Chapter 5
Consequences of Convexity

In this chapter, we wish to explore the geometric aspects of the Hahn–Banach Theorem. The crucial property, it turns out, is local convexity. We will first recall some notions from general topology and then introduce the concept of a topological vector space. These spaces, which include Banach spaces, are sufficiently complex that we can say something interesting about their structure. Banach spaces are topological vector spaces where the topology is determined by a complete norm, and in this chapter we will get some idea of how they fit into a more general topological framework.

5.1 General Topology

Let E be a set. A *topology* τ on E is a collection of subsets called *open sets* satisfying the following three criteria:

1. The collection τ contains both E and the empty set $\emptyset$.
2. If $\{U_i\}_{i \in I}$ is a (possibly uncountable) family of sets in τ, then $\bigcup_{i \in I} U_i$ is in τ.
3. If U and V are in τ, then $U \cap V$ is in τ.

When E is equipped with a topology τ, we call the pair (E, τ) a *topological space*. When there is no ambiguity, we will suppress the τ and simply write E for the topological space (E, τ), and say U is open in E when $U \in \tau$.

If (E, τ) and (F, τ') are topological spaces, then a function $f : E \to F$ is called *continuous* if $f^{-1}(U)$ is open in E whenever U is open in F. For x a point in E, a *neighborhood* of x is any subset N of E for which there exists an open set U such that $x \in U$ and $U \subset N$.

Example 5.1 Let (E, τ) be a topological space. If every set in E is open, then τ is called the *discrete topology*. If $\tau = \{E, \emptyset\}$, then τ is called the *indiscrete topology*.

Example 5.2 Let (E, τ) and (F, τ') be two topological spaces. The product $E \times F$ can be given the *product topology*, denoted $\tau \times \tau'$, as follows: $W \subseteq E \times F$ is open if

$$W = \bigcup_{i \in I} (U_i \times V_i),$$

where $U_i \in \tau$ and $V_i \in \tau'$ for each $i \in I$, where I is a (possibly uncountable) index set.

Let (E, τ) be a topological space. The topology τ is said to have a *base* of open sets $\{U_i\}_{i \in I}$ if for each open set $V \in \tau$, there exists an index set $J \subseteq I$ such that $V = \bigcup_{i \in J} U_i$. When τ has a base, we say the base *generates* the topology τ. In Example 5.2, the product topology on $E \times F$ is generated by the base

$$\{U \times V : U \in \tau, \ V \in \tau'\}.$$

Example 5.3 (Product topology). Let I be a (possibly uncountable) index set. For each $i \in I$, let (E_i, τ_i) be a topological space. The *product topology* on the product $\prod_{i \in I} (E_i, \tau_i)$ is a topology with a base consisting of sets of the form

$$U_{i_1} \times \cdots \times U_{i_n} \times \prod_{i \in I \setminus \{i_1, \ldots, i_n\}} E_i,$$

where U_{i_j} is open in E_{i_j} for $i_j \in I$, $n \in \mathbb{N}$, and $j \in \{1, \ldots, n\}$. Observe that all but finitely many elements of the product are the entire space. This example contains Example 5.2 as a special case, because the product in that example is finite.

Example 5.4 Suppose M is a set with a metric d. We will show that the metric d determines a topology on M. For each $x \in M$ and $r > 0$, let

$$B(x, r) = \{z \in M : d(x, z) < r\}.$$

This set is the *open ball about x of radius r*. We declare a subset V of M to be open if for each $x \in V$, there is an $r > 0$ such that $B(x, r) \subseteq V$. The collection of all such open sets forms a topology on M called the *metric topology on M generated by the metric d*, or just the *metric topology* on M, if the metric d is understood.

Suppose V is open in the metric topology. For each $x \in V$, there exists a number $r_x > 0$ such that $B(x, r_x) \subseteq V$. Thus, $V = \bigcup_{x \in V} B(x, r_x)$, and so the collection of open balls forms a base for the metric topology.

Let E be a topological space and let $x \in E$. A *local base at x* is a collection η of open sets, all of which contain x, such that any neighborhood U of x contains an element of η.

