1$. Equivalently, $|f(x)| \leq MS$, and we are done. □

As a consequence of the above result, we obtain

7.10 Corollary.
Let E be a locally convex space. Any continuous linear functional $\phi_o\colon E_o \to \mathbb{K}$ defined on a linear subspace E_o of E extends to a continuous linear functional $\phi\colon E \to \mathbb{K}$.

Proof. Let $\mathcal{P}$ be a collection of semi-norms on E that induces the topology of this space (see Theorem 7.7). Then the collection $\mathcal{P}_o \stackrel{\text{def}}{=} \{p \restriction E_o\colon\ p \in \mathcal{P}\}$ induces the topology of E_o. So, we infer from Proposition 7.9 that

$$|\phi_o(x)| \leq M \sum_{p\in P} p(x) \qquad (x \in E_o)$$

for certain constant $M > 0$ and a finite set $P \subset \mathcal{P}$. Mimicing the proof of Theorem 5.10 (that was presented for norms, but works perfectly also for semi-norms), we can construct a linear functional $\phi\colon E \to \mathbb{K}$ that extends ϕ_o and satisfies $|\phi(x)| \leq M \sum_{p\in P} p(x)$ for all $x \in E$. Consequently, ϕ is continuous (by Proposition 7.9), and we are done. □

A natural question of when a TVS is normable was answered by Kolmogorov with the aid of the following (classical) notion.

7.11 Definition.
A set A in a topological vector space E is *bounded* (in the sense of TVS's) if for any 0-neighbourhood U in E there is a positive real number r such that $A \subset rU$.

7.12 Example.
Since $E = \bigcup_{n=1}^{\infty} n\,\mathrm{int}(U)$ for any 0-neighbourhood U in a TVS E, it follows that compact subsets of TVS's are bounded.

7.13 Proposition.
Let $(a_n)_{n=1}^{\infty}$ be a fixed sequence of positive real numbers that converge to 0. For a subset A of a topological vector space E the following conditions are equivalent:

(i) *A is bounded (in the sense of Definition 7.11);*

(ii) *$a_n x_n \to 0$ $(n \to \infty)$ for any $(x_n)_{n=1}^{\infty} \subset A$;*

(iii) *$t_n x_n \to 0$ $(n \to \infty)$ for all $(x_n)_{n=1}^{\infty} \subset A$ and $(t_n)_{n=1}^{\infty} \subset \mathbb{K}$ such that $\lim_{n\to\infty} t_n = 0$.*

(proof—exercise)

7.14 Theorem. (Kolmogorov normability theorem)
A topological vector space is normable iff it is T_0 and contains a 0-neighbourhood that is both convex and bounded.

Proof. Since all balls in normed vector spaces are bounded (in the sense of Definition 7.11 [!]), we only need to prove the 'if' part of the theorem. To this end, fix a $\mathrm{T_2VS}$ E and a convex 0-neighbourhood W that is bounded. Reducing the set W if necessary, we may and do assume that W is absolutely convex and open (note that a subset of a bounded set is bounded as well). It follows from the boundedness of W that the sets $\{rW\colon\ r > 0\}$ form a basis, to be denoted by $\mathscr{B}$, of 0-neighbourhoods of E. So, it follows from the proof of Theorem 7.7 that the collection $\{p_K\colon\ K \in \mathscr{B}\}$ induces the topology of E. But $p_{rW} = \frac{p_W}{r}$ (exercise), which implies that the topology of E is induces by the semi-norm p_W. Since E is T_0, we infer that p_W is a norm and we are done. □

The next two results are another central tools in functional analysis.

7.15 Theorem. (Separation of open convex sets)
Let U and V be two disjoint convex sets in a TVS E. If U is open, there exists a continuous linear functional $\phi\colon E \to \mathbb{K}$ and a real number t such that

$$\operatorname{Re}\phi(u) < t \leq \operatorname{Re}\phi(v) \qquad (u \in U,\ v \in V). \tag{7:3}$$

Proof. If U or V is empty, it suffices to set $\phi = 0$ (and $t = \pm 1$). Below we assume that both these sets are non-empty. We also fix $a \in U$ and $b \in V$.

Set $w = b - a$ and observe that w is non-zero and $W \stackrel{\text{def}}{=} U + (-V) + w$ is a 0-neighbourhood that is both open and convex. Let $p \stackrel{\text{def}}{=} p_W$ and $\psi\colon \mathbb{R}\cdot w \ni tw \mapsto tp(w) \in \mathbb{R}$. Since p is sublinear, we infer that $p(-tw) + p(tw) \geq 0$ for $t > 0$. Equivalently, $p(-tw) \geq -tp(w) = \psi(-tw)$ for all $t > 0$ and hence $\psi \leq p \restriction \mathbb{R}w$. It follows from the Hahn-Banach theorem that ψ extends to an $\mathbb{R}$-linear functional $\phi\colon E \to \mathbb{R}$ such that $\phi \leq p$. In particular, $\phi(w) = p(w)$ and $\phi(\varepsilon(-W \cap W)) \subset [-\varepsilon, \varepsilon]$ for each $\varepsilon > 0$ (why? cf. item (c) of Proposition 7.4). We conclude that ϕ is continuous. Further, since $U \cap V = \varnothing$, it follows that $0 \notin U + (-V)$ and, consequently, $w \notin W$. Hence $p(w) \geq 1$. So, if $u \in U$ and $v \in V$ are arbitrary, then $u - v + w \in W$ and thus $\phi(u - v + w) \leq p(u - v + w) \leq 1 \leq p(w) = \phi(w)$, which yields

$$\phi(u) \leq \phi(v) \qquad (u \in U,\ v \in V).$$

Set $t = \inf \phi(V)$. Then $\phi(u) \leq t \leq \phi(v)$ for all $u \in U$ and $v \in V$. Finally, note that for any $u \in U$ there is $\varepsilon > 0$ such that $u + \varepsilon w \in U$ (because U is open). Consequently, $\phi(u + \varepsilon w) \leq t$. This inequality, combined with $\phi(w) \geq 1$, yields that $\phi(u) < t$.

If the field is real, the proof is finished. And when $\mathbb{K} = \mathbb{C}$, it is sufficient to define $\Phi\colon E \to \mathbb{C}$ by $\Phi(x) = \phi(x) - i\phi(ix)$ to obtain a continuous $\mathbb{C}$-linear functional Φ for which $\operatorname{Re}(\Phi) = \phi$, which finishes the proof in that case. □

7.16 Theorem. (Separation of closed convex sets)
Let A and B be two disjoint convex sets in a ***locally convex*** *space E. If A is compact and B is closed, then there exists a continuous linear functional $\phi\colon E \to \mathbb{K}$ and two real numbers s and t with such that*

$$\operatorname{Re}\phi(a) \leq s < t \leq \operatorname{Re}\phi(b) \qquad (a \in A,\ b \in B). \tag{7:4}$$

Proof. As in the proof of the previous result, we may and do assume that both A and B are non-empty. Observe that $D \overset{\text{def}}{=} E \setminus B$ is an open set that contains A. So, for any point $a \in A$ there exists an open 0-neighbourhood W_a such that $a + W_a + W_a \subset D$. It follows from the compactness of A that there exists a finite non-empty set $F \subset A$ such that $A \subset \bigcup_{a \in F}(a + W_a)$. Since E is locally convex, there exists a 0-neighbourhood V that is open and convex and contained in $\bigcap_{a \in F} W_a$. We claim that

$$(A + V) \cap B = \varnothing. \tag{7:5}$$

Indeed, if $b \in A$ and $v \in V$, then there is $a \in F$ such that $b \in a + W_a$. Consequently, $b + v \in a + W_a + W_a \subset D$ and hence $b + v \notin B$. Now noting that $A + V$ is open and convex, we may apply Theorem 7.15 to obtain a continuous linear functional $\phi\colon E \to \mathbb{K}$ and a real number t such that $\operatorname{Re}\phi(A + V) \subset (-\infty, t)$ and $\operatorname{Re}\phi(B) \subset [t, \infty)$. Since A is compact, $s \overset{\text{def}}{=} \sup \operatorname{Re}\phi(A)$ is less than t, and the conclusion follows. □

As a consequence of the above result, we obtain important

7.17 Corollary.
For any two distinct points a and b of a locally convex T_2-space E there exists a continuous linear functional $\phi\colon E \to \mathbb{K}$ such that $\phi(a) \neq \phi(b)$.

Proof. Just apply Theorem 7.16 to $A \overset{\text{def}}{=} \{a\}$ and $B = \{b\}$. □

Our last topic of this chapter is related to extreme points in compact convex sets (which naturally generalise the notion of vertices of planar polygons). A formal definition is given below.

7.18 Definition.
Let A be a convex set in a vector space. A point $a \in A$ is said to be *extreme* (in A) if there are no points $x, y \in A$ distinct from a for which $a = \frac{x+y}{2}$. In other words, a is extreme in A iff:

$$\big(x, y \in A,\ a = \frac{x+y}{2}\big) \implies x = y = a.$$

The set of all extreme points of A is denoted by $\operatorname{ext}(A)$. We will denote its *closed convex hull* (that is, the smallest set containing $\operatorname{ext}(A)$ that is both closed and convex) by $\overline{\operatorname{conv}}\operatorname{ext}(A)$.

More generally, a convex subset B of A is called a *face* (of A) if the following condition holds:

$$\big(x, y \in A,\ \frac{x+y}{2} \in B\big) \implies x, y \in B.$$

7.19 Example.
All the properties listed below are left as easy exercises.

(A) $\operatorname{ext}([0,1]) = \{0, 1\}$.

(B) For any non-zero Hilbert space H, $\operatorname{ext}(\bar{B}_H) = \partial\bar{B}_H$.

(C) $\operatorname{ext}(\bar{B}_{c_0}) = \varnothing$.

(D) If U is an open convex set in a non-zero topological vector space, then $\operatorname{ext}(U) = \varnothing$.

7.20 Proposition.
Let A be a convex set in a vector space E.

(a) *A point $b \in A$ is extreme in A iff $\{b\}$ is a face of A.*

(b) *If B and C are convex subsets of A such that $C \subset B$ and C is a face of B and B is a face of A, then C is a face of A as well.*

(c) *If* $\phi\colon A \to F$ *is an affine function (where* F *is an arbitrary vector space) and* $b \in \operatorname{ext}(\phi(A))$*, then* $\phi^{-1}(\{b\})$ *is a face of* A*.*

(d) *The intersection of an arbitrary non-empty collection of faces of* A *is a face of* A *as well.*

(*proof—exercise*)

The following lemma is a special case of a more general theorem that we will establish next (see Theorem 7.22 below). However, this lemma is a key part of the proof of the latter result.

7.21 Lemma.
If K *is a compact convex non-empty set in a locally convex* T_2*-space* E*, then* $\operatorname{ext}(K) \neq \varnothing$*.*

Proof. In this proof we treat E as a real vector space, even if it is complex.

It follows from Zorn's lemma and from item (d) of Proposition 7.20 that among all non-empty closed faces of K there exists a minimal set (w.r.t. the inclusion), say L. It is sufficient to show that L consists of a single point (why?). To this end, assume (on the contrary) that there are distinct points a and b that belong to L. We infer from Corollary 7.17 that there is a continuous linear functional $\phi\colon E \to \mathbb{R}$ such that $\phi(a) \neq \phi(b)$. Then $I \stackrel{\text{def}}{=} \phi(L)$ is a non-degenerate compact interval in $\mathbb{R}$, say $I = [p, q]$ with $p < q$. It follows from Proposition 7.20 that $L \cap \phi^{-1}(\{q\})$ is a (non-empty closed) face of K, and it is a proper subset of L, which contradicts minimality of L. □

7.22 Theorem. (Kreĭn-Milman Theorem)
Let K *be a compact convex set in a locally convex* T_2*-space* E*. Then*

$$K = \overline{\operatorname{conv}} \operatorname{ext}(K).$$

Proof. As in the previous proof, we treat E as a real vector space, even if it is complex.

Of course, $L \stackrel{\text{def}}{=} \overline{\operatorname{conv}} \operatorname{ext}(K)$ is contained in K. If these two sets differ, take an arbitrary point $a \in K$ that is not in L and apply Theorem 7.16 to obtain a continuous linear functional $\phi\colon E \to \mathbb{R}$ such that $\phi(a) < \min \phi(L)$. Then $I \stackrel{\text{def}}{=} \phi(K)$ is a compact non-degenerate interval in $\mathbb{R}$, say $I = [p, q]$ with $p < q$. Note that $p \notin \phi(L)$ and therefore $S \stackrel{\text{def}}{=} K \cap \phi^{-1}(\{p\})$ is disjoint from L. Moreover, S is a non-empty closed face of K. So, it follows from Lemma 7.21 that S has an extreme point, say c. Then $c \in \operatorname{ext}(K) \setminus L$ (thanks to Proposition 7.20), which contradicts the definition of L and finishes the whole proof. □

Later, in Theorem 10.13 (p. 51), we will prove a result that is, in a sense, the converse of the above theorem.

7.23 ***Remark.***
The assertions of both Lemma 7.21 and Theorem 7.22 are valid in a slightly more general setting: instead of requiring that the space E is locally convex, a sufficient assumption is that all the continuous linear functionals on E separate points of K. Both the proofs presented above work perfectly under this weaker assumption.

'Generalisation' of Theorem 7.22 (to the non-locally convex context) discussed above is superficial as any compact convex set K with the property that all continuous real-valued affine functions on K separate points of K is actually *isomorphic* (in the category of topological convex spaces with continuous affine functions as morphisms) to a compact convex set in a certain locally convex T_2-space. (A proof of this observation is not too difficult as is left to interested readers.)

In the next chapter, in Example 8.20 (p. 41), we will give an example of Kreĭn-Milman Theorem application.

8 Weak and weak* topologies

We begin this part with an abstract and quite general context that later on will be applied to normed vector spaces.

8.1 Definition.
We say two vector spaces X and Y (over the same field) form a *dual pair* w.r.t. a function $B\colon X \times Y \to \mathbb{K}$ if all the following conditions are fulfilled:

- B is bilinear;
- for any non-zero $x \in X$ there exists $y \in Y$ with $B(x, y) \neq 0$;
- for any non-zero $y \in Y$ there exists $x \in X$ with $B(x, y) \neq 0$.

Whenever X and Y form a dual pair (w.r.t. B), we define (in a similar way) two topologies—one on X and the other on Y:

- $\sigma(X, Y)$ = the topology on X induced by a separating collection of semi-norms $\{p_y\}_{y\in Y}$ where $p_y\colon X \ni x \mapsto |B(x, y)| \in \mathbb{R}_+$;
- $\sigma(Y, X)$ = the topology on Y induced by a separating collection of semi-norms $\{q_x\}_{x\in X}$ where $q_x\colon Y \ni y \mapsto |B(x, y)| \in \mathbb{R}_+$.

In this way we obtain two locally convex T_2-spaces: $(X, \sigma(X, Y))$ and $(Y, \sigma(Y, X))$.
In practice (with dual pairs), instead of '$B(x, y)$' one usually writes '$\langle x, y\rangle$.'

8.2 Definition.
Let E be a locally convex T_2-space. The *dual space of* E (or briefly, the *dual* of E), to be denoted E^*, is the vector space of all continuous linear functionals on E.

8.3 Example.
Let E be a locally convex T_2-space. Then E and E^* form a dual pair w.r.t. the function:

$$E \times E^* \ni (x, \phi) \mapsto \phi(x) \in \mathbb{K},$$

which simply follows from Corollary 7.17.

The name 'dual pair' is explained by the following

8.4 Proposition.
Let X and Y form a dual pair w.r.t. $\langle\cdot, -\rangle\colon X \times Y \to \mathbb{K}$ and be equipped with their topologies of the dual pair. Then the assignments

$$X \ni x \mapsto \langle x, \cdot\rangle \in Y^*$$
$$Y \ni y \mapsto \langle \cdot, y\rangle \in X^*$$

correctly define linear isomorphisms between respective vector spaces.

Proof. It is sufficient to show the assertion for the first assignment. To this end, for each $x \in X$ set $\mathfrak{e}_x\colon Y \ni y \mapsto \langle x, y\rangle \in \mathbb{K}$ and note that $\mathfrak{e}_x$ is a linear functional on Y. Continuing notation introduced in Definition 8.1, observe that $|\mathfrak{e}_x(y)| = q_x(y)$ and hence $\mathfrak{e}_x$ is continuous (thanks to Proposition 7.9). We infer that $\Phi\colon X \ni x \mapsto \mathfrak{e}_x \in Y^*$ is a well defined linear operator. Since the collection $\{q_x\}_{x\in X}$ is separating, it follows that Φ is one-to-one. To show that Φ is surjective, take any $\phi \in Y^*$. Another application of Proposition 7.9 shows that for some constant $M > 0$ and a finite number of vectors $x_1, \dots, x_N \in X$ we have:

$$|\phi(y)| \leq M \sum_{k=1}^{N} |\mathfrak{e}_{x_k}(y)| \qquad (y \in Y).$$

In particular, the kernel $\mathcal{N}(\phi)$ of ϕ contains $\bigcap_{k=1}^{N} \mathcal{N}(\mathfrak{e}_{x_k})$. So, it follows from a basic result from linear algebra that then ϕ is a linear combination of $\mathfrak{e}_{x_1}, \dots, \mathfrak{e}_{x_N}$; that is, $\phi = \sum_{k=1}^{N} \alpha_k \mathfrak{e}_{x_k} = \Phi(\sum_{k=1}^{N} \alpha_k x_k)$, and we are done. ∎

8.5 Definition.
Let E be a locally convex T_2-space. Then E and E^* form a dual pair in a canonical way described in Example 8.3. *Weak* topology on E is the topology $\sigma(E, E^*)$. Similarly, *weak** topology on E^* is defined as $\sigma(E^*, E)$.