In Example 5.4, any point in the metric space M has a local base. For $x \in M$, the collection of open balls $B(x, r)$ for all $r > 0$ forms a local base at x. In fact, if we consider the collection of sets $\eta = \{B(x, 1/n) : n \in \mathbb{N}\}$, then η is a countable local base at x.

A topological space (E, τ) with a countable local base at every point $x \in E$ is called *first countable*. Further, (E, τ) is called *second countable* if τ has a countable base. From Example 5.4 (and the comments following it), we see that any metric

space M is first countable; however, M will not be second countable unless M is separable. (See Exercise 5.9.)

Let (E, τ) be a topological space. We say (E, τ) is *metrizable* if there exists a metric d such that d generates the topology τ. That is, if the open balls in (E, d) form a base for the topology τ. We call (E, τ) a *Hausdorff space* if for any distinct points x and y in E there exist open sets U and V in τ such that $x \in U$, $y \in V$, and $U \cap V = \emptyset$.

Example 5.5   Any metrizable space is a Hausdorff space. To see this, suppose (M, d) is a metric space and let x and y be two distinct points in M. Since $x \neq y$, it follows that $d(x, y) > 0$. Let $\delta = d(x, y)$. Furthermore, let $U = B(x, \delta/2)$ and $V = B(y, \delta/2)$. Then U and V are open in the metric topology on M, $x \in U$, $y \in V$, and $U \cap V = \emptyset$.

Example 5.6   Any nonempty set E with the *discrete topology* (see Example 5.1) is metrizable. Define a metric on E by

$$d(x, y) = \begin{cases} 0 & \text{if } x = y, \\ 1 & \text{if } x \neq y, \end{cases} \quad \text{for } (x, y) \in E \times E.$$

It is easy to see that d is, in fact, a metric. This metric is called the *discrete metric* on E and it is not hard to show that d generates the discrete topology on E.

Of particular interest to us is the notion of *compactness*. A topological space is said to be *compact* if any open cover contains a finite open subcover. To be more precise, let X be a topological space. Then X is compact if for any collection $\mathcal{U}$ of open sets such that $X \subseteq \bigcup_{U \in \mathcal{U}} U$ there exists a finite collection $\{U_1, \dots, U_n\}$ of elements in $\mathcal{U}$ such that $X \subseteq U_1 \cup \cdots \cup U_n$.

For a (not necessarily compact) topological space, we define a *compact subset* in a similar way: A subset E of a topological space X is *compact* if any cover of E by sets open in X admits a finite subcover of E.

Some well-known properties of compact sets are treated in the exercises at the end of this chapter. (See Exercise 5.2.)

A topological space is said to be *locally compact* if every point has a compact neighborhood. Naturally, all compact spaces are locally compact, but the converse need not be true. For example, the real line $\mathbb{R}$ with its standard topology is locally compact, but not compact.

A notion of fundamental importance in topology is that of a *convergent sequence*. If X is a topological space, and $(x_n)_{n=1}^{\infty}$ is a sequence of elements from X, then $(x_n)_{n=1}^{\infty}$ is said to *converge* to a point $x \in X$ if for every open neighborhood U of x there exists an $N \in \mathbb{N}$ such that $x_n \in U$ for all $n \geq N$. In such a case, we say x is the limit of the sequence $(x_n)_{n=1}^{\infty}$ and we write $x = \lim_{n \to \infty} x_n$. (Note that this notion of a limit agrees with the standard definition of a limit in a metric space.)

In general, the limit of a sequence need not be unique. The spaces we consider, however, are Hausdorff spaces, and limits are necessarily unique in a Hausdorff space. (See Exercise 5.8.)

A subset U of a topological space X is called *sequentially open* if every sequence $(x_n)_{n=1}^\infty$ that converges to a point in U is eventually in U. That is, if there exists some $N \in \mathbb{N}$ such that $x_n \in U$ for all $n \geq N$. We call X a *sequential space* if every sequentially open set is open. Any first countable topological space is a sequential space. In particular, any metric space is a sequential space.

5.2 Topological Vector Spaces

We now consider topological spaces with additional structure, namely an underlying linear structure.

Let X be a vector space over the field $\mathbb{K}$ (which is either $\mathbb{R}$ or $\mathbb{C}$). A topology τ on X is called a *vector topology* if the maps

$$(\lambda, x) \mapsto \lambda x, \quad \lambda \in \mathbb{K}, \ x \in X,$$

and

$$(x_1, x_2) \mapsto x_1 + x_2, \quad (x_1, x_2) \in X \times X,$$

are both continuous. That is, if both scalar multiplication and addition are continuous in the topology on X. In this case, (X, τ) is called a *topological vector space*.

Example 5.7 Any normed vector space X is a topological vector space, where the topology is given by the base of open balls:

$$x + \lambda(\operatorname{int} B_X), \quad \lambda > 0, \ x \in X.$$

Equivalently, the topology on X is generated by the metric d given by the formula $d(x, y) = \|x - y\|$ for $(x, y) \in X \times X$.

A vector topology is determined by a base of neighborhoods at the origin, since sets can be translated and scaled continuously. We will denote the origin by 0. Let η be a base of neighborhoods of the origin in a topological vector space (X, τ). A set $V \in \eta$ is called *absorbent* if $X = \bigcup_{n=1}^\infty n V$. A set $V \in \eta$ is called *balanced* if $\lambda V \subseteq V$ for all scalars λ such that $|\lambda| \leq 1$.

Lemma 5.8 *In a topological vector space, any open neighborhood of the origin is absorbent.*

Proof Let X be a topological vector space. Suppose V is an open neighborhood of 0 and let $x \in X$. Scalar multiplication is continuous, and so the map $\lambda \mapsto \lambda x$ is continuous. Consequently, the set $\{\lambda : \lambda x \in V\}$ is open in $\mathbb{K}$. By assumption, V is a neighborhood of 0, and so $0 \in \{\lambda : \lambda x \in V\}$. We have established that the set $\{\lambda : \lambda x \in V\}$ is open in $\mathbb{K}$ and contains 0. Thus, it must contain $\frac{1}{n}$ for a sufficiently large $n \in \mathbb{N}$. We conclude that $\frac{x}{n} \in V$, and consequently $x \in nV$. Therefore, V is absorbent. □

Proposition 5.9 *Any topological vector space has a base of neighborhoods η of the origin such that for all $V \in \eta$: (i) V is balanced, (ii) V is absorbent, and (iii) there exists $W \in \eta$ such that $W + W \subseteq V$.*

Proof Let (X, τ) be a topological vector space and let U be a neighborhood of the origin. Let $s : \mathbb{K} \times X \to X$ be scalar multiplication, so that $s(\lambda, x) = \lambda x$ for all $\lambda \in \mathbb{K}$ and $x \in X$. By assumption, s is continuous. Thus, since U is open in X, the preimage $s^{-1}(U)$ is open in $\mathbb{K} \times X$. Certainly, $(0, 0) \in s^{-1}(U)$, and so there exists some $\delta > 0$ and an open neighborhood W of 0 in X such that $\delta B_{\mathbb{K}} \times W \subseteq s^{-1}(U)$. Therefore, $s(\delta B_{\mathbb{K}} \times W) \subseteq U$, and hence $\alpha W \subseteq U$ for all $|\alpha| \leq \delta$. Let

$$V = \bigcup_{\alpha \in \delta B_{\mathbb{K}}} \alpha W.$$

Then V is open, balanced, and contained in U. For each open neighborhood of 0, such a V can be constructed. Let η be the collection of all such balanced sets. Then *(i)* follows from the construction and *(ii)* follows from Lemma 5.8.

It remains to verify *(iii)*. Let $V \in \eta$. By the continuity of addition, there exist two open neighborhoods U_1 and U_2 of $0 \in X$ such that $U_1 + U_2 \subseteq V$. Let $U = U_1 \cap U_2$. Then U is an open neighborhood of 0 such that $U + U \subseteq V$. As demonstrated earlier in this proof, U contains a subset $W \in \eta$, and this W is the required set. □

Proposition 5.10 *Let X be a topological vector space with η a base of open sets about the origin. Then X is a Hausdorff space if and only if $\bigcap_{V \in \eta} V = \{0\}$.*

Proof Without loss of generality, we may assume that η satisfies the conclusions of Proposition 5.9.