It is worth underlying that the weak topology of E is weaker than the given one.

As a special case of Proposition 8.4, we obtain the following important

8.6 Theorem. (Weak and weak* continuous linear functionals)
Let E be a locally convex T_2-space.

(a) *A linear functional $\phi\colon E \to \mathbb{K}$ is continuous in the weak topology of E iff it is so in the given topology of E.*

(b) *Weak* continuous linear functionals on E^* are precisely the evaluation functionals (that is, linear functionals of the form $\phi \mapsto \phi(x)$ where $x \in X$ is arbitrarily fixed).*

(*proof—exercise*)

Very often weak topologies differ from the given ones. However, the collections of closed convex sets coincide in both these topologies, as shown by

8.7 Theorem. (Weakly closed convex sets)
A convex set in a locally convex T_2-space E is closed iff it is weakly closed.

Proof. For clarity, denote by τ and ω, respectively, the given and the weak topologies on E. Since $\mathrm{id}\colon (E,\tau) \to (E,\omega)$ is continuous, it follows that an arbitrary weakly closed set is closed w.r.t. τ. To see the converse for convex sets, consider a non-empty closed convex set W w.r.t. τ and fix any vector $v \in E \setminus W$. It follows from Theorem 7.15 that there exists $\phi \in E^*$ such that $\operatorname{Re}\phi(v) < \inf \operatorname{Re}\phi(W)$. In particular, $\phi(v) \notin \mathrm{cl}(\phi(W))$. But ϕ is weakly continuous (thanks to Theorem 8.6) and therefore the image of the weak closure of W under ϕ is contained in $\mathrm{cl}(\phi(W))$. So, we infer that v does not lie in the weak closure of W, which shows that W is weakly closed. □

As an immediate consequence of Corollary 7.10, we obtain an important (as well as basic) property of locally convex spaces:

8.8 Proposition.
A linear subspace F of a locally convex T_2-space E is a locally convex T_2-space as well. Moreover, the weak topology of F coincides with the topology induced from the weak topology of E.

(*proof—exercise*)

8.9 Definition.
Let A be an arbitrary subset of a locally convex T_2-space E. The *polar* of A is the set $A^\circ \subset E^*$ consisting of all $\phi \in E^*$ such that $\phi(A) \subset \bar{B}_{\mathbb{K}}$; that is, $\phi \in E^*$ belongs to the polar A° of A iff

$$\forall a \in A\colon\ |\phi(a)| \leq 1.$$

Similarly, the *prepolar* of a set $B \subset E^*$ is the set ${}^\circ B \in E$ consisting of all $x \in E$ such that

$$\forall \phi \in B\colon\ |\phi(x)| \leq 1.$$

Finally, the *bipolar* of A ($\subset E$), denoted by $A^{\circ\circ}$, is the set ${}^\circ(A^\circ)$.

Recall that the *absolutely convex hull* of A is the set

$$\operatorname{abs\,conv}(A) \stackrel{\text{def}}{=} \Big\{\sum_{k=1}^{n} t_k a_k\colon\ n \geq 1,\ t_1,\dots,t_n \in \mathbb{K},\ a_1,\dots,a_n \in A,\ \sum_{k=1}^{n}|t_k| \leq 1\Big\}$$

(it is the smallest set among all absolutely convex supersets of A in E). We will denote by $\overline{\operatorname{abs\,conv}}(A)$ the closure of the above set.

8.10 Theorem. (Bipolar Theorem)
Let E be a locally convex T_2-space. Then for any non-empty set $A \subset E$,

$$A^{\circ\circ} = \overline{\operatorname{abs\,conv}}(A).$$

Proof. It is straightforward that:

- the prepolar of any subset of E^* is always absolutely convex and closed; and
- $A \subset A^{\circ\circ}$.

So, we infer that $W \stackrel{\text{def}}{=} \overline{\operatorname{abs\,conv}}(A)$ is contained in the bipolar of A. To convince oneself that actually these two sets coincide, take any $a \notin W$. It follows from Theorem 7.16 that there exists $\phi \in E^*$ such that $\sup \operatorname{Re}\phi(W) < \operatorname{Re}\phi(a)$. Since W is absolutely closed, we get that $M \stackrel{\text{def}}{=} \sup \operatorname{Re}\phi(W)$ coincides with $\sup\{|\phi(w)|\colon\ w \in W\}$. In particular, $M \geq 0$. Let r be any real number such that $M < r < \operatorname{Re}\phi(a)$. Then $r > 0$ and $\frac{\phi}{r} \in A^\circ$. However, $\frac{|\phi(a)|}{r} > 1$ and hence a does not belong to the prepolar of A°. In other words, $a \notin A^{\circ\circ}$, which finishes the proof. □

Now we will study in a more detail weak and weak* topologies in the realm of normed vector spaces.

8.11 Proposition.
For any normed vector space E the operator $\kappa_E\colon E \to E^{**}$ is a topological embedding when E and E^{**} are equipped with, respectively, the weak and the weak topologies.*

(*proof—exercise*)

Sequences that are convergent in the weak or weak* topology are of great importance. Basic properties of them are established below.

8.12 Theorem.

(A) *A weakly convergent sequence in a normed vector space is bounded.*

(B) *A weak* convergent sequence in a* ***Banach space*** *is bounded.*

Proof. Both the items are immediate consequences of the Uniform Boundedness Principle (Theorem 5.1, p. 20; see also Corollary 5.16). Indeed, if E is a normed vector space (resp. a Banach space) and $(x_n)_{n=1}^\infty \subset E$ weakly converges to $x_0 \in E$ (resp. $(\phi_n)_{n=1}^\infty$ weak* converges to $\phi_0 \in E^*$), we consider bounded linear functionals $\omega_n\colon E^* \to \mathbb{K}$ given by $\omega_n = \kappa_E(x_n)$ where $n \geq 0$ (resp. $\omega_n\colon E \to \mathbb{K}$ given by $\omega_n(x) = \phi_n(x),\ n \geq 0$). It follows from our assumptions that $\omega_1, \omega_2, \dots$ converge pointwise to ω_0. So, Theorem 5.1 implies that $\sup_{n>0} \|\omega_n\| < \infty$. But $\|\omega_n\| = \|x_n\|$ (resp. $\|\omega_n\| = \|\phi_n\|$), and we are done. □

8.13 ***Remark.***
We leave it as an interesting exercise to give an example of a norm unbounded weak* convergent sequence in the dual Banach space of a certain incomplete normed vector space.

8.14 Example.
As $C(K)$-spaces are universal for normed vector spaces (which means that each normed vector space is linearly isometric to a linear subspace of a certain space of the form $C(K)$ for compact K), it is of great importance to know

which sequences are weakly convergent in such spaces (cf. Proposition 8.8). Below we give a full characterisation of them in a slightly more general context.

Let X be a non-empty locally compact Hausdorff space. Functions $f_1, f_2, \ldots \in C_0(X)$ converge weakly to a function $g \in C_0(X)$ iff both the following conditions are satisfied:

- $\lim_{n\to\infty} f_n(x) = g(x)$ for all $x \in X$; and
- $\sup_{n>0} \|f_n\|_\infty < \infty$.

To see the necessity of these conditions, consider evaluation functionals $\mathfrak{e}_x \colon C_0(X) \ni f \mapsto f(x) \in \mathbb{K}$ to infer the former, and use Theorem 8.12 to get the latter. Conversely, if both these conditions are satisfied, then $f_1, f_2, \ldots$ converge weakly to g, thanks to the Riesz representation theorem: each bounded linear functional $\phi\colon C_0(X) \to \mathbb{K}$ has an integral form $\phi(f) = \int_X f(x)\tau(x)\,\mathrm{d}\mu(x)$ for some probabilistic Borel measure μ on X and a bounded Borel function $\tau\colon X \to \mathbb{K}$. Now it follows from the Lebesgue's dominated convergence theorem that $\lim_{n\to\infty} \int_X f_n(x)\tau(x)\,\mathrm{d}\mu(x) = \int_X g(x)\tau(x)\,\mathrm{d}\mu(x)$ (here we need the boundedness of the sequence $(f_n)_{n=1}^\infty$ [!]). In other words, $\lim_{n\to\infty} \phi(f_n) = \phi(g)$, and we are done.

As an immediate consequence of Theorem 8.7, we obtain

8.15 Theorem. (Mazur)
Let a sequence $(x_n)_{n=1}^\infty$ of vectors of a normed vector space E be weakly convergent to $a \in E$. Then for each $n > 0$ there exist non-negative real numbers $t_1^{(n)}, \ldots, t_n^{(n)}$ such that $\sum_{k=1}^n t_k^{(n)} = 1$ and

$$\lim_{n\to\infty} \Big\|a - \sum_{k=1}^n t_k^{(n)} x_n\Big\| = 0.$$

Proof. We may and do assume that $a = 0$. For each $n > 0$ let W_n be the set of all vectors of the form $\sum_{k=1}^n t_k x_k$ where $t_1, \ldots, t_n$ are non-negative reals that sum up to 1. Set

$$c_n \stackrel{\text{def}}{=} \inf\{\|w\|\colon\ w \in W_n\}.$$

Then $W_n \subset W_{n+1}$ and $c_n \geq c_{n+1} \geq 0$. Moreover, the norm closure W of $\bigcup_{n=1}^\infty W_n$ is a closed convex set in E and hence it is also weakly closed, by Theorem 8.7. In particular, $0 \in W$. We infer that $\lim_{n\to\infty} c_n = 0$ and the conclusion follows. □

8.16 Proposition.
A linear subspace V of E^* (where E is a normed vector space) is weak closed iff $V = ({}^\perp V)^\perp$.*

More generally, for any set $A \subset E^*$, $({}^\perp A)^\perp$ coincides with the weak closure of* $\operatorname{lin}(A)$.

Proof. It is sufficient to show only the second claim. It follows from its definition that the annihilator is always a weak* closed linear subspace. Hence the weak* closure W of $\operatorname{lin}(A)$ is contained in $({}^\perp A)^\perp$. On the other hand, for any $\phi \notin W$ there is a weak* continuous linear functional $\xi\colon E^* \to \mathbb{K}$ such that $\sup \operatorname{Re}\xi(W) < \operatorname{Re}\xi(\phi)$ (by Theorem 7.16). Since W is a linear subspace, we infer that $\xi(W) = \{0\}$. Further, it follows from Theorem 8.6 that ξ is of the form $\xi(\alpha) = \alpha(b)$ $(\alpha \in E^*)$ for some $b \in E$. Then $b \in {}^\perp A$ and, consequently, $\phi \notin ({}^\perp A)^\perp$ (as $\phi(b) \neq 0$), which finishes the proof. □

Another central tool of functional analysis is stated below.

8.17 Theorem. (Banach-Alaoglu Theorem)
For any normed vector space E the closed unit ball of E^* is compact in the weak topology.*

Proof. This result is an immediate consequence of the Tychonoff's theorem (on the compactness of the product of compact spaces). Indeed, let E serve as a set of indices. For any non-zero $p \in E$ set $K_p \stackrel{\text{def}}{=} \|p\|\bar{B}_{\mathbb{K}}$ and let $K_0 \stackrel{\text{def}}{=} \{0\}$.

Observe that the function

$$\Phi\colon E^* \ni \phi \mapsto (\phi(p))_{p\in E} \in \prod_{p\in E} \mathbb{K} = \mathbb{K}^E$$

is a topological embedding when E^* is equipped with the weak* topology and $\mathbb{K}^E$ with the product one. So, $\bar{B}_{E^*}$ is weak* compact iff its image via Φ is a compact set in $\mathbb{K}^E$. Since $\Phi(\bar{B}_{E^*}) \subset \prod_{p\in E} K_p$ ($\subset \mathbb{K}^E$) and $T \stackrel{\text{def}}{=} \prod_{p\in E} K_p$ is compact (by the Tychonoff's theorem), we only need to check that $\Phi(\bar{B}_{E^*})$ is closed in $\mathbb{K}^E$, which is immediate because $\Phi(\bar{B}_{E^*})$ consists precisely of all $(t_p)_{p\in E} \in \mathbb{K}^E$ that satisfy all the following conditions (for all $p, q \in E$ and $s \in \mathbb{K}$):

- $t_p \in K_p$;
- $t_{p+q} = t_p + t_q$;
- $t_{sp} = st_p$.

□

8.18 Theorem.
Let E be a normed vector space and Ω stand for the ball $\bar{B}_{E^*}$ equipped with the weak topology. The following conditions are equivalent:*

(i) *Ω is metrisable;*

(ii) *Ω is first countable;*

(iii) *E is separable.*

Proof. First assume that E is separable. Let $D = \{d_1, d_2, \ldots\}$ be a dense set in B_E. Then the formula

$$\|\phi\|_D \stackrel{\text{def}}{=} \sum_{n=1}^{\infty} \frac{1}{2^n} |\phi(d_n)| \qquad (\phi \in E^*)$$

correctly defines a (new) norm on E^* (which is, in general, non-equivalent to the original norm of E^*). To show (i), it is sufficient to check that this norm induces a metric that is compatible with the weak* topology (only) on Ω. To this end, fix $(\phi_\sigma)_{\sigma\in\Sigma} \subset \Omega$ and $\psi \in \Omega$. Observe that if $(\phi_\sigma)_{\sigma\in\Sigma}$ converges to ψ (in Ω), then

$$\lim_{\sigma\in\Sigma} \|\phi_\sigma - \psi\|_D = 0 \tag{8:1}$$

(as $\|\phi_\sigma - \psi\|_D \leq \sum_{k=1}^n |\phi_\sigma(d_k) - \psi(d_k)| + 2^{1-n}$). Conversely, if (8:1) holds, then $\lim_{\sigma\in\Sigma} \phi_\sigma(d_k) = \psi(d_k)$ for any $k > 0$. Now fix arbitrary $a \in B_E$ and $\varepsilon > 0$. Then there is $k > 0$ such that $\|a - d_k\| \leq \varepsilon/3$. Further, there exists $\sigma_0 \in \Sigma$ such that $|\phi_\sigma(d_k) - \psi(d_k)| \leq \varepsilon/3$ whenever $\sigma \geq \sigma_0$. Then, for all such σ,

$$|\phi_\sigma(a) - \psi(a)| \leq |\phi_\sigma(a - d_k)| + |\phi_\sigma(d_k) - \psi(d_k)| + |\psi(d_k - a)| \leq 2\|a - d_k\| + |\phi_\sigma(d_k) - \psi(d_k)| \leq \varepsilon,$$

which shows that $\lim_{\sigma\in\Sigma} \phi_\sigma(a) = \psi(a)$. Since a runs over all points of B_E, we simply conclude that $\phi_\sigma \stackrel{\Omega}{\to} \psi$, and we are done.

Now assume that Ω is first countable. This implies that there exist a sequence $(F_n)_{n=1}^\infty$ of finite subsets of E and a sequence $(\varepsilon_n)_{n=1}^\infty$ of positive real numbers such that the sets

$$U_n \stackrel{\text{def}}{=} \{\phi \in \Omega |\ \forall x \in F_n\colon\ |\phi(x)| < \varepsilon_n\} \qquad (n > 0)$$

form a basis of neighourhoods of the origin in Ω. As $V \stackrel{\text{def}}{=} \overline{\operatorname{lin}}(\bigcup_{n=1}^\infty F_n)$ ($\subset E$) is separable, it is enough to show that $V = E$. To this end, take any $\psi \in V^\perp$. If ψ was non-zero, take any $b \in E$ with $\psi(b) = \|\psi\|$ and observe that $V \stackrel{\text{def}}{=} \{\phi \in \Omega\colon\ |\phi(b)| < 1\}$ is a neighourhood of the origin in Ω that excludes $\frac{\psi}{\|\psi\|}$ and, consequently, does not contain any of the sets U_n (as $\frac{\psi}{\|\psi\|} \in \bigcap_{n=1}^\infty U_n$). The above argument shows that $V^\perp = \{0\}$ and hence $V = E$, by Proposition 5.25 (p. 25). As (ii) trivially follows from (i), the whole proof is finished. □

As a consequence of the last two results, we obtain

8.19 Theorem. (Banach-Mazur Theorem)
Every separable normed vector space is linearly isometric to a linear subspace of $C([0,1],\mathbb{K})$.

Proof. Let E be a separable normed vector space. Denote by K the ball $\bar{B}_{E^*}$ equipped with the weak* topology. We have already known that $K \neq \varnothing$ is compact and metrisable. Moreover, being a convex subset of a locally convex space, it is both connected and locally connected (exercise). So, it follows from the Hahn-Mazurkiewicz theorem that there exists a continuous surjection $u\colon [0,1] \to K$. We will involve the map u in the last step of the proof.

For any $x \in E$ define $\hat{x}\colon K \to \mathbb{K}$ by the rule $\hat{x}(\phi) = \phi(x)$. Observe that $\hat{x} \in C(K,\mathbb{K})$ and that

$$T\colon E \ni x \mapsto \hat{x} \in C(K,\mathbb{K})$$

is a linear operator. Moreover, T is isometric, thanks to Corollary 5.13 (p. 23). So, to finish the proof, it is sufficient to notice that the function $S\colon C(K,\mathbb{K}) \ni f \mapsto f \circ u \in C([0,1],\mathbb{K})$ is a linear isometry (because u is surjective) and to consider the composition $S \circ T\colon E \to C([0,1],\mathbb{K})$ (that is also a linear isometry). □

We are now ready to give an important example related to the Kreĭn-Milman Theorem (Theorem 7.22).

8.20 Example.
Let X be a compact Hausdorff space. Denote by $\mathscr{B}(X)$ and $\mathcal{M}_r(X)$, respectively, the σ-algebra of all Borel subsets of X and the real vector space of all signed (that is, real-valued) regular Borel measures on X. For any $\mu \in \mathcal{M}_r(X)$ let $\phi_\mu\colon C(X,\mathbb{R}) \to \mathbb{R}$ be given by $\phi_\mu(f) \stackrel{\text{def}}{=} \int_X f(x)\,\mathrm{d}\mu(x)$. According to the Riesz representation theorem, the assignment

$$\mathcal{M}_r(X) \ni \mu \mapsto \phi_\mu \in C(X,\mathbb{R})^*$$

correctly defines a bijective linear operator Φ. The image K of the set $\mathrm{Prob}_r(X)$ of all regular probabilistic Borel measures on X under Φ consists precisely of all linear functionals $L\colon C(X,\mathbb{R}) \to \mathbb{R}$ that are non-negative (that is, such that $L(f) \geq 0$ for all non-negative $f \in C(X,\mathbb{R})$) and send the function constantly equal to 1 to the scalar 1. As such, K is weak* compact (and, of course, convex). One transports the weak* topology of K to $\mathrm{Prob}_r(X)$ (via Φ) and calls this topology on $\mathrm{Prob}_r(X)$ *weak*. (Such a terminology is present e.g. in dynamical systems). In this way $\mathrm{Prob}_r(X)$ becomes a compact convex set w.r.t. its weak topology. Convergence in this space is characterised as follows: a net $(\mu_\sigma)_{\sigma\in\Sigma} \subset \mathrm{Prob}_r(X)$ tends (weakly) to $\lambda \in \mathrm{Prob}_r(X)$ iff

$$\lim_{\sigma\in\Sigma} \int_X f\,\mathrm{d}\mu_\sigma = \int_X f\,\mathrm{d}\lambda \qquad (f \in C(X,\mathbb{R})). \tag{8:2}$$

When X is metrisable, one proves that all probabilistic Borel measures on X are regular and thus one writes simply $\mathrm{Prob}(X)$ instead of $\mathrm{Prob}_r(X)$ in that case. $\mathrm{Prob}(X)$ (for metrisable X) is metrisable in the weak topology.

It is not difficult to show that $\mathrm{ext}(\mathrm{Prob}_r(X))$ consists precisely of all measures δ_a $(a \in X)$ where

$$\delta_a(A) = \begin{cases} 1 & a \in A \\ 0 & a \notin A \end{cases}$$

(exercise). In particular, $\mathrm{ext}(\mathrm{Prob}_r(X))$ is compact and *naturally* homeomorphic to X. It follows from the Kreĭn-Milman theorem that all probabilistic measures that are supported on finite sets form a (weakly) dense subset of $\mathrm{Prob}_r(X)$.

8.21 Theorem. (Goldstine's theorem)
For every normed vector space E, the set $\kappa_E(\bar{B}_E)$ is weak dense in $\bar{B}_{E^{**}}$.*

Proof. For simplicity, set $K \stackrel{\text{def}}{=} \kappa_E(\bar{B}_E)$ and consider $X \stackrel{\text{def}}{=} E^{**}$ with the weak* topology. Since K is absolutely convex and X is locally convex, we infer from Theorem 8.10 that the weak* closure of K coincides with its bipolar. However, it follows from item (b) of Theorem 8.6 that X^* may naturally be identified with E. Under this identification, K° coincides with $\bar{B}_E$ (thanks to Corollary 5.13, p. 23). Consequently, $K^{\circ\circ} = B_{E^{**}}$ and we are done. □

As a consequence, we obtain the following result, sometimes attributed to Banach.

8.22 Theorem.
A normed vector space is a reflexive Banach space iff its closed unit ball is weakly compact.

Proof. If E is a reflexive Banach space, then $\bar{B}_E$ in the weak topology is homeomorphic to $\bar{B}_{E^{**}}$ in the weak* topology via κ_E (by Proposition 8.11). So, the weak compactness of $\bar{B}_E$ follows from Theorem 8.17. Conversely, if $\bar{B}_E$ is weakly compact, then $\kappa_E(\bar{B}_E) = \bar{B}_{E^{**}}$, thanks to Theorem 8.21 and (again) Proposition 8.11. Consequently, κ_E is surjective and we are done. □

There is also another characterisation of reflexivity (for Banach spaces), but its proof is much more difficult (therefore we skip it):

8.23 Theorem. (James' theorem)
A normed vector space is a reflexive Banach space iff any its bounded linear functional attains its norm. More precisely, a normed vector space E is a reflexive Banach space iff for any $\phi \in E^*$ there exists $x \in \bar{B}_E$ such that $\phi(x) = \|\phi\|$.

(*without proof*)

Basic consequences of Theorem 8.22 are listed in the following

8.24 Corollary.

(A) *A closed linear subspace of a reflexive Banach space is reflexive as well.*

(B) *If $T\colon E \to F$ is a bounded surjective linear operator between two Banach spaces and E is reflexive, then so is F. In particular, a Banach space that is isomorphic (as a Banach space) to a reflexive Banach space is reflexive as well.*

Proof. If V is a closed linear subspace of a reflexive Banach space E, then V is weakly closed in E (by Theorem 8.7). Consequently, $(\bar{B}_V =)\ \bar{B}_E \cap V$ is weakly compact (as a weakly closed subset of a weakly compact set—see Theorem 8.22) and the conclusion follows from Theorem 8.22.

Now if $T\colon E \to F$ is as specified in (B), then T is open (by Theorem 5.4, p. 21). So, there is $r > 0$ such that $\bar{B}_F \subset rT(\bar{B}_E)$. Note that T is continuous in the weak topologies of E and F (why?); and $\bar{B}_F$ is weakly closed in F (again by Theorem 8.7). Therefore $T(\bar{B}_E)$ is weakly compact and so is $\bar{B}_F$. Consequently, F is reflexive (again thanks to Theorem 8.22). □

8.25 Definition.
Let $T\colon E \to F$ be a bounded linear operator between normed vector spaces. The *adjoint operator* of T, denoted $T^*\colon F^* \to E^*$, is defined by $T^*(\phi) = \phi \circ T$ $(\phi \in F^*)$.

8.26 Proposition.
For any bounded linear operator $T\colon E \to F$ between normed vector spaces, the operator T^* is bounded as well and $\|T^*\| = \|T\|$. Moreover,

$$T^{**} \circ \kappa_E = \kappa_F \circ T. \tag{8:3}$$

Proof. Observe that $\|T^*(\phi)\| = \|\phi \circ T\| \leq \|\phi\| \cdot \|T\|$, which implies that $\|T^*\| \leq \|T\|$. In particular, T^{**} is well defined and $\|T^{**}\| \leq \|T^*\|$. Note that for any $\phi \in F^*$ and $x \in E$ we have:

$$\begin{aligned}(T^{**}(\kappa_E(x)))(\phi) &= (\kappa_E(x) \circ T^*)(\phi) = (\kappa_E(x))(T^*(\phi)) = (\kappa_E(x))(\phi \circ T) \\ &= (\phi \circ T)(x) = \phi(T(x)) = (\kappa_F(T(x)))(\phi),\end{aligned}$$

which implies that $T^{**}(\kappa_E(x)) = \kappa_F(T(x))$ or, equivalently, $T^{**} \circ \kappa_E = \kappa_F \circ T$. In particular, for any $x \in E$, $\|T(x)\| = \|\kappa_F(T(x))\| = \|T^{**}(\kappa_E(x))\| \le \|T^{**}\| \cdot \|\kappa_E(x)\| \le \|T^*\| \cdot \|x\|$ and hence $\|T\| \le \|T^*\|$. □

Equation (8:3) says that, when identifying E with a subspace of E^{**} via κ_E and similarly for F, the operator T^{**} extends T. In particular, if E is reflexive, T and T^{**} 'coincide.'

8.27 Theorem.
Let E and F be two normed vector spaces.

(A) *For any $T \in \mathscr{L}(E, F)$ the operator $T^*\colon F^* \to E^*$ is continuous in the weak* topologies of E^* and F^*.*

(B) *For a linear operator $S\colon F^* \to E^*$ the following conditions are equivalent:*

(i) *S is continuous in the weak* topologies;*

(ii) *there exists a bounded linear operator $T\colon E \to F$ such that $S = T^*$.*

Proof. We start from (A). Let $(\phi_\sigma)_{\sigma\in\Sigma} \subset F^*$ converges to $\phi \in F^*$ in the weak* topology. This means that this net converges pointwise to ϕ. Since T is continuous, we infer that the functionals $T \circ \phi_\sigma$ converge pointwise to $T \circ \phi$. Equivalently, the functionals $T^*(\phi_\sigma)$ converge to $T^*(\phi)$ in the weak* topology, and we are done.

We now turn to (B). Thanks to (A), we only need to show that (ii) follows from (i). To this end, we assume S satisfies (i). For any $x \in E$ the functional $\mathfrak{e}_x\colon E^* \to \mathbb{K}$ ($\mathfrak{e}_x(\phi) = \phi(x)$) is weak* continuous and hence so is $\mathfrak{e}_x \circ S$. We conclude from Theorem 8.6 that there is a (unique) vector $v \in F$ such that $\mathfrak{e}_x \circ S = \mathfrak{e}_v$ ($\mathfrak{e}_v\colon F \ni \xi \mapsto \xi(v) \in \mathbb{K}$). The uniqueness of v allows us to set $T(x) \stackrel{\text{def}}{=} v$. In this way we obtain a function $T\colon E \to F$. The equation $\mathfrak{e}_x \circ S = \mathfrak{e}_{Tx}$ means that

$$S(\phi) = \phi \circ T \qquad (\phi \in F^*). \tag{8:4}$$

Since the functionals from F^* separate the points of F, we infer from (8:4) that T is linear. So, it remains to check that T is bounded (because then (8:4) shows that $S = T^*$). To this end, take any sequence $x_1, x_2, \ldots \in E$ convergent to 0. We claim that then $(T(x_n))_{n=1}^\infty$ converges weakly to 0 (in F). Indeed, for any $\phi \in F^*$ we have (thanks to (8:4)) $\phi(T(x_n)) = (S(\phi))(x_n) \to 0$ ($n \to \infty$). Now we apply Theorem 8.12 to conclude that the sequence $(T(x_n))_{n=1}^\infty$ is norm bounded. In this way we have shown that T transforms sequences convergent to 0 into bounded sequences. A linear operator (between normed vector spaces) with this property is automatically continuous (exercise!). □

9 Kreĭn-Šmulian and Eberlein[-Šmulian] theorems

The aim of this chapter is to prove the following two celebrated results of functional analysis:

9.1 Theorem. (Kreĭn-Šmulian Theorem, 1940)
Let E be a Banach space and $A \subset E^*$ a convex set. Then A is weak closed iff $A \cap n\bar{B}_{E^*}$ is weak* closed (or, equivalently, weak* compact) for each $n > 0$.*

It is worth underlying that the above result is false (in general) when E is only a normed vector space.

9.2 Theorem. (Eberlein[-Šmulian] Theorem, 1947)
A subset of a Banach space is weakly sequentially compact iff it is weakly compact.

More generally, if A is a subset of a Banach space E such that:

(wcpc) *for any infinite subset D of A there exists a point $d \in E$ such that $U \cap D$ is infinite for any weak neighbourhood U of d in E,*

then the weak closure K of A in E is both weakly compact and sequentially weakly compact, and has the following property:

- *for any $B \subset K$ and any point z from the weak closure of B there exists a sequence $b_1, b_2, \ldots \in B$ that converges weakly to z.*

Weak sequential compactness of weakly compact subsets of Banach spaces was discovered by Šmulian in 1940 (actually this is a rather easy observation, see Corollary 9.10 below). The converse (that is, weak compactness of weakly sequentially compact sets) was obtained by Eberlein in 1947. F. Albiac and N.J. Kalton in their book *Topics in Banach Space Theory* (Graduate Texts in Mathematics **233**, Second Edition, Springer, 2016) call this result *Eberlein-Šmulian theorem*. They write there: "*The Eberlein-Šmulian theorem was probably the deepest result of earlier (pre-1950) Banach space theory.*"

The proofs of the above theorems will be preceded by auxiliary lemmas. To simplify statements, we fix a Banach space E (from now to the end of this chapter) and call a set $A \subset E^*$ *clob** if $A \cap n\bar{B}_{E^*}$ is weak* closed for any $n > 0$.

9.3 Lemma.
Any clob set $A \subset E^*$ is closed in the norm topology, and $r(A+u)$ is clob* as well for any $r > 0$ and $u \in E^*$.*

Proof. The norm closedness of A easily follows from the boundedness of norm convergent sequences. Fix an integer $N > 0$ such that $\|u\| \leq N$ as well as $\frac{1}{r} \leq N$, set $K \stackrel{\text{def}}{=} \bar{B}_{E^*}$ and note that for any $n > 0$, $\frac{n}{r}K - u \subset (n+1)NK$. Since $\frac{n}{r}K - u$ is weak* closed (e.g. by the Banach-Alaoglu theorem), we infer that the set $(\frac{n}{r}K - u) \cap A$ $(= (\frac{n}{r}K - u) \cap ((n+1)NK \cap A))$ is weak* closed as well (as A is clob*). Thus, so is $nK \cap r(A+u) = r[(\frac{n}{r}K - u) \cap A] + ru$, and we are done. □

9.4 Lemma.
For any $r > 0$, the collection $\mathcal{F}_r \stackrel{\text{def}}{=} \{F^\circ \colon\ \varnothing \neq F \subset \frac{1}{r}\bar{B}_E \text{ finite}\}$ consists of weak closed sets, is downward directed and*

$$\bigcap_{Q \in \mathcal{F}_r} Q = r\bar{B}_{E^*}.$$

Proof. It is clear that F° is weak* closed and that $F^\circ \cap D^\circ = (F \cup D)^\circ$, which yields downward directedness. Denote by $\mathcal{S}$ the collection of all finite non-empty subsets of $\frac{1}{r}\bar{B}_E$. Then $\bigcap_{Q \in \mathcal{F}_r} Q = \bigcap_{F \in \mathcal{S}} F^\circ = (\bigcup_{F \in \mathcal{S}} F)^\circ = (\frac{1}{r}\bar{B}_E)^\circ = r(\bar{B}_E)^\circ = r\bar{B}_{E^*}$. □

9.5 Lemma.
If $A \subset E^*$ is a clob set disjoint from $\bar{B}_{E^*}$, then there is a vector $x \in E$ such that*

$$\operatorname{Re}\phi(x) \geq 1 \qquad (\phi \in A). \tag{9:1}$$

Proof. We may and do assume that $A \neq \varnothing$. First we will construct inductively a sequence of finite non-empty sets $F_0, F_1, \ldots \subset E$ such that for all $n \geq 0$:

(1_n) $F_n \subset 2^{1-n}\bar{B}_E$; and

(2_n) $A \cap 2^n\bar{B}_{E^*} \cap \bigcap_{k=0}^n F_k^\circ = \varnothing$.

We start from $F_0 \stackrel{\text{def}}{=} \{0\}$. Observe that (1_0)–(2_0) hold (as $A \cap \bar{B}_{E^*} = \varnothing$). Now assume that $n > 0$ and $F_0, \ldots, F_{n-1}$ have been defined. Set $K \stackrel{\text{def}}{=} A \cap 2^n\bar{B}_E \cap \bigcap_{k=0}^{n-1} F_k^\circ$. Since A is clob*, K is weak* compact (by the Banach-Alaoglu theorem). Denote by $\mathcal{S}$ the collection of all finite non-empty subsets of $2^{1-n}\bar{B}_E$. It follows from Lemma 9.4 that $K \cap \bigcap_{S \in \mathcal{S}} S^\circ = K \cap 2^{n-1}\bar{B}_{E^*} = \varnothing$, where the last equality follows from (2_{n-1}). So, we conclude from properties listed in Lemma 9.4 and weak* compactness of K that there is a set $F_n \in \mathcal{S}$ such that $K \cap F_n^\circ = \varnothing$. In this way (1_n)–(2_n) are satisfied.