Assume X is a Hausdorff space. Certainly $0 \in \bigcap_{V \in \eta} V$. Suppose $x \neq 0$. We will show that $x \notin \bigcap_{V \in \eta} V$. Since X is a Hausdorff space, there are open sets U and W such that $0 \in U$ and $x \in W$ and $U \cap W = \emptyset$. By assumption, η is a base of open sets about the origin, and consequently there exists a set $V_0 \in \eta$ such that $V_0 \subseteq U$. It follows that $x \notin V_0$, and so $x \notin \bigcap_{V \in \eta} V$. Therefore, $\bigcap_{V \in \eta} V = \{0\}$.

Now assume $\bigcap_{V \in \eta} V = \{0\}$. We will show that X is a Hausdorff space. Let x and y be elements of X that cannot be separated by disjoint open sets. Let $V \in \eta$. By Proposition 5.9, there exists a set $W \in \eta$ such that $W + W \subseteq V$. By assumption, $x + W$ and $y + W$ are not disjoint. Then there exist elements w_1 and w_2 in W such that

$$x + w_1 = y + w_2.$$

Therefore, $x - y = w_2 - w_1 \in W - W$. The set W is balanced, and so we conclude $x - y \in W + W \subseteq V$. This is true for every $V \in \eta$, and so $x - y \in \bigcap_{V \in \eta} V = \{0\}$. It follows that $x = y$, and consequently X is a Hausdorff space. □

In our discussions of normed spaces, a key notion was that of the dual space. In the more general context of topological vector spaces, this will remain true.

Definition 5.11 Let X be a topological vector space. The *dual space* X^* consists of all continuous linear scalar-valued functionals on X.

5.3   Some Metrizable Examples

In this section, we consider some examples of real topological vector spaces which
are metrizable, but do not have a norm structure.

Example A: $L_p(0, 1)$, $0 < p < 1$

If $0 < p < \infty$, the symbol $L_p(0, 1)$ denotes the collection of all (equivalence classes
of) Lebesgue measurable real-valued functions f on $[0, 1]$ such that

$$\|f\|_p = \left(\int_0^1 |f(t)|^p \, dt \right)^{1/p} < \infty.$$

If $p \geq 1$, then $L_p(0, 1)$ is a Banach space. If $0 < p < 1$, however, then $\| \cdot \|_p$
does not determine a norm, because it is no longer subadditive. On the other hand,
if $0 < p < 1$, then it is true that

$$\|f + g\|_p^p \leq \|f\|_p^p + \|g\|_p^p.$$

This fact follows from the proposition below.

Proposition 5.12   *If* $\{a, b\} \subseteq \mathbb{R}$ *and* $0 < p < 1$, *then* $|a + b|^p \leq |a|^p + |b|^p$.

Proof Without loss of generality, assume that a and b are nonnegative real numbers
such that $a + b = 1$. Let $a = t$ and $b = 1 - t$, and let $f(t) = t^p + (1 - t)^p$. We will
show that $f(t) \geq 1$ for all $t \in [0, 1]$. Since $f(0) = f(1) = 1$, it will suffice to show
that f is a concave function.

   We require only techniques of elementary differential calculus. Calculating the
first derivative of f, we have $f'(t) = pt^{p-1} - p(1 - t)^{p-1}$, and so f has one critical
point, which is at $t = 1/2$. Differentiating a second time, we have

$$f''(t) = p(p - 1)t^{p-2} + p(p - 1)(1 - t)^{p-2}.$$

Therefore, $f''(1/2) = 2^{3-p} p(p - 1) < 0$, and so f has a local maximum at $t = 1/2$.
The result follows.   □

   From the above proposition, we conclude that $\|f + g\|_p^p \leq \|f\|_p^p + \|g\|_p^p$ whenever
$0 < p < 1$. Consequently,

$$d(f, g) = \|f - g\|_p^p$$

determines a metric on $L_p(0, 1)$ when $0 < p < 1$. It follows that if $0 < p < 1$, then
$L_p(0, 1)$ is a metrizable space, if not a normed space.