Having the sets $F_0, F_1, \ldots$, we arrange all the elements of $\bigcup_{n=0}^\infty F_n$ in a sequence $(p_n)_{n=1}^\infty$ in a way such that

$$\lim_{n\to\infty} \|p_n\| = 0 \tag{9:2}$$

(this is possible by (1_n) and the finiteness of each F_n). It follows from (2_n) that

$$A \cap \{p_n \colon\ n > 0\}^\circ = \varnothing. \tag{9:3}$$

Further, (9:2) enables us to define correctly a (bounded) linear operator $P\colon E^* \ni \phi \mapsto (\phi(p_n))_{n=1}^\infty \in c_0$. Note that $P(A)$ is convex and disjoint from B_{c_0} (the latter property is implied by (9:3)). So, it follows from Theorem 7.15 (p. 33) that there exists a bounded linear non-zero functional $\xi\colon c_0 \to \mathbb{K}$ such that $\sup \operatorname{Re}\xi(B_{c_0}) \leq \inf \operatorname{Re}\xi(P(A))$. Without loss of generality, we may and do assume that $\|\xi\| = 1$. Then $\sup \operatorname{Re} \xi(B_{c_0}) = 1$ (why?) and $\operatorname{Re} \xi(P(\phi)) \geq 1$ for any $\phi \in A$. Finally, ξ is of the form $\xi((w_n)_{n=1}^\infty) = \sum_{n=1}^\infty a_n w_n$ where $(a_n)_{n=1}^\infty \subset \mathbb{K}$ and $\sum_{n=1}^\infty |a_n| = 1$. We define $x \stackrel{\text{def}}{=} \sum_{n=1}^\infty a_n p_n$ (the series converges in E—why?). Then, for any $\phi \in A$:

$$\operatorname{Re} \phi(x) = \operatorname{Re}\Big(\sum_{n=1}^\infty a_n \phi(p_n)\Big) = \operatorname{Re} \xi((\phi(p_n))_{n=1}^\infty) = \operatorname{Re} \xi(P(\phi)) \geq 1.$$

□

Now we are ready to give

Proof of Theorem 9.1. Fix $u \in E^* \setminus A$. It is suffient to show that u does not belong to the weak* closure of A. It follows from Lemma 9.3 that A is norm closed. So, there is $r > 0$ such that $A \cap (u + \frac{1}{r}\bar{B}_{E^*}) = \varnothing$. Then $\bar{B}_{E^*} \cap r(A - u) = \varnothing$ as well. Another application of Lemma 9.3 yields that $r(A - u)$ is clob*. Thus, we infer from Lemma 9.4 that there exists $x \in E$ such that $\operatorname{Re} \phi(x) \geq 1$ for all $\phi \in r(A - u)$. In other words, $r(A - u) \subset H \stackrel{\text{def}}{=} \{\phi \in E^*\colon\ \operatorname{Re} \phi(x) \geq 1\}$. Since H is weak* closed and does not contain the zero functional, we see that 0 is not in the weak* closure of $r(A-u)$. Equivalently, u is not in the weak* closure of A, and we are done. □

9.6 Corollary.
If E is a Banach space and $F \subset E^*$ is a convex cone (that is, if F is convex and $rF \subset F$ for all $r > 0$), then F is weak closed iff $F \cap \bar{B}_{E^*}$ is weak* closed.*

(proof—exercise)

In particular, the above result applies to linear subspaces F of E^*.

9.7 Corollary.
If E is a separable Banach space, then a convex set $A \subset E^*$ is weak closed iff it is sequentially weak* closed.*

Proof. If A is sequentially weak* closed, then $A \cap n\bar{B}_{E^*}$ is weak* closed (for any $n > 0$), since $n\bar{B}_{E^*}$ is metrisable in the weak* topology (as it is homeomorphic to $\bar{B}_{E^*}$ equipped with the weak* topology, and E is separable—see Theorem 8.18). So, weak* closedness of A follows from Theorem 9.1. □

The following consequence of the Kreĭn-Šmulian theorem is very useful.

9.8 Theorem.
Let E and F be Banach spaces. A linear operator $T\colon E^* \to F^*$ is continuous in the weak topologies of E^* and F^* iff so is $T \restriction \bar{B}_{E^*}$.*

Proof. Assume the restriction S of T to the closed unit ball is continuous in the weak* topologies. First consider any $v \in F$ and set $\xi = \kappa_F(v)$. We claim that $\xi \circ T$ is continuous in the weak* topology. Indeed, since $\xi \circ T$ is a linear functional, it follows from Corollary 2.15 (p. 5) that this function is weak* continuous iff its kernel N is weak* closed. But $N \cap \bar{B}_{E^*} = S^{-1}(\mathcal{N}(\xi))$ is weak* closed (since both S and ξ are weak* continuous). Now an application of Corollary 9.6 yields that N is weak* closed and, consequently, $\xi \circ T$ is weak* continuous.

Now take any net $(\phi_\sigma)_{\sigma\in\Sigma} \subset E^*$ convergent to 0 in the weak* topology. We only need to show that $(T(\phi_\sigma))_{\sigma\in\Sigma}$ converges to 0 in the weak* topology of F^*. This is equivalent to the statement that $\lim_{\sigma\in\Sigma}(T(\phi_\sigma))(v) = 0$ for any $v \in F$, which is valid thanks to the first part of the proof (as $(T(\phi_\sigma))(v) = (\kappa_F(v) \circ T)(\phi_\sigma)$). □

Now we pass to weak compactness. We begin with a simple

9.9 Lemma.
A weakly compact set in a separable Banach space is weakly metrisable.

Proof. Let K be a weakly compact subset of a separable Banach space X. For any non-zero vector $x \in X$ take a functional $\phi_x \in X^*$ with $\phi_x(x) \neq 0$. Since X is separable (and metrisable), the cover $\{\phi_x^{-1}(\mathbb{K} \setminus \{0\})\}_{x \in X \setminus \{0\}}$ of $X \setminus \{0\}$ has a countable subcover (of $X \setminus \{0\}$). This means that there is a sequence $\alpha_1, \alpha_2, \ldots \in X^*$ such that for any two distinct vectors x and y in X one can find an index $n > 0$ with $\alpha_n(x) \neq \alpha_n(y)$. In particular, the function

$$\Phi\colon K \ni x \mapsto (\alpha_n(x))_{n=1}^{\infty} \in \mathbb{K}^{\omega}$$

is one-to-one. It is also continuous when K is equipped with the weak topology and $\mathbb{K}^{\omega}$ with the product one. Hence Φ is a topological embedding (since K is compact in the weak topology) into a metrisable space, and we are done. □

9.10 Corollary. (Šmulian theorem)
A weakly compact set in a Banach space is sequentially weakly compact.

Proof. Let $x_1, x_2, \ldots$ be vectors from a weakly compact set K in E. Then the closure F of a linear span of this sequence is a separable Banach space as well as a weakly closed set in E (by Theorem 8.7). Since the weak topology of F coincides with the topology induced from the weak topology of E, we conclude that $K \cap F$ is a weakly compact set in F. So, Lemma 9.9 applies and $K \cap F$ is weakly metrisable. Being metrisable and compact (in the weak topology), $K \cap F$ is sequentially compact. Therefore our given sequence has a subsequence that is weakly convergent to a certain point from $K \cap F$, and we are done. □

9.11 Lemma.
For any finite-dimensional linear subspace V of E^{**} there is a finite set $I \subset \bar{B}_{E^*}$ such that for any $\xi \in V$:

$$\frac{\|\xi\|}{2} \leq \max\{|\xi(\phi)|\colon \ \phi \in I\}.$$

Proof. The unit sphere S of V (that is, the set of all unit vectors in V) is compact and therefore there exists a finite set $J \subset S$ that is a $\frac{1}{4}$-net in S; that is, for any $\beta \in S$ there is $\mu \in J$ such that

$$\|\mu - \beta\| \leq \frac{1}{4}. \tag{9:4}$$

Further, for any $\mu \in J$ there is a unit vector $\psi_\mu \in E^*$ such that $\mu(\psi_\mu)$ is a real number greater than $\frac{3}{4}$. We define I as the set of all ψ_μ where μ runs over all elements of J. Now take an arbitrary functional $\beta \in S$ and choose $\mu \in J$ such that (9:4) holds. Then for $\phi \stackrel{\text{def}}{=} \psi_\mu$ $(\in I)$ we get

$$|\beta(\phi)| \geq |\mu(\phi)| - |(\beta - \mu)(\phi)| \geq \frac{3}{4} - \frac{1}{4} = \frac{1}{2}\|\beta\|,$$

and the whole conclusion easily follows. □

9.12 Lemma.
Let A be a bounded subset of E that satisfies (wcpc) *and W denote the weak* closure of $\kappa_E(A)$ in E^{**}. Then W is weak* compact and for any $\xi \in W$ there is a sequence of vectors from A that converge weakly to some $c \in E$ such that $\kappa_E(c) = \xi$. In particular, the weak closure of A is weakly compact.*

Proof. Boundedness of A implies that W is weak* compact (by the Banach-Alaoglu theorem). We fix $\xi \in W$ and will show that there exists a sequence $(a_n)_{n=1}^{\infty} \subset A$ that converges weakly in E to certain $c \in E$ such that $\kappa_E(c) = \xi$.

Since ξ belongs to the weak* closure of $\kappa_E(A)$, it follows that:

($\star$) for any finite set $C \subset E^*$ and $\varepsilon > 0$ there is $a \in A$ such that $|\xi(\phi) - \phi(a)| \leq \varepsilon$ for all $\phi \in C$.

We will now construct inductively a sequence $a_1, a_2, \ldots$ of vectors from A and an increasing sequence of finite subsets $F_1 \subset F_2 \subset \ldots$ of $\bar{B}_{E^*}$ such that for any $n > 0$:

$$|\xi(\phi) - \phi(a_n)| \leq \frac{1}{n} \qquad (\phi \in F_n) \tag{9:5}$$

and

$$\|\eta\| \leq 2\max\{|\eta(\phi)|\colon\ \phi \in F_n\} \qquad (\eta \in \operatorname{lin}\{\xi - \kappa_E(a_j)\colon\ 0 \leq j < n\}) \tag{9:6}$$

where $a_0 \stackrel{\text{def}}{=} 0$. Setting $F_0 \stackrel{\text{def}}{=} \{0\}$, we assume that $a_0, \dots, a_{k-1}$ and $F_0 \subset F_1 \subset \dots \subset F_{k-1} \subset \bar{B}_{E^*}$ have been already defined. Applying Lemma 9.11 to $V = \operatorname{lin}\{\xi - \kappa_E(a_j)\colon\ 0 \leq j < k\}$, we obtain a finite set $I \subset \bar{B}_{E^*}$ such that (9:6) holds for $n = k$ and $F_k \stackrel{\text{def}}{=} F_{k-1} \cup I$. Now $(\star)$ with $C = F_k$ and $\varepsilon = \frac{1}{k}$ gives us a point $a_k \in A$ such that (9:5) holds for $n = k$.

In this way both the sequences $(a_n)_{n=1}^\infty$ and $(F_n)_{n=1}^\infty$ have been constructed. Now (wcpc) of A yields a point $c \in E$ such that any weak neighbourhood of c in E contains infinitely many entries of $(a_n)_{n=1}^\infty$ (such c exists even if the set $D \stackrel{\text{def}}{=} \{a_n\colon\ n > 0\}$ is finite). In particular, c belongs to the weak closure of D. Since the norm closure $\bar{Z}$ of $Z \stackrel{\text{def}}{=} \operatorname{lin}(D)$ is weakly closed (by Theorem 8.7), we conclude that $c \in \bar{Z}$.

Fix for a moment $\phi \in M \stackrel{\text{def}}{=} \bigcup_{n=1}^\infty F_n$, say $\phi \in F_m$, and $\varepsilon > 0$. Since $\{x \in E\colon\ |\phi(x) - \phi(c)| < \varepsilon\}$ is a weak neighbourhood of c, we conclude that there are infinitely many indices $n > m$ such that $|\phi(a_n) - \phi(c)| \leq \varepsilon$. But then, for any such n (since $F_m \subset F_n$):

$$|\xi(\phi) - \phi(c)| \leq |\xi(\phi) - \phi(a_n)| + |\phi(a_n) - \phi(c)| \leq \frac{1}{n} + \varepsilon,$$

by (9:5). Letting $n \to \infty$, we get $|\xi(\phi) - \phi(c)| \leq \varepsilon$ and, consequently:

$$\xi(\phi) = \phi(c) \qquad (\phi \in M). \tag{9:7}$$

On the other hand, (9:6) yields that for any $\eta \in X \stackrel{\text{def}}{=} \operatorname{lin}\{\xi - \kappa_E(a_n)\colon\ n \geq 0\} = \operatorname{lin}(\{\xi\} \cup \{\kappa_E(a_n)\colon\ n > 0\})$ (the last equality is valid because $a_0 = 0$):

$$\|\eta\| \leq 2\sup\{|\eta(\phi)|\colon\ \phi \in M\}. \tag{9:8}$$

Observe that the set of all $\eta \in E^{**}$ that satisfy (9:8) is norm closed in E^{**}. (Indeed, $M \subset \bar{B}_{E^*}$ and therefore the right-hand side of (9:8) defines as semi-norm q on E^{**} such that $q(\eta) \leq 2\|\eta\|$ for any $\eta \in E^{**}$.) In particular, (9:8) is valid for all $\eta \in \bar{X}$ (where $\bar{X}$ is the norm closure of X). Finally, since $\kappa_E(D) \subset X$, we infer that $\kappa_E(\bar{Z}) \subset \bar{X}$ and, consequently, $\kappa_E(c) \in \bar{X}$. So, (9:8) applied to $\eta = \xi - \kappa_E(c)$ $(\in \bar{X})$ combined with (9:7) gives $\xi = \kappa_E(c)$.

The above argument shows that $W \subset \kappa_E(E)$. Since W is weak* compact, it follows from Proposition 8.11 that the set $K \stackrel{\text{def}}{=} \kappa_E^{-1}(W)$ is weakly compact. Since $D \subset K$, also $c \in K$. Now Lemma 9.9 applied to $K \cap \bar{Z}$ implies that $(a_n)_{n=1}^\infty$ has a subsequence that is weakly convergent to c (since c is a limit point of the sequence $(a_n)_{n=1}^\infty$), and finally we are done. □

Proof of Theorem 9.2. Observe that each weakly sequentially compact set satisfies (wcpc). By Corollary 9.10, the same is true for weakly compact sets. So, it remains to prove the additional claim. To this end, let A be as specified in the theorem. We claim A in bounded. Indeed, if A is unbounded, Corollary 5.16 yields that there exists $\phi \in E^*$ such that $\phi(A)$ is unbounded. Equivalently, there exists a one-to-one sequence $(a_n)_{n=1}^\infty \subset A$ such that $\lim_{n\to\infty} |\phi(a_n)| = \infty$. Observe that then for any vector $b \in E$, its weak neighbourhood $\{x \in E\colon\ |\phi(x)| < |\phi(b)| + 1\}$ contains only a finite number of points from $D \stackrel{\text{def}}{=} \{a_n\colon\ n > 0\}$ and, consequently, D witnesses that (wcpc) does not hold for A. Thus, A is bounded.

It follows from Lemma 9.12 that the weak closure K of A is weakly compact. Hence, K is weakly sequentially compact as well (by Corollary 9.10). Now take any $B \subset K$ and a point z that belongs to the weak closure of B in E. Then $\kappa_E(z)$ is in the weak* closure of $\kappa_E(B)$. Since B satisfies (wcpc) (as a subset of a weakly compact set), thus Lemma 9.12 applies to B (in place of A). So, we conclude that there is a sequence $(b_n)_{n=1}^\infty \subset B$ that converges weakly to some $c \in E$ such that $\kappa_E(c) = \kappa_E(z)$. But then $c = z$ and the proof is finished. □

The following result is a direct consequence of the Eberlein theorem and Theorem 8.22.

9.13 Corollary.
A Banach space is reflexive iff every bounded sequence has a weakly convergent subsequence.

(*proof—exercise*)

9.14 Example.
According to Conway, the following example is due to von Neumann. Let $A = \{x(m,k)\colon\ 1 \leq m < k\}$ where $x(m,k) = (x_n(m,k))_{n=1}^\infty$ and $x_m(m,k) = 1$, $x_k(m,k) = m$ and $x_n(m,k) = 0$ otherwise. Then 0 belongs to the weak closure of A in ℓ_2, but there is no sequence of elements of A that converges weakly to 0. The details are left to interested readers.

We end this chapter with another result attributed to Kreĭn and Šmulian:

9.15 Theorem. (Kreĭn-Šmulian theorem on weak compactness)
If a subset of a Banach space is weakly compact, then its closed convex hull is weakly compact as well.

10 Vector integral

In this chapter we present the most classical version of the so-called *Pettis* integral. The details are specified below.

10.1 Definition.
Let E be a locally convex T_2-space and $(\Omega, \mathfrak{M}, \mu)$ a measure space. A function $f\colon \Omega \to E$ is *Pettis integrable* (or, more precisely, *Pettis μ-integrable*) if:

- for any $\phi \in E^*$, the function $\phi \circ f\colon \Omega \to \mathbb{K}$ is both $\mathfrak{M}$-measurable and μ-integrable; and
- there exists a vector $a \in E$ such that for all $\phi \in E^*$:

  (10:1) $$\int_\Omega \phi \circ f \,\mathrm{d}\mu = \phi(a).$$

  (Since E^* separates the points of E, the above a is uniquely determined by (10:1).)

If these two conditions hold, we call the above vector a the *Pettis integral of f* and denote it by $\int_\Omega f\,\mathrm{d}\mu$. So, if f is Pettis integrable, then:
$$\forall \phi \in E^*\colon\ \phi\Big(\int_\Omega f(\omega)\,\mathrm{d}\mu(\omega)\Big) = \int_\Omega \phi(f(\omega))\,\mathrm{d}\mu(\omega).$$

For simplicity, in the above context, we will call a function $f\colon \Omega \to E$ *weakly measurable* (or, more precisely, *weakly $\mathfrak{M}$-measurable*) if $\phi \circ f$ is $\mathfrak{M}$-measurable for all $\phi \in E^*$.

We begin with a few basic properties of the concepts introduced in Definition 10.1.

10.2 Proposition.
For a function $f = (f_1, \dots, f_n)\colon \Omega \to \mathbb{K}^n$ where $(\Omega, \mathfrak{M}, \mu)$ is a measure space (and $n > 0$ is finite), the following conditions are equivalent:

(i) *f is Pettis integrable;*

(ii) *$f_j\colon \Omega \to \mathbb{K}$ is $\mathfrak{M}$-measurable and μ-integrable (in the ordinary sense) for $j = 1, \dots, n$.*

Moreover, if f is Pettis integrable, then
$$\int_\Omega f\,\mathrm{d}\mu = \Big(\int_\Omega f_1\,\mathrm{d}\mu, \dots, \int_\Omega f_n\,\mathrm{d}\mu\Big).$$

(*proof—exercise*)

10.3 Proposition.
Let $(\Omega, \mathfrak{M})$ be a measurable space and E a locally convex T_2-space.

(A) *For any measure μ on $\mathfrak{M}$, the set of all Pettis μ-integrable functions $f\colon \Omega \to E$ is a vector space (with pointwise operations) and the Pettis integral (w.r.t. μ) is a linear operator on that space.*

(B) *For any weakly $\mathfrak{M}$-measurable function $f\colon \Omega \to E$ the set M of all non-negative measures $\mu\colon \mathfrak{M} \to [0,\infty]$ such that f is Pettis μ-integrable is a cone (in the sense that $s\mu + r\nu \in M$ for all $\mu, \nu \in M$ and real scalars $s, r \geq 0$) and the function $M \ni \mu \mapsto \int_\Omega f\,\mathrm{d}\mu \in E$ is additive and positively homogeneous.*

(C) *Let* $(\Lambda, \mathfrak{N})$ *be a measurable space,* $\mu\colon \mathfrak{M} \to [0,\infty]$ *a measure and* $\tau\colon \Omega \to \Lambda$ *be a measurable function (that is,* $\tau^{-1}(N) \in \mathfrak{M}$ *for any* $N \in \mathfrak{N}$*). Let* $\nu\colon \mathfrak{N} \to [0,\infty]$ *stand for the transport of* μ *via* τ*; that is,* $\nu(B) \stackrel{\text{def}}{=} \mu(\tau^{-1}(B))$ *for all* $B \in \mathfrak{N}$*. For any weakly measurable function* $f\colon \Lambda \to E$*, the function* $f \circ \tau$ *is weakly measurable as well, and* f *is Pettis* ν*-integrable iff* $f \circ \tau$ *is Pettis* μ*-integrable. Moreover, if* f *is Pettis integrable, then*

$$\int_\Lambda f \,\mathrm{d}\nu = \int_\Omega f \circ \tau \,\mathrm{d}\mu.$$

(D) *For any Pettis* μ*-integrable function* $f\colon \Omega \to E$ *(where* $\mu\colon \mathfrak{M} \to [0,\infty]$ *is a measure) and any continuous linear operator* $T\colon E \to F$ *(where* F *is a locally convex* T_2*-space) the function* $T \circ f$ *is Pettis* μ*-integrable as well and* $\int_\Omega T \circ f \,\mathrm{d}\mu = T(\int_\Omega f \,\mathrm{d}\mu)$*.*

(proof—exercise)

10.4 Proposition. (Generalised dominated convergence theorem)
Let $(\Omega, \mathfrak{M}, \mu)$ *be a finite measure space and* E *a locally convex* T_2*-space. Further, let* $f_n\colon \Omega \to E$ $(n > 0)$ *as well as* $g\colon \Omega \to E$ *be Pettis* μ*-integrable. If the vectors* $f_n(\omega)$ *converge weakly to* $g(\omega)$ *for* μ*-almost all* $\omega \in \Omega$ *and the set* $B \stackrel{\text{def}}{=} \bigcup_{n=1}^\infty f_n(\Omega)$ *is bounded (in the sense of TVS's), then*

$$\int_\Omega f_n \,\mathrm{d}\mu \to \int_\Omega g \,\mathrm{d}\mu \qquad (n \to \infty)$$

in the weak topology.

Proof. Fix $\phi \in E^*$. It follows from our assumptions that the functions $\phi \circ f_n$ converge pointwise μ-almost everywhere to $\phi \circ g$ and are uniformly bounded (as $\phi(B)$ is bounded in $\mathbb{K}$). So, we infer from the (classical) Lebesgue's dominated convergence theorem that $\lim_{n\to\infty} \int_\Omega \phi \circ f_n \,\mathrm{d}\mu = \int_\Omega \phi \circ g \,\mathrm{d}\mu$, which the conclusion of the proposition easily follows from. □

The following simple observation appears to be a key property of the Pettis integral.

10.5 Lemma.
Let $(\Omega, \mathfrak{M}, \mu)$ *be a probabilistic space and* W *be a closed convex set in a locally convex* T_2*-space* E*. If* $u\colon \Omega \to E$ *is Pettis integrable and* $u(\omega) \in W$ *for* μ*-almost all* $\omega \in \Omega$*, then* $\int_\Omega u \,\mathrm{d}\mu \in W$*.*

Proof. Let $a \in E \setminus W$ be arbitrary. It follows from Theorem 7.16 (p. 33) that there exists a functional $\phi \in E^*$ such that $\operatorname{Re}\phi(a) < \inf \operatorname{Re}\phi(W)$. Then $(\phi(\int_\Omega f \,\mathrm{d}\mu) =)$ $\int_\Omega \phi \circ f \,\mathrm{d}\mu \neq \phi(a)$, since $\operatorname{Re}\phi \circ f > \operatorname{Re}\phi(a)$ μ-almost everywhere and $\mu(\Omega) = 1$. Consequently, $\int_\Omega f \,\mathrm{d}\mu \neq a$ and we are done. □

10.6 Lemma.
If $f\colon \Omega \to K$ *is weakly* $\mathfrak{M}$*-measurable where* $(\Omega, \mathfrak{M})$ *is a measurable space and* K *is a weakly compact set in a locally convex* T_2*-space, then* $u \circ f\colon \Omega \to \mathbb{C}$ *is* $\mathfrak{M}$*-measurable for any weakly continuous function* $u\colon K \to \mathbb{C}$*.*

Proof. Below we consider K with the topology induced from the weak topology of E. The set A of all weakly continuous functions $u\colon K \to \mathbb{C}$ for which $u \circ f$ is $\mathfrak{M}$-measurable is a closed unital subalgebra of $C(K, \mathbb{C})$ such that $\bar{u} \in A$ for any $u \in A$. Moreover, A contains $E^* \restriction K$ (cf. Theorem 8.6, p. 37) and therefore A separates the points of K. Consequently, it follows from the Stone-Weierstrass theorem that $A = C(K, \mathbb{C})$ and we are done. □

10.7 Theorem. (Pettis integrable functions)
Let $(\Omega, \mathfrak{M}, \mu)$ *be a probabilistic space and* E *a locally convex* T_2*-space. If* $f\colon \Omega \to E$ *is weakly* $\mathfrak{M}$*-measurable and the weak closure* K *of the convex hull of* $f(\Omega)$ *is weakly compact, then* f *is Pettis integrable and* $\int_\Omega f \,\mathrm{d}\mu \in K$*.*

Proof. For further purposes (of this chapter), instead of using only $\phi \in E^*$, we will deal here with the set Ξ of all weakly continuous affine functions $\phi\colon K \to \mathbb{K}$. For any such ϕ let C_ϕ consist of all $a \in K$ such that $\phi(a) = \int_\Omega \phi \circ f \,\mathrm{d}\mu$. (Note that the last integral exists since $\phi \circ f$ is measurable—by Lemma 10.6—and bounded.) In other words, $C_\phi = K \cap \phi^{-1}(\{\int_\Omega \phi \circ f \,\mathrm{d}\mu\})$. Hence C_ϕ is a weakly compact set. Now fix a finite number of functions $\phi_1, \ldots, \phi_n \in \Xi$ and set $\psi \stackrel{\text{def}}{=} (\phi_1, \ldots, \phi_n)\colon K \to \mathbb{K}^n$. We infer from Proposition 10.2 that the function $F \stackrel{\text{def}}{=} \psi \circ f\colon \Omega \to \mathbb{K}^n$ is Pettis integrable. Moreover, since $F(\Omega) \subset \psi(K)$ and $\psi(K)$ is both convex and compact (as the image of such a set under a weakly continuous affine map), an application of Lemma 10.5 yields that $\int_\Omega F \,\mathrm{d}\mu \in \psi(K)$. In particular, there exists a point $b \in K$ such that $(\int_\Omega \phi_1 \circ f \,\mathrm{d}\mu, \ldots, \int_\Omega \phi_n \circ f \,\mathrm{d}\mu) = (\phi_1(b), \ldots, \phi_n(b))$. Consequently, $b \in \bigcap_{j=1}^n C_{\phi_j}$. So, the family $\{C_\phi\}_{\phi \in \Xi}$ is centered and we conclude from the weak compactness of K that $\Delta \stackrel{\text{def}}{=} \bigcap_{\phi \in \Xi} C_\phi$ is non-empty. A notice that $\phi(c) = \int_\Omega \phi \circ f \,\mathrm{d}\mu$ for any $c \in \Delta$ ($\subset K$) finishes the proof (recall that Ξ contains all functions of the form $\phi \restriction K$ where $\phi \in E^*$, thanks to Theorem 8.6, p. 37). □

Direct consequences of the above result follow.

10.8 Corollary.
Let E, $(\Omega, \mathfrak{M}, \mu)$ and $f\colon \Omega \to E$ be, respectively, a Banach space, a probabilistic space and a weakly measurable function. In each of the following cases f is Pettis integrable:

(a) *the weak closure of $f(\Omega)$ is weakly compact; or*

(b) *f is norm bounded and E is reflexive.*

Proof. Item (a) follows from Theorems 10.7 and 9.15, whereas (b) is a direct consequence of Theorems 10.7 and 8.22. □

10.9 Corollary.
Let E, $(\Omega, \mathfrak{M}, \mu)$ and $f\colon \Omega \to E$ be, respectively, a locally convex T_2-space, a probabilistic space and a weakly measurable function such that the closure K of the convex hull of $f(\Omega)$ is compact. Further, let F be a locally convex T_2-space and $P\colon K \to F$ a continuous affine function. Then both f and $P \circ f$ are Pettis μ-integrable and $\int_\Omega P \circ f \,\mathrm{d}\mu = P(\int_\Omega f \,\mathrm{d}\mu)$.

Proof. Both the sets K and $L \stackrel{\text{def}}{=} P(K)$ are convex and compact. Consequently, the topologies of these sets (induced from the given topologies of the entire spaces E and F) coincide with their weak topologies. In particular, K and L are weakly compact and P is continuous in the weak topologies of K and L. We infer that $P \circ f$ is weakly measurable. Further, it follows from the proof of Theorem 10.7 that both f and $P \circ f$ are Pettis μ-integrable and

(10:2) $$\psi\Big(\int_\Omega f \,\mathrm{d}\mu\Big) = \int_\Omega \psi \circ f \,\mathrm{d}\mu$$

for any continuous affine function $\psi\colon K \to \mathbb{K}$ (why?). So, for any $\phi \in E^*$ (10:2) holds for $\psi = \phi \circ P$, which yields $\phi(P(\int_\Omega f \,\mathrm{d}\mu)) = \phi(\int_\Omega P \circ f \,\mathrm{d}\mu)$. Since the functionals from E^* separate the points of E, the conclusion of the result follows. □

10.10 Corollary.
For every probabilistic Borel measure μ on a compact Hausdorff space X and any weakly compact convex set K in a locally convex T_2-space, each weakly continuous function $u\colon X \to K$ is Pettis μ-integrable.

(*proof—exercise*)

Now we focus on a special context (that is quite useful in functional analysis)—namely, on Pettis integration on measures spaces of the form (X, μ) where X is a compact Hausdorff space and μ is a regular probabilistic Borel measure on X.

For simplicity, we call a function $f\colon X \to E$ *wccc* (where X is a compact T_2-space and E is locally convex and Hausdorff) if it is **w**eakly **c**ontinuous and the weak closure of the **c**onvex hull of $f(X)$ is weakly **c**ompact. It follows from Corollary 10.10 that each wccc function is Pettis μ-integrable for any $\mu \in \mathrm{Prob}_r(X)$ (for the notation, consult Example 8.20, p. 41). We use $WC(X, E)$ to denote the set of all wccc functions from X into E.

10.11 Theorem.
Let X and E be, respectively, compact T_2-space and a locally convex T_2-space.

(A) *The set $WC(X,E)$ is a linear subspace of the vector space of all weakly continuous functions (from X into E) and for any $\mu \in \mathrm{Prob}_r(\mu)$ the function $WC(X,E) \ni f \mapsto \int_X f\,\mathrm{d}\mu \in E$ is linear and has the following property:*

($\star$) *If $f_1, f_2, \ldots \in WC(X,E)$ converge pointwise to $g \in WC(X,E)$ and the weak closure of $\bigcup_{n=1}^{\infty} f_n(X)$ is weakly compact, then the vectors $\int_X f_n\,\mathrm{d}\mu$ converge weakly to $\int_X g\,\mathrm{d}\mu$.*

(B) *For any $f \in WC(X,E)$, the function $\mathrm{Prob}_r(X) \ni \mu \mapsto \int_X f\,\mathrm{d}\mu \in E$ is affine and continuous in the weak topologies of both $\mathrm{Prob}_r(X)$ and E (cf. Example 8.20, p. 41).*

Proof. Item (A) is partially a special case of Propositions 10.3 and 10.4 and is left to the reader—here we focus only the continuity postulated in (B). To this end, we fix a net $(\mu_\sigma)_{\sigma\in\Sigma} \subset \mathrm{Prob}_r(X)$ that converges to $\lambda \in \mathrm{Prob}_r(X)$ in the weak topology of that space (cf. Example 8.20, p. 41). To show that then the vectors $\int_X f\,\mathrm{d}\mu_\sigma$ converge weakly to $\int_X f\,\mathrm{d}\lambda$, it is enough to check that $\lim_{\sigma\in\Sigma} \phi(\int_X f\,\mathrm{d}\mu_\sigma) = \phi(\int_X f\,\mathrm{d}\lambda)$ for any $\phi \in E^*$ (why?). But this is an immediate consequence of (8:2), because $\phi \circ f \in C(X,\mathbb{K})$. The details are left to the reader. □

10.12 Corollary.
Let K be a compact set in a locally convex T_2-space and let X stand for the closure of $\mathrm{ext}(K)$. Then the function

$$\Lambda\colon \mathrm{Prob}_r(X) \ni \mu \mapsto \int_X \mathrm{id}\,\mathrm{d}\mu \in K$$

is a continuous (in the weak topology of its domain) affine surjection.

Proof. Thanks to Theorem 10.11, we only need to show that Λ is surjective (indeed, all the values of Λ lie in K—by Lemma 10.5—and the original topology of K coincides with the weak topology, thanks to the compactness of K). To this end, observe that $\int_X \mathrm{id}\,\mathrm{d}\delta_a = a$ for any $a \in X$ where δ_a is a unique probabilistic measure supported on $\{a\}$ (defined in Example 8.20, p. 41). So, $\mathrm{ext}(K) \subset \Lambda(\mathrm{Prob}_r(X))$. Finally, $\Lambda(\mathrm{Prob}_r(X))$ is a closed compact set contained in K and therefore it coincides with K, by the Kreĭn-Milman theorem (Theorem 7.22, p. 35). □

10.13 Theorem. ('Converse' of the Kreĭn-Milman Theorem)
Let Z be a subset of a compact convex set K in a locally convex T_2-space E. Then $K = \overline{\mathrm{conv}}(Z)$ iff $\mathrm{ext}(K) \subset \bar{Z}$.

Proof. The 'if' part of the theorem is a direct consequence of Theorem 7.22 (p. 35). To see the 'only if' part, set $X \stackrel{\text{def}}{=} \bar{Z}$ and consider the function $\Lambda\colon \mathrm{Prob}_r(X) \ni \mu \mapsto \int_X \mathrm{id}\,\mathrm{d}\mu \in K$. Similarly as presented in the previous proof, one shows that Λ is continuous and affine, and $X \subset M \stackrel{\text{def}}{=} \Lambda(\mathrm{Prob}_r(X))$. So, it follows from our assumption (that $K = \overline{\mathrm{conv}}(Z)$) that $M = K$. In particular, for any $b \in \mathrm{ext}(K)$ the set $F \stackrel{\text{def}}{=} \Lambda^{-1}(\{b\})$ is non-empty. We infer from Proposition 7.20 (p. 34) that F is a face of $\mathrm{Prob}_r(X)$. Being compact convex and non-empty, F has an extreme point (thanks to Lemma 7.21, p. 35), say μ. Then (again by Proposition 7.20) μ is an extreme point of $\mathrm{Prob}_r(X)$, which means that μ has the form $\mu = \delta_a$ for some $a \in X$. But then $b = \Lambda(\mu) = a$ and, consequently, $b \in X$. □

*10.14 **Remark**.*
It follows from Theorem 10.7 that for any measure $\mu \in \mathrm{Prob}(K)$ where K is a compact set in a locally convex T_2-space, the identity on K is Pettis μ-integrable and $\int_K \mathrm{id}\,\mathrm{d}\mu$ belongs to K. The point $\int_K \mathrm{id}\,\mathrm{d}\mu$ is called the *barycenter* of μ.

11 Riesz representation theorem for $C(K)$-spaces

In this chapter X is reserved to denote a non-empty locally compact Hausdorff space, $\mathscr{B}(X)$ stands for the σ-algebra of all Borel subsets of X (that is, $\mathscr{B}(X)$ is the smallest σ-algebra that contains all open subsets of X) and $\mathcal{P}(Y)$ (where Y is an arbitrary set) is used to denote the collection of all subsets of Y.

11.1 Definition.
A function $f\colon X \to \mathbb{R}$ is *compactly supported* if the set $\operatorname{supp} f \overset{\text{def}}{=} \overline{f^{-1}(\mathbb{R}\setminus\{0\})}$, called the *support* of f, is compact. By $C_c(X)$ we will denote the real vector space of all continuous real-valued functions on X that are compactly supported.

11.2 ***Remark.***
Although locally compact T_2 spaces may not be normal, in all such spaces compact sets can be separated by functions from $C_c(X)$. Namely, if K and L are two disjoint compact subsets of X, then there exists a function $f \in C_c(X)$ such that $f \restriction K \equiv 0$ and $f \restriction L \equiv 1$. The above property is a direct consequence of the existence of a one-point compactification of X (which is T_4). We will involve this property in this chapter many times.

11.3 Definition.
Let $\mathcal{K}(X)$ denote the collection of all compact subsets of X (including the empty set). A *content* on X is a set function $\mu\colon \mathcal{K}(X) \to \mathbb{R}_+$ such that the following three conditions are fulfilled:

- [(finite) additivity] $\mu(K\cup L) = \mu(K)+\mu(L)$ for any disjoint sets $K, L \in \mathcal{K}(X)$;
- [monotonicity] $\mu(K) \le \mu(L)$ whenever $K, L \in \mathcal{K}(X)$ satisfy $K \subset L$;
- [subadditivity] $\mu(K\cup L) \le \mu(K)+\mu(L)$ for any $K, L \in \mathcal{K}(X)$.

A content μ on X is said to be *regular* if for any $K \in \mathcal{K}(X)$:

$$\mu(K) = \inf\{\mu(L)\colon\ L \in \mathcal{K}(X),\ K \subset \operatorname{int} L\}.$$

11.4 Definition.
Let $\mathfrak{M}$ be a σ-algebra of subsets of X that contains $\mathscr{B}(X)$ and let $\mu\colon \mathfrak{M} \to [0,\infty]$ be a measure. A set $A \in \mathfrak{M}$ is said to be:

- *inner regular* if $\mu(A) = \sup\{\mu(K)\colon\ K \in \mathcal{K}(X),\ K \subset A\}$;
- *outer regular* if $\mu(A) = \inf\{\mu(U)\colon\ A \subset U \text{ — open in } X\}$;
- *regular* if A is both inner and outer regular.

The measure μ is called *Radon* if all the following conditions are satisfied:

- $\mu(K) < \infty$ for any $K \in \mathcal{K}(X)$;
- each set $A \in \mathfrak{M}$ is outer regular;
- each open set in X is inner regular.

A non-negative measure on $\mathfrak{M}$ that is finite on compact sets and such that all measurable sets are regular is said to be *regular*.

11.5 Definition.
A set function $\rho\colon \mathcal{P}(Y) \to [0,\infty]$ is said to be an *outer measure* if

- $\rho(\varnothing) = 0$; and
- $\rho(A) \le \rho(B)$ whenever $A \subset B \subset Y$; and
- $\rho(\bigcup_{n=1}^\infty A_n) \le \sum_{n=1}^\infty \rho(A_n)$ for any $A_1, A_2, \ldots \subset Y$.

If ρ is an outer measure, then a set $A \subset Y$ is *ρ-measurable* (in the sense of Carathéodory) if

(11:1) $$\forall B \subset Y\colon\ \rho(B) = \rho(B \cap A) + \rho(B \setminus A).$$

The collection of all ρ-measurable subsets of Y will be denoted by $\mathfrak{M}(\rho)$.

In this chapter we will use the following classical theorem from measure theory (for a proof, see Chapter 12).

11.6 Theorem. (Carathéodory theorem on outer measures)
If $\rho\colon \mathcal{P}(Y) \to [0, \infty]$ is an outer measure, then $\mathfrak{M}(\rho)$ is a σ-algebra and the restriction of ρ to $\mathfrak{M}(\rho)$ is a measure.

We begin with basic properties of contents and Radon measures.

11.7 Proposition.
Let ν be a content on X and let $\mu\colon \mathcal{K}(X) \to \mathbb{R}$ be given by $\mu(K) = \inf\{\nu(L)\colon\ L \in \mathcal{K}(X),\ K \subset \operatorname{int} L\}$. Then μ is a regular content on X such that $\nu \leq \mu$.

Proof. It follows from the monotonicity of ν that $\nu \leq \mu$. It is also clear that μ admits only non-negative real values (why?) and is both monotone and subadditive. So, we only need to show that μ is additive and regular. To this end, fix two disjoint sets K and L from $\mathcal{K}(X)$. Then there are two disjoint sets P and Q such that $K \subset \operatorname{int} P$ and $L \subset \operatorname{int} Q$ (why?). Then for any compact set $M \subset X$ whose interior contains $K \cup L$ we obtain $\nu(M) \geq \nu(M \cap (P \cup Q)) = \nu(M \cap P) + \nu(M \cap Q) \geq \mu(K) + \mu(L)$ and, consequently, $\mu(K \cup L) \geq \mu(K) + \mu(L)$. On the other hand, if $K \subset \operatorname{int} R$ and $L \subset \operatorname{int} S$ where $R, S \in \mathcal{K}(X)$, then $\nu(R) + \nu(S) \geq \nu(R \cap P) + \nu(S \cap Q) = \nu(R \cap (P \cup Q)) \geq \mu(K \cup L)$ (as $K \cup L \subset \operatorname{int}(R \cap (P \cup Q))$). So, $\mu(K) + \mu(L) \geq \mu(K \cup L)$, which shows that μ is a content.

Finally, for any $K \in \mathcal{K}(X)$ and $\varepsilon > 0$ there is a compact set L such that $K \subset \operatorname{int} L$ and $\nu(L) \leq \mu(K) + \varepsilon$. There exists $M \in \mathcal{K}(X)$ such that $K \subset \operatorname{int} M$ and $M \subset \operatorname{int} L$ (why?). Then $\mu(M) \leq \nu(L) \leq \mu(K) + \varepsilon$ and therefore μ is regular. □

11.8 Proposition.

(A) *If μ is a Radon measure on X, then $\mu \restriction \mathcal{K}(X)$ is a regular content.*

(B) *If μ and λ are two Borel measures on X that are Radon and coincide on $\mathcal{K}(X)$, then $\mu = \lambda$.*

Proof. To show (A), we only need to check that the restriction ρ of μ to $\mathcal{K}(X)$ is regular. To this end, we fix $K \in \mathcal{K}(X)$ and $\varepsilon > 0$. It follows from the outer regularity of K that $\mu(U) \leq \mu(K) + \varepsilon$ for some open set $U \supset K$. Then there is a compact set $L \subset U$ such that $K \subset \operatorname{int} L$ (why?). We conclude that $\rho(L) \leq \mu(U) \leq \rho(K) + \varepsilon$ and therefore ρ is regular.

Now we pass to (B). Since each open set U in X is inner regular w.r.t. both μ and λ, we infer that $\mu(U) = \lambda(U)$. Finally, since each Borel set A in X is outer regular w.r.t. both μ and λ, we conclude that $\mu(A) = \lambda(A)$ as well. □

11.9 Proposition.
Let $\mu\colon \mathscr{B}(X) \to [0, \infty]$ be a Radon measure. If $A \in \mathscr{B}(X)$ satisfies $\mu(A) < \infty$, then A is regular.
In particular, finite Radon measures are regular.

Proof. Since A is outer regular, there is an open set $U \supset A$ such that $\mu(U) < \infty$. Further, since U is locally compact and T_2 and $\mathscr{B}(U) \subset \mathscr{B}(X)$ and the restriction of μ to U is Radon (on U), we may and do assume that $U = X$; that is, we assume μ is finite. Now let $W = X \cup \{\omega\}$ be a one-point compactification of X. (So, W is a compact Hausdorff space.) We extend μ to a finite Borel measure $\lambda\colon \mathscr{B}(W) \to \mathbb{R}_+$ by the formula $\lambda(B) \stackrel{\text{def}}{=} \mu(B \cap W)$. Observe that each Borel set in W is inner regular (w.r.t. λ). We will now show that each set $B \in \mathscr{B}(W)$ is outer regular (so, in particular, λ is Radon). If $\omega \notin B$, we have nothing to do. So, assume $\omega \in B$ and fix $\varepsilon > 0$. There exists an open (in X) set $V \subset X$ such that $B \cap X \subset V$ and $\mu(V) \leq \mu(B \cap X) + \varepsilon$. Further, since U is inner regular (w.r.t. μ), there exists a compact set $K \subset U$ such that $\mu(U) \leq \mu(K) + \varepsilon$. Then $V \cup (W \setminus K)$ is open in W and contains B, and $\lambda(V \cup (W \setminus K)) \leq \lambda(V) + \lambda(W \setminus K) = \mu(V) + \mu(U \setminus K) = \mu(V) + \mu(U) - \mu(K) \leq \mu(B \cap X) + 2\varepsilon = \lambda(B) + 2\varepsilon$, which shows that B is outer regular.

Now it follows from the outer regularity of $W \setminus A$ that for each $\varepsilon > 0$ there is an open (in W) set $V \supset W \setminus A$ such that $\lambda(V) \leq \lambda(W \setminus A) + \varepsilon$. Equivalently (since λ is a finite measure!), $\lambda(A) \leq \lambda(W \setminus V) + \varepsilon$. So, to finish the proof, it remains to note that $W \setminus V$ is compact and contained in A, and hence belongs to $\mathcal{K}(X)$. □

11.10 Proposition.

(A) *If μ is a Radon measure, then so is $r\mu$ for any real $r > 0$.*

(B) *A sum of two Radon (resp. regular) Borel measures is Radon (resp. regular) as well.*

(C) *Let μ and λ be two Borel non-negative measures such that μ is absolutely continuous w.r.t. λ and λ is Radon (resp. regular). If*

- *μ is finite and λ is σ-finite; or*
- *X is σ-compact,*

then μ is Radon (resp. regular) iff μ is finite on compact sets.

It is worth noting here that item (C) is false in general (that is, when X is not σ-compact and either λ is not σ-finite or μ is not finite).

Proof of Proposition 11.10. Item (A) is left as an easy exercise. To show (B), we fix two Radon (resp. regular) Borel measures μ and ν on X. It is sufficient to check that $A \in \mathscr{B}(X)$ is outer (resp. inner) regular w.r.t. $\mu + \nu$ provided it is so w.r.t. both μ and ν.

First assume A is outer regular w.r.t. μ and ν. If $\mu(A) + \nu(A) = \infty$, we have nothing to do. Thus, we assume that both $\mu(A)$ and $\nu(A)$ are finite. For a fixed ε there are two open sets U and V that contain A and satisfy $\mu(U) \leq \mu(A) + \varepsilon$ and $\nu(U) \leq \nu(A) + \varepsilon$. Then $U \cap V$ is an open superset of A such that $\mu(U \cap V) + \nu(U \cap V) \leq \mu(A) + \nu(A) + 2\varepsilon$, which shows that A is outer regular w.r.t. $\mu + \nu$.

Now assume A is inner regular w.r.t. μ and ν. Fix a real number m such that $m < \mu(A) + \nu(A)$. Then there exist two real numbers p and q such that $m = p + q$, $p < \mu(A)$ and $q < \nu(A)$. It follows from the inner regularity of A (w.r.t. to μ and ν) that there are two compact sets K and L contained in A such that $p \leq \mu(K)$ and $q \leq \mu(L)$. Then $K \cup L \in \mathcal{K}(X)$ is compact and $m = p + q \leq \mu(K \cup L) + \nu(K \cup L)$, and we are done.

Now we pass to (C). We only need to show that if μ is finite on compact sets, then it is Radon (resp. regular). As in the proof of (B), it suffices to show that $A \in \mathscr{B}(X)$ is outer (resp. inner) regular w.r.t. μ if it is so w.r.t. λ.

First assume A is inner regular w.r.t. λ. Write $A = \bigcup_{n=1}^{\infty} A_n$ where $A_n \in \mathscr{B}(X)$ are of finite measure λ. For each $n > 0$ there exists a sequence $K_1^{(n)}, K_2^{(n)}, \ldots \subset A_n$ of compact sets such that $\lim_{m\to\infty} \lambda(K_m^{(n)}) = \lambda(A_n)$. We infer that $\lambda(A_n \setminus \bigcup_{m=1}^{\infty} K_m^{(n)}) = 0$ (why?) and, consequently, $\mu(A_n \setminus \bigcup_{m=1}^{\infty} K_m^{(n)}) = 0$. Then $\mu(A \setminus \bigcup_{n,m} K_m^{(n)}) = 0$ as well and therefore $\mu(\bigcup_{n,m} K_m^{(n)}) = \mu(A)$. So, $L_n \stackrel{\text{def}}{=} \bigcup_{k=1}^{n} \bigcup_{m=1}^{n} K_m^{(n)}$ is compact, contained in A and $\lim_{n\to\infty} \mu(L_n) = \mu(A)$, which shows that A is inner regular w.r.t. μ.

Now assume $A \in \mathscr{B}(X)$ is arbitrary (then A is outer regular w.r.t. λ). If $\mu(A) = \infty$, we have nothing to do. Thus, we assume $\mu(A) < \infty$ and fix $\varepsilon > 0$. Notice that λ is σ-finite (even if we only assume that X is σ-compact). So, we can write $A = \bigcup_{n=1}^{\infty} A_n$ where $A_n \in \mathscr{B}(X)$ satisfy $\lambda(A_n) < \infty$. Fix for a moment $N > 0$. There is a decreasing sequence $U_1^{(N)}, U_2^{(N)}, \ldots$ of open sets such that $A_N \subset \bigcap_{n=1}^{\infty} U_n^{(N)}$, $\lambda(U_1^{(N)}) < \infty$ and $\lim_{n\to\infty} \lambda(U_n^{(N)}) = \lambda(A_N)$ (why?). Then $\lambda((\bigcap_{n=1}^{\infty} U_n^{(N)}) \setminus A_N) = 0$ (why?) and, consequently, $\mu((\bigcap_{n=1}^{\infty} U_n^{(N)}) \setminus A_N) = 0$. So,

$$\mu(A_N) = \mu\Big(\bigcap_{n=1}^{\infty} U_n^{(N)}\Big). \tag{11:2}$$

First assume μ is finite. Then (11:2) implies that

$$\mu(A_N) = \lim_{n\to\infty} \mu(U_n^{(N)}) \tag{11:3}$$

(why?). We infer that there is an index $m > 0$ such that $\mu(U_m^{(N)}) \leq \mu(A_N) + \frac{\varepsilon}{2^N}$. Setting $V_N \stackrel{\text{def}}{=} U_m^{(N)}$, we obtain $A \subset \bigcup_{n=1}^{\infty} V_n$ and $\mu((\bigcup_{n=1}^{\infty} V_n) \setminus A) \leq \sum_{n=1}^{\infty} \mu(V_n \setminus A_n) \leq \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon$. Hence $\mu(\bigcup_{n=1}^{\infty} V_n) \leq \mu(A) + \varepsilon$, which shows that A is outer regular w.r.t. μ, provided μ is finite.

Finally, assume X is σ-compact. Then X can be expressed as $X = \bigcup_{n=1}^{\infty} K_n$ where each K_n is compact and $K_n \subset \operatorname{int} K_{n+1}$ (exercise). Then, in the above proof we can set $A_n = A \cap K_n$ and (for each $N > 0$) $U_1^{(N)} = \operatorname{int} K_{N+1}$. In such a situation, (11:2) implies that (11:3) holds as well (why?). So, the argument presented above (below (11:3)) works perfectly also in that case, which finishes the proof. □

11.1 From a regular content to a Radon measure

The aim of this section is to prove the following

11.11 Theorem. (Extending content to a Radon measure)
Every regular content extends to a unique Radon Borel measure.

The proof will be divided into a few steps. To shorten statements, we fix a regular content $\rho\colon \mathcal{K}(X) \to \mathbb{R}_+$. We use τ to denote the topology of X (that is, τ stands for the collection of all open sets in X). We define $\rho'\colon \tau \to [0,\infty]$ and $\rho^*\colon \mathcal{P}(X) \to [0,\infty]$ as follows:

$$\rho'(U) \stackrel{\text{def}}{=} \sup\{\rho(K)\colon\ K \in \mathcal{K}(X),\ K \subset U\} \qquad (U \in \tau)$$

and

$$\rho^*(A) \stackrel{\text{def}}{=} \inf\{\rho'(U)\colon\ U \in \tau,\ A \subset U\} \qquad (A \subset X).$$

11.12 Lemma.
The function ρ^* is an outer measure that extends both ρ and ρ'. Moreover,

$$\rho^*\Big(\bigcup_{n=1}^{\infty} U_n\Big) = \sum_{n=1}^{\infty} \rho^*(U_n) \tag{11:4}$$

for any countable collection $\{U_n\}_{n>0}$ of pairwise disjoint open sets.

Proof. Since ρ is additive (and real-valued), we infer that $\rho(\varnothing) = 0$. It is clear that ρ' is monotone and, consequently, so is ρ^*. In particular, ρ^* extends ρ' and for any compact set K and open set U:

- $K \subset U \implies \rho(K) \le \rho'(U)$;
- $U \subset K \implies \rho'(U) \le \rho(K)$.

The former property implies that $\rho^*(K) \ge \rho(K)$ for any compact set K. On the other hand, for any $K \in \mathcal{K}(X)$ and $\varepsilon > 0$ there exists a set $L \in \mathcal{K}(X)$ such that $K \subset \operatorname{int} L$ and $\rho(L) \le \rho(K)+\varepsilon$. Then $\rho^*(K) \le \rho'(\operatorname{int} L) \le \rho(L) \le \rho(K)+\varepsilon$, which shows that ρ^* extends ρ.

Consider a sequence of open sets $V_1, V_2, \dots$ in X. We will show that

$$\rho^*\Big(\bigcup_{n=1}^{\infty} V_n\Big) \le \sum_{n=1}^{\infty} \rho^*(V_n). \tag{11:5}$$

To this end, fix a compact set $K \subset \bigcup_{n=1}^{\infty} V_n$. It follows from its compactness that $K \subset \bigcup_{n=1}^{N} V_n$ for some finite $N > 0$. There exists a partition of unity (on K) $\{v_n\colon K \to [0,1]\}_{n=1}^{N}$ (consisting of continuous functions on K that sum up pointwise on K to 1) such that $v_n^{-1}((0,1]) \subset V_n$ for $n = 1, \dots, N$. For such n set $L_n \stackrel{\text{def}}{=} v_n^{-1}([\frac{1}{N},1])$. L_n is a compact subset of V_n and $K = \bigcup_{n=1}^{N} L_n$ (why?). Since ρ is subadditive, a simple induction argument shows that $\rho(K) \le \sum_{n=1}^{N} \rho(L_n)$ $(\le \sum_{n=1}^{N} \rho^*(V_n))$. So, $\rho^*(K) \le \sum_{n=1}^{\infty} \rho^*(V_n)$. Since K was arbitrary, we get (11:5).

With the aid of (11:5) we will easily show that ρ^* is an outer measure. To this end, fix a sequence $A_1, A_2, \dots$ of subsets of X. If $\sum_{n=1}^{\infty} \rho^*(A_n) = \infty$, then we have nothing to do. So, assume the last series converges and fix $\varepsilon > 0$. For any n take an open set $V_n \supset A_n$ such that $\rho^*(V_n) \le \rho^*(A_n) + \frac{\varepsilon}{2^n}$. Then $\bigcup_{n=1}^{\infty} A_n \subset \bigcup_{n=1}^{\infty} V_n$ and (thanks to (11:5)) $\rho^*(\bigcup_{n=1}^{\infty} A_n) \le \rho^*(\bigcup_{n=1}^{\infty} V_n) \le \sum_{n=1}^{\infty} \rho^*(V_n) \le \varepsilon + \sum_{n=1}^{\infty} \rho^*(A_n)$, which implies that ρ^* is an outer measure.

Finally, assume that $U_1, U_2, \dots$ are pairwise disjoint open sets in X. Taking into account (11:5), we only need to show that the right-hand side of (11:4) is not less than its right-hand side. We may and do assume that $\rho^*(\bigcup_{n=1}^{\infty} U_n) < \infty$ (in particular, $\rho^*(U_n) < \infty$ for all n). Fix $\varepsilon > 0$ and for each $n > 0$ take a compact set $L_n \subset U_n$ such that $\rho^*(U_n) \le \rho(L_n) + \frac{\varepsilon}{2^n}$. Then for each $N > 0$ the set $K \stackrel{\text{def}}{=} \bigcup_{n=1}^{N} L_n$ is compact and contained in $V \stackrel{\text{def}}{=} \bigcup_{n=1}^{\infty} U_n$. So, $\rho^*(V) \ge \rho(K)$. Since the sets L_n are pairwise disjoint, we infer from the additivity of ρ that $\rho(K) = \sum_{n=1}^{N} \rho(L_n)$. Thus, letting $N \to \infty$, we obtain

$$\rho^*(V) \ge \sum_{n=1}^{\infty} \rho(L_n) \ge \sum_{n=1}^{\infty} (\rho^*(U_n) - \frac{\varepsilon}{2^n}) = \Big(\sum_{n=1}^{\infty} \rho^*(U_n)\Big) - \varepsilon,$$

which finishes the proof. □

Since ρ^* is an outer measure, we may consider the σ-algebra $\mathfrak{M} \overset{\text{def}}{=} \mathfrak{M}(\rho^*)$ introduced in Definition 11.5.

11.13 Lemma.
$\mathcal{K}(X) \subset \mathfrak{M}$.

Proof. Fix $K \in \mathcal{K}(X)$ and let A be an arbitrary subset of X. Since ρ^* is an outer measure, we have $\rho^*(A) \leq \rho^*(A \cap K) + \rho^*(A \setminus K)$. To prove the reverse inequality, we may assume $\rho^*(A) < \infty$ and fix $\varepsilon > 0$. There is an open set $U \supset A$ such that $\rho^*(U) \leq \rho^*(A) + \varepsilon$. Since $U \setminus K$ is open (and ρ^* extends ρ'), there exists a compact set $L \subset U \setminus K$ for which $\rho^*(U \setminus K) \leq \rho(L) + \varepsilon$. Then $K \cap L = \varnothing$ and, consequently, $U \setminus L$ is an open set containing $A \cap K$. Similarly, there is a compact set $M \subset U \setminus L$ satisfying $\rho^*(U \setminus L) \leq \rho(M) + \varepsilon$. Then the sets L and M are disjoint and therefore $\rho(L \cup M) = \rho(L) + \rho(M)$. We conclude that $\rho^*(A \cap K) + \rho^*(A \setminus K) \leq \rho^*(U \setminus L) + \rho^*(U \setminus K) \leq \rho(M) + \varepsilon + \rho(L) + \varepsilon = \rho(L \cup M) + 2\varepsilon \leq \rho^*(U) + 2\varepsilon \leq \rho^*(A) + 3\varepsilon$, and we are done. $\square$

11.14 Lemma.
$\mathfrak{M} \supset \tau$.

Proof. Fix an open set U and an arbitrary set $A \subset X$. As in the previous proof, we only need to check that $\rho^*(A \cap U) + \rho^*(A \setminus U) \leq \rho^*(A)$ provided that $\rho^*(A) < \infty$. To this end, fix $\varepsilon > 0$ and take an open set $V \supset A$ such that $\rho^*(V) \leq \rho^*(A) + \varepsilon$. Then also $\rho^*(A \cap U) \leq \rho^*(V \cap U)$ and $\rho^*(A \setminus U) \leq \rho^*(V \setminus U)$. Since $V \cap U$ is open, there is a compact set $K \subset U \cap V$ that satisfies $\rho(U \cap V) \leq \rho^*(K) + \varepsilon$. Note that $V \setminus U \subset V \setminus K$ and $K = V \cap K$. So, it follows from Lemma 11.13 that $\rho^*(A \cap U) + \rho^*(A \setminus U) \leq \rho^*(V \cap U) + \rho^*(V \setminus U) \leq \rho(K) + \varepsilon + \rho^*(V \setminus K) = \rho^*(V \cap K) + \rho^*(V \setminus K) + \varepsilon = \rho^*(V) + \varepsilon \leq \varrho^*(A) + 2\varepsilon$, which finishes the proof. $\square$

Proof of Theorem 11.11. We have already established that ρ^* is an outer measure such that $\tau \subset \mathfrak{M}(\rho^*)$. So, it follows from Theorem 11.6 that $\mathscr{B}(X) \subset \mathfrak{M}(\rho^*)$ and that $\mu \overset{\text{def}}{=} \rho^* \restriction \mathscr{B}(X)$ is a measure. Further, Lemma 11.12 implies that μ extends ρ. In particular, μ is finite on compact sets. Other axioms of a Radon measure follow from the very definition of ρ^*, whereas uniqueness of μ follows from Proposition 11.8. $\square$

11.2 From non-negative linear functionals to Radon measures

11.15 Definition.
A linear functional $L\colon C_c(X) \to \mathbb{R}$ is said to be *non-negative* if $L(f) \geq 0$ for all non-negative $f \in C_c(X)$. Equivalently, L is non-negative if $L(f) \leq L(g)$ whenever $f, g \in C_c(X)$ satisfy $f \leq g$.

11.16 Example.
Let μ be a Borel non-negative measure on X that is finite on compact sets. Then the assignment $f \mapsto \int_X f \, \mathrm{d}\mu$ correctly defines a non-negative linear functional on $C_c(X)$. The aim of this section is to show that there are no other non-negative linear functionals on $C_c(X)$.

11.17 Proposition.
If X is compact, then each non-negative linear functional on $C(X)$ is automatically continuous.

Proof. Let $L\colon C(X) \to \mathbb{R}$ be linear and non-negative. Denoting by $j\colon X \to \mathbb{R}$ the function constantly equal to 1, we have:

$$-\|f\| \cdot j \leq f \leq \|f\| \cdot j \qquad (f \in C(X)).$$

So, $-\|f\| L(j) \leq L(f) \leq \|f\| L(j)$ and, consequently, $\|L\| = L(j)$. $\square$

The main goal of this section is to prove the following classical result from functional analysis.

11.18 Theorem. (Riesz representation theorem for positive functionals)
For any non-negative linear functional $L\colon C_c(X) \to \mathbb{R}$ there exists a unique Radon measure $\mu\colon \mathscr{B}(X) \to [0,\infty]$ such that

$$L(f) = \int_X f \,\mathrm{d}\mu \qquad (f \in C_c(X)). \tag{11:6}$$

As in the previous section, we divide the proof into a few steps.

11.19 Lemma.
Let μ and ν be two Radon Borel measures on X. If

$$\int_X f \,\mathrm{d}\mu = \int_X f \,\mathrm{d}\nu$$

for all non-negative $f \in C_c(X)$, then $\mu = \nu$.

Proof. Fix a compact set K and a Radon measure λ on X. Denote by $\mathcal{F}$ the collection of all compactly supported continuous functions $f\colon X \to [0,1]$ that are constantly equal to 1 on K. Observe that $\lambda(K) \le \int_X f \,\mathrm{d}\lambda$ for any $f \in \mathcal{F}$. So,

$$\lambda(K) \le \inf\Big\{ \int_X f \,\mathrm{d}\lambda\colon\ f \in \mathcal{F} \Big\}. \tag{11:7}$$

On the other hand, for each $\varepsilon > 0$ there exists an open set $U \supset K$ such that $\lambda(U) \le \lambda(K) + \varepsilon$. Next, there is $f \in \mathcal{F}$ that vanishes outside U (why?). Then $\int_X f \,\mathrm{d}\lambda \le \lambda(U) \le \lambda(K) + \varepsilon$. This argument, combined with (11:7), yields

$$\lambda(K) = \inf\Big\{ \int_X f \,\mathrm{d}\lambda\colon\ f \in \mathcal{F} \Big\}.$$

So, we conclude that μ and ν coincide on $\mathcal{K}(X)$ and hence $\mu = \nu$, by Proposition 11.8. □

To shorten statements, from now on to the end of this section we fix a non-negative linear functional $L\colon C_c(X) \to \mathbb{R}$. For each $K \in \mathcal{K}(X)$ we denote by $\mathcal{F}(K)$ the set of all functions $f\colon X \to [0,1]$ from $C_c(X)$ that are constantly equal to 1 on K. Further, we define $\rho\colon \mathcal{K}(X) \to \mathbb{R}$ by

$$\rho(K) = \inf\{ L(f)\colon\ f \in \mathcal{F}(K) \}. \tag{11:8}$$

11.20 Lemma.
The function ρ is a regular content.

Proof. It follows from the non-negativity of L that rho is non-negative. Since $\mathcal{F}(K) \subset \mathcal{F}(L)$ for any $K, L \in \mathcal{K}(X)$ such that $K \supset L$, we conclude that ρ is monotone. Further, if $K, L \in \mathcal{K}(X)$ are arbitrary and $f \in \mathcal{F}(K)$ and $g \in \mathcal{F}(L)$, then $h \stackrel{\text{def}}{=} \min(f{+}g, 1)$ belongs to $\mathcal{F}(K{\cup}L)$ and therefore $\rho(K{\cup}L) \le L(h) \le L(f{+}g) = L(f){+}L(g)$. Now passing to infima on the right-hand side of this inequality gives us subadditivity of ρ. Finally, if K and L are disjoint, there are two functions $u \in \mathcal{F}(K)$ and $v \in \mathcal{F}(L)$ such that $u \cdot v \equiv 0$ (why?), which implies that $u + v \le 1$. So, if $f \in \mathcal{F}(K \cup L)$ is arbitrary, then $fu \in \mathcal{F}(K)$, $fv \in \mathcal{F}(L)$ and $f(u+v) \le f$. Consequently, $\rho(K) + \rho(L) \le L(fu) + L(fv) = L(f(u+v)) \le L(f)$. Passing to infimum on the right-hand side of this inequality yields $\rho(K) + \rho(L) \le \rho(K \cup L)$, and we are done.

Now to show that ρ is regular, we fix $K \in \mathcal{K}(X)$ and $\varepsilon > 0$. There exists $f \in \mathcal{F}(K)$ such that $L(f) \le \rho(K) + \varepsilon$. Choose $r > 1$ such that $rL(f) \le L(f) + \varepsilon$ and consider $g \stackrel{\text{def}}{=} \min(rf, 1)$ and $M \stackrel{\text{def}}{=} g^{-1}(\{1\})$. Observe that M is a compact set such that $K \subset \operatorname{int} M$ (why?) and $g \in \mathcal{F}(M)$, which implies that $\rho(M) \le L(g) \le L(rf) = rL(f) \le L(f) + \varepsilon \le L(f) + 2\varepsilon$, and we are done. □

Having the above result, we apply Theorem 11.11 to get a Radon measure μ on X that extends ρ. We will now show that (11:6) holds. The main part of the proof is contained in the following

11.21 Lemma.
$\int_X f(x)\,\mathrm{d}\mu(x) \le L(f)$ for any non-negative $f \in C_c(X)$.

Proof. We may and do assume that $f \not\equiv 0$. Set $K \stackrel{\text{def}}{=} \operatorname{supp} f$ and $C \stackrel{\text{def}}{=} 2\|f\|_\infty$, and fix $\varepsilon > 0$. Observe that there are only countably many values $r \ge 0$ for which $\mu(f^{-1}(\{r\})) > 0$ (why?). So, we infer that there are a finite number of reals $0 = r_0 < r_1 < \ldots < r_N \le C$ such that $f(X) \subset [0, r_N]$, $r_k - r_{k-1} \le \varepsilon$ and $\mu(f^{-1}(\{r_k\})) = 0$ for any $k > 0$. Set $U_k \stackrel{\text{def}}{=} f^{-1}((r_{k-1}, r_k))$ $(k = 1, \ldots, N)$ and $V \stackrel{\text{def}}{=} \bigcup_{k=1}^N U_k$, and note that:

- $V \subset K$; and
- all the sets U_k are open and pairwise disjoint; and
- $\int_X f\,\mathrm{d}\mu = \int_V f\,\mathrm{d}\mu$.

In particular, $\mu(U_k) < \infty$ and therefore there exist a compact set $M_k \subset U_k$ such that $\mu(U_k) \le \mu(M_k) + \frac{\varepsilon}{N}$ and $g_k \in \mathcal{F}(M_k)$ that vanishes outside U_k (why?). Set $h \stackrel{\text{def}}{=} \sum_{k=1}^N r_{k-1} g_k$. Note that $h \le f$ (why?) and, consequently, $L(h) \le L(f)$. Moreover, $\mu(M_k) = \rho(M_k) \le L(g_k)$. On the other hand, setting $Q \stackrel{\text{def}}{=} \bigcup_{k=1}^N M_k$, we obtain $\mu(V \setminus Q) \le \varepsilon$ and for any $x \in Q$ there is a unique $k \in \{1, \ldots, N\}$ satisfying $x \in M_k$—then $f(x) \le r_k \le \varepsilon + r_{k-1} \le \varepsilon + h(x)$ (exercise). Hence:

$$\int_X f\,\mathrm{d}\mu = \int_V f\,\mathrm{d}\mu = \int_Q f\,\mathrm{d}\mu + \int_{V\setminus Q} f\,\mathrm{d}\mu \le \int_Q (\varepsilon + h)\,\mathrm{d}\mu + C\mu(V \setminus Q) \le \varepsilon\mu(Q) + \sum_{k=1}^N r_{k-1}\mu(M_k) + C\varepsilon$$
$$\le \varepsilon(\mu(K) + C) + \sum_{k=1}^N r_{k-1} L(g_k) = \varepsilon(\mu(K) + C) + L(h) \le \varepsilon(\mu(K) + C) + L(f),$$

which finishes the proof. ∎

Proof of Theorem 11.18. Uniqueness of μ follows from Lemma 11.19. To establish its existence, we continue the notation introduced in this section. So, we have a Radon measure μ on X such that

$$\mu(K) = \inf\{L(f)\colon\ f \in \mathcal{F}(K)\} \qquad (K \in \mathcal{K}(X)). \tag{11:9}$$

To show (11:6), it suffices to verify that $\int_X f\,\mathrm{d}\mu \le L(f)$ for any $f \in C_c(X)$ (because then we may apply this inequality to both f and $-f$ to get the equality). To this end, we fix $f \in C_c(X)$ and $\varepsilon > 0$. Since L is homogeneous, we may and do assume that $\|f\|_\infty \le 1$. Set $K \stackrel{\text{def}}{=} \operatorname{supp} f$. We infer from (11:9) that there exists $g \in \mathcal{F}(K)$ such that $L(g) \le \mu(K) + \varepsilon$. Notice that then $f + g \ge 0$ (as $|f| \le 1$) and $\mu(K) \le \int_X g\,\mathrm{d}\mu$. So, it follows from Lemma 11.21 that $\int_X (f + g)\,\mathrm{d}\mu \le L(f + g)$, which yields

$$\int_X f\,\mathrm{d}\mu = \int_X (f+g)\,\mathrm{d}\mu - \int_X g\,\mathrm{d}\mu \le L(f+g) - \mu(K) = L(f) + L(g) - \mu(K) \le L(f) + \varepsilon,$$

and the proof is finished. ∎

11.3 Signed and complex-valued measures

11.22 Definition.
A set function $\mu\colon \mathfrak{M} \to \mathbb{K}$ (where $\mathfrak{M}$ is a σ-algebra of subsets of Y) is said to be a *scalar measure* if $\mu(\bigcup_{n=1}^\infty A_n) = \sum_{n=1}^\infty \mu(A_n)$ for any sequence $A_1, A_2, \ldots$ of pairwise disjoint sets from $\mathfrak{M}$. When $\mathbb{K} = \mathbb{R}$, a scalar measure is called *signed*; and when $\mathbb{K} = \mathbb{C}$, one speaks about *complex* measures.

The *variation* of a scalar measure μ is a set function $|\mu|\colon \mathfrak{M} \to [0, \infty]$ given by

$$|\mu(A)| \stackrel{\text{def}}{=} \sup\Big\{\sum_{n=1}^\infty |\mu(B_n)|\colon\ B_1, B_2, \ldots \in \mathfrak{M} \text{ are pairwise disjoint and contained in } A\Big\} \qquad (A \in \mathfrak{M}).$$

The quantity $|\mu|(Y)$ is called the *total variation* of μ.

11.23 Theorem.
*The variation of a scalar measure is a **finite** measure.*

We divide the proof of the above result into separate lemmas. Since each signed measure is complex as well, without loss of generality, we may assume that $\mathbb{K} = \mathbb{C}$. For simplicity, we fix a complex measure μ defined of a σ-algebra $\mathfrak{M}$ of subsets of Y.

11.24 Lemma.
The variation of μ is a measure.

Proof. Let $A_1, A_2, \dots$ be pairwise disjoint sets from $\mathfrak{M}$. Observe that if $C_1, C_2, \dots \in \mathfrak{M}$ are pairwise disjoint subsets of $\bigcup_{n=1}^{\infty} A_n$, then $C_k \bigcap A_1, C_k \bigcap A_2, \dots$ (for each $k > 0$) are pairwise disjoint as well and their union coincides with C_k. So, we conclude that $\mu(C_k) = \sum_{n=1}^{\infty} \mu(C_k \bigcap A_n)$. At the same time, $C_1 \cap A_n, C_2 \cap A_n, \dots$ are pairwise disjoint and contained in A_n (separately for each n). Thus, it follows that

$$\sum_{k=1}^{\infty} |\mu(C_k)| = \sum_{k=1}^{\infty} \Big| \sum_{n=1}^{\infty} \mu(C_k \cap A_n) \Big| \leq \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} |\mu(C_k \cap A_n)| = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} |\mu(C_k \cap A_n)| \leq \sum_{n=1}^{\infty} |\mu|(A_n).$$

Passing to the supremum on the very left-hand side of the above inequality, we obtain $|\mu|(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} |\mu|(A_n)$. On the other hand, if $N > 0$ is arbitrarily fixed and for each $n \in \{1, \dots, N\}$, $B_1^{(n)}, B_2^{(n)}, \dots \in \mathfrak{M}$ are pairwise disjoint and contained in $\mathfrak{M}$, then all the sets $B_m^{(n)}$ (where $n \leq N$ and $m > 0$) are pairwise disjoint subsets of $\bigcup_{n=1}^{\infty} A_n$ (and there are countably many of them). Consequently, $\sum_{n=1}^{N} (\sum_{k=1}^{\infty} |\mu(B_k^{(n)})|) \leq |\mu|(\bigcup_{n=1}^{\infty} A_n)$. Again, passing to the supremum in each of the summands of the left-hand side of the last inequality, we obtain $\sum_{n=1}^{N} |\mu|(A_n) \leq |\mu|(\bigcup_{n=1}^{\infty} A_n)$. Now letting $N \to \infty$, we obtain $\sum_{n=1}^{\infty} |\mu|(A_n) \leq |\mu|(\bigcup_{n=1}^{\infty} A_n)$, which shows that $|\mu|$ is a measure. □

Now we pass to the most intriguing part of Theorem 11.23—that is, we will show that $|\mu|(Y) < \infty$. The next result is a crucial step.

11.25 Lemma.
For any $A \in \mathfrak{M}$, $|\mu|(A) \leq 4\sqrt{2} \sup\{|\mu(B)|\colon\ B \in \mathfrak{M},\ B \subset A\}$.

Proof. For simplicity, denote by $\phi_1, \phi_2, \phi_3, \phi_4 \colon \mathbb{C} \to \mathbb{R}$ four $\mathbb{R}$-linear functionals given by $\phi_1(z) = -\phi_2(z) \stackrel{\text{def}}{=} \operatorname{Re} z$ and $\phi_3(z) = -\phi_4(z) \stackrel{\text{def}}{=} \operatorname{Im} z$. For $k = 1, 2, 3, 4$ set $I_k \stackrel{\text{def}}{=} \{z \in \mathbb{C}\colon\ \phi_k(z) \geq \max(|\operatorname{Re} z|, |\operatorname{Im} z|)\}$. Note that $\mathbb{C} = \bigcup_{k=1}^{4} I_k$. Now take a sequence $B_0, B_1, \dots \in \mathfrak{M}$ of pairwise disjoint subsets of A and set $J_k \stackrel{\text{def}}{=} \{n \in \mathbb{N}\colon\ \mu(B_n) \in I_1\} \setminus \bigcup_{j=0}^{k-1} J_j$ ($J_0 \stackrel{\text{def}}{=} \varnothing$) and $D_k \stackrel{\text{def}}{=} \bigcup_{n \in J_k} B_n$ ($k = 1, 2, 3, 4$). Then $D_k \in \mathfrak{M}$, $\mu(D_k) = \sum_{n \in J_k} \mu(B_n)$ ($\sum_{n \in \varnothing} \stackrel{\text{def}}{=} 0$) and $\mathbb{N} = \bigsqcup_{j=1}^{4} J_k$. In particular,

$$\phi_k(\mu(D_k)) = \sum_{n \in J_k} \phi_k(\mu(B_n)).$$

It folows from the very definition of the sets I_k and J_k that for each $k = 1, 2, 3, 4$ and $n \in J_k$, $|\mu(B_n)| \leq \sqrt{2}\phi_k(B_n)$. Consequently,

$$\sum_{n=0}^{\infty} |\mu(B_n)| = \sum_{k=1}^{4} \sum_{n \in J_k} |\mu(B_k)| \leq \sqrt{2} \sum_{k=1}^{4} \sum_{n \in J_k} \phi_k(\mu(B_n)) = \sqrt{2} \sum_{k=1}^{4} \phi_k(\mu(D_k))$$
$$\leq 4\sqrt{2} \sup\{|\mu(C)|\colon\ C \in \mathfrak{M},\ C \subset A\}.$$

Passing to the supremum on the left-hand side of the above inequality concludes the proof. □

11.26 Lemma.
If $A \in \mathfrak{M}$ satisfies $|\mu|(A) = \infty$, then there exists $B \in \mathfrak{M}$ such that $B \subset A$, $|\mu|(B) = \infty$ and $|\mu(A \setminus B)| \geq 1$.

Proof. It follows from Lemma 11.25 that $\sup\{|\mu(C)|\colon\ C \in \mathfrak{M},\ C \subset A\} = \infty$. So, there is $D \in \mathfrak{M}$ contained in A such that $|\mu(D)| \geq |\mu(A)|+1$. Then $|\mu(A \setminus D)| = |\mu(A)-\mu(D)| \geq |\mu(D)|-|\mu(A)| \geq 1$ and $|\mu|(D)+|\mu|(A\setminus D) = |\mu|(A) = \infty$ (thanks to Lemma 11.24). So, there is $B \in \{D, A \setminus D\}$ that satisfies $|\mu|(B) = \infty$. Then automatically $|\mu(A \setminus B)| \geq 1$, by the above estimations. □

Proof of Theorem 11.23. It follows from Lemma 11.24 that $|\mu|$ is a measure. To convince oneself that it is finite, we argue by a contradiction and apply infinitely many times Lemma 11.26: if $|\mu|(Y) = \infty$, then starting from $A_0 \stackrel{\text{def}}{=} Y$, for each $n > 0$ we find $A_n \in \mathfrak{M}$ such that $A_n \subset A_{n-1}$, $|\mu(A_n)| = \infty$ and

$$|\mu(B_n)| \geq 1 \tag{11:10}$$

where $B_n \stackrel{\text{def}}{=} A_{n-1} \setminus A_n$. Note that then the sets $B_1, B_2, \ldots$ are pairwise disjoint. So, $\mu(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu(B_n)$ and, consequently, $\lim_{n\to\infty} \mu(B_n) = 0$, which contradicts (11:10) and finishes the proof. □

11.27 Proposition.

(A) *Each complex measure μ has a unique representation in the form $\mu = \mu_1 + i\mu_2$ where μ_1 and μ_2 are two signed measures.*

(B) *Each scalar measure is the difference of two finite non-negative measures.*

(C) *The set of all scalar messures on a fixed σ-algebra is a vector space and the total variation is a norm on this space.*

Proof. Part (A) trivially follows from the property that if μ is a complex measure, then $\bar{\mu}$ is a complex measure as well where $\bar{\mu}(A) = \overline{\mu(A)}$; and $\frac{\mu+\bar{\mu}}{2}$ and $\frac{\mu-\bar{\mu}}{2i}$ are signed measures.

Further, for each scalar measure μ on $\mathfrak{M}$ and any $A \in \mathfrak{M}$, one has $|\mu(A)| \leq |\mu|(A)$. It folows that if μ is signed, then both $\mu_+ \stackrel{\text{def}}{=} \frac{1}{2}(|\mu| + \mu)$ and $\mu_- \stackrel{\text{def}}{=} \frac{1}{2}(|\mu| - \mu)$ are non-negative measures, which easily implies (B). Finally, it is clear that a linear combination of two scalar measure is a scalar measure as well. We leave it as an exercise that the assignment $\mu \mapsto |\mu|(Y)$ defines a norm. □

11.28 Definition.
Let $\mathcal{M}(\mathfrak{M}, \mathbb{K})$ be the vector space of all scalar measures on $\mathfrak{M}$ (with values in $\mathbb{K}$), equipped with the norm of total variation.

If μ is a signed measure on $\mathfrak{M}$ and $f\colon Y \to \mathbb{K}$ is bounded and $\mathfrak{M}$-measurable, we define the integral of f w.r.t. μ as $\int_Y f\,\mathrm{d}\mu \stackrel{\text{def}}{=} \int_Y f\,\mathrm{d}\mu_+ - \int_Y f\,\mathrm{d}\mu_-$ (cf. the proof of Proposition 11.27). It is easy to check that this integral is bilinear as a function in two variables f and μ.

11.29 Theorem.
For any σ-algebra $\mathfrak{M}$, $\mathcal{M}(\mathfrak{M}, \mathbb{K})$ is a Banach space.

(*proof—exercise*)

11.30 Theorem.
For any $\mu \in \mathcal{M}(\mathfrak{M}, \mathbb{K})$ there exists an $\mathfrak{M}$-measurable function $u\colon Y \to \mathbb{K}$ such that $|u| \equiv 1$ and $\mu(A) = \int_A u\,\mathrm{d}|\mu|$ for any $A \in \mathfrak{M}$. Moreover, for such u and any bounded $\mathfrak{M}$-measurable function $f\colon Y \to \mathbb{K}$:

$$\int_Y f\,\mathrm{d}\mu = \int_Y f u\,\mathrm{d}|\mu|.$$

Proof. For simplicity, set $\lambda \stackrel{\text{def}}{=} |\mu|$.

Firt assume μ is signed. Observe that then both μ_+ and μ_- (see the proof of Proposition 11.27) are absolutely continuous w.r.t. λ. So, it follows from the classical Radon-Nikodym theorem that $\mu_\pm(A) = \int_A u_\pm \,\mathrm{d}\lambda$ for some non-negative λ-integrable functions u_+ and u_-. Then

$$\mu(A) = \int_A u \,\mathrm{d}\lambda \qquad (A \in \mathfrak{M}) \tag{11:11}$$

where $u \stackrel{\text{def}}{=} u_+ - u_-$.

Now assume μ is a complex measure. Fix for a moment $A \in \mathfrak{M}$ such that $\lambda(A) = 0$. Then for any $B \in \mathfrak{M}$ contained in A we also have $|\mu(B)| = 0$ and therefore $\operatorname{Re}\mu(B) = 0$ and $\operatorname{Im}\mu(B) = 0$. So, both $\operatorname{Re}\mu$ and $\operatorname{Im}\mu$ vanish at any set contained in A, which implies that $|\operatorname{Re}\mu|(A) = |\operatorname{Im}\mu|(A) = 0$. In other words, $|\operatorname{Re}\mu|$ and $|\operatorname{Im}\mu|$ are absolutely continuous w.r.t. λ. Another usage of the Radon-Nikodym theorem gives us two non-negative λ-integrable functions g_1 and g_2 such that $|\operatorname{Re}\mu|(A) = \int_A g_1 \,\mathrm{d}\lambda$ and, similarly, $|\operatorname{Im}\mu|(A) = \int_A g_2 \,\mathrm{d}\lambda$. It folows from the first part of the proof that $\operatorname{Re}\mu(A) = \int_A v_1 \,\mathrm{d}|\operatorname{Re}\mu|$ and $\operatorname{Im}\mu(A) = \int_A v_2 \,\mathrm{d}|\operatorname{Im}\mu|$. Combining all these properties together and using standard techniques of the classical measure theory, we infer that (11:11) holds for $u \stackrel{\text{def}}{=} v_1 g_1 + i v_2 g_2$.

Now assume (11:11) holds (and the field is arbitrary). Take any set $A \in \mathfrak{M}$ such that $\lambda(A) > 0$. It follows from the last cited formula that $\frac{1}{\lambda(A)}|\int_A u \,\mathrm{d}\lambda| = \frac{|\mu(A)|}{\lambda(A)} \leq 1$. Consequently, $|u| \leq 1$ λ-a.e. (exercise). So, changing u on a measure zero set, we may and do assume that $|u| \leq 1$. Now fix $A \in \mathfrak{M}$ and take any sequence $B_1, B_2, \ldots \in \mathfrak{M}$ of pairwise disjoint subsets of A. Then $\sum_{n=1}^\infty |\mu(B_n)| = \sum_{n=1}^\infty |\int_{B_n} u \,\mathrm{d}\lambda| \leq \sum_{n=1}^\infty \int_{B_n} |u| \,\mathrm{d}\lambda = \int_{\bigcup_{n=1}^\infty B_n} |u| \,\mathrm{d}\lambda \leq \int_A |u| \,\mathrm{d}\lambda$. Passing to the supremum on the very left-hand side of this inequality, we obtain $\lambda(A) \leq \int_A |u| \,\mathrm{d}\lambda$. In particular, this last inequality is valid for $A = Z \stackrel{\text{def}}{=} \{x\colon\ |u(x)| < 1\}$, from which it follows that $\lambda(Z) = 0$. So, $|u| \equiv 1$ λ-a.e. So, we may and do assume that $|u(x)| = 1$ for all x.

The additional claim of the theorem is left to the reader. □

11.31 Example.
We leave as an exercise that if λ is an arbitrary non-negative measure (not necessarily finite), f is a scalar-valued λ-integrable function and μ is given by $\mu(A) = \int_A f \,\mathrm{d}\lambda$, then $|\mu|(A) = \int_A |f| \,\mathrm{d}\lambda$.

Now let's get back to the context of locally compact spaces and regular measures.

11.32 Definition.
A scalar measure defined on all Borel sets of X is said to be *regular* if so is its variation. The set of all regular scalar measures defined on $\mathscr{B}(X)$ is denoted by $\mathcal{M}_r(X, \mathbb{K})$. (When $\mathbb{K} = \mathbb{R}$, we also write $\mathcal{M}_r(X)$ instead of $\mathcal{M}_r(X, \mathbb{R})$.)

11.33 Proposition.

(A) *The space $\mathcal{M}_r(X, \mathbb{K})$ is a closed vector subspace of $\mathcal{M}(\mathscr{B}(X), \mathbb{K})$.*

(B) *For any $\mu \in \mathcal{M}_r(X, \mathbb{K})$ the function $L\colon C_0(X, \mathbb{K}) \ni f \mapsto \int_X f \,\mathrm{d}\mu \in \mathbb{K}$ is a bounded linear functional such that $\|L\| \leq \|\mu\|$.*

11.4 From bounded linear functionals to scalar regular measures

The aim of this section is to prove the following

11.34 Theorem. (Riesz representation theorem for bounded functionals)
Let E stand for one of $C_c(X, \mathbb{K})$ or $C_0(X, \mathbb{K})$. For any bounded linear functional ϕ on E there is a unique $\mu \in \mathcal{M}_r(X, \mathbb{K})$ such that

$$\phi(f) = \int_X f \,\mathrm{d}\mu \qquad (f \in E). \tag{11:12}$$

Moreover, $\|\phi\| = \|\mu\|$. In particular, the dual E^* of E is linearly isometric to $\mathcal{M}_r(X, \mathbb{K})$.

12 Appendix: proof of the Carathéodory theorem on outer measures