   The metric d is even complete. The proof of this fact is similar to the case when $1 \leq
p < \infty$. Observe that the proof of the Cauchy Summability Criterion (Lemma 2.24)
requires only subadditivity of the norm, a property which is shared with $\| \cdot \|_p^p$ when
$0 < p < 1$. Consequently, we can use Lemma 2.24 to prove that $L_p(0, 1)$ is complete
when $0 < p < 1$. The details of the proof are left to the reader. (See Exercise 5.12.)

Let $B = \{f : \|f\|_p < 1\}$. Then the collection of open sets $(2^{-n}B)_{n=1}^{\infty}$ determines a countable base at 0 which satisfies the conclusions of Proposition 5.9. The first two properties are clear. To see property *(iii)*, simply observe that $2^{-N}B + 2^{-N}B \subseteq B$ whenever $N > 1/p$.

We will now compute $L_p(0,1)^*$ for $0 < p < 1$. Suppose ϕ is a continuous linear functional on $L_p(0,1)$. Then ϕ is bounded on ∂B, so that

$$\|\phi\| = \sup_{\|f\|_p=1} |\phi(f)| < \infty.$$

This should be taken as the definition of $\|\phi\|$ in this context. The function ϕ is a linear functional, but not on a normed space, and consequently the notation $\|\phi\|$ has not yet been given a meaning.

Let $f \in L_p(0,1)$ be such that $\|f\|_p = 1$. The map

$$t \mapsto \int_0^t |f(s)|^p \, ds, \quad t \in [0,1],$$

is continuous with range $[0,1]$. Therefore, by the Intermediate Value Theorem, there exists some $a \in [0,1]$ such that $\int_0^a |f(s)|^p \, ds = 1/2$.

Define two functions g and h in $L_p(0,1)$ by

$$g = f \, \chi_{(0,a)} \quad \text{and} \quad h = f \, \chi_{(a,1)}.$$

By the choice of a,

$$\|g\|_p = \left(\int_0^a |f(s)|^p \, ds \right)^{1/p} = \left(\frac{1}{2} \right)^{1/p},$$

and similarly, $\|h\|_p = (1/2)^{1/p}$.

By the linearity of ϕ, together with the definition of $\|\phi\|$, we have the two bounds $|\phi(g)| \le \|\phi\|(1/2)^{1/p}$ and $|\phi(h)| \le \|\phi\|(1/2)^{1/p}$. Thus, again using the linearity of ϕ,

$$|\phi(f)| \le 2 \cdot \|\phi\|(1/2)^{1/p} = \|\phi\| \, 2^{1-\frac{1}{p}}.$$

Taking the supremum over all functions $f \in L_p(0,1)$ with $\|f\|_p = 1$, we have

$$\|\phi\| \le \|\phi\| \, 2^{1-\frac{1}{p}}.$$

However, this can happen only if $\phi = 0$. This implies that $L_p(0,1)^* = \{0\}$.

The preceding remark guarantees that $L_p(0,1)$ does not satisfy a Hahn–Banach Theorem if $0 < p < 1$. On the other hand, since $L_p(0,1)$ is a complete metric space, even when $0 < p < 1$, we can apply the Baire Category Theorem (Theorem 4.1). It is also possible to prove a version of the Open Mapping Theorem (Theorem 4.29) and the Closed Graph Theorem (Theorem 4.35) for these spaces.

Example B: $L_0(0, 1)$

We denote by $L_0(0, 1)$ the set of all (equivalence classes of) scalar-valued Lebesgue measurable functions on $[0, 1]$. (As usual, we identify functions if they agree almost everywhere.) The topology on $L_0(0, 1)$ is determined by convergence in Lebesgue measure. More precisely, we define a set to be open when it is sequentially open, and a sequence converges when it converges in Lebesgue measure. Recall that a sequence $(f_n)_{n=1}^{\infty}$ of measurable functions *converges in Lebesgue measure* to a measurable function f if for every $\epsilon > 0$,

$$\lim_{n \to \infty} m\{t : |f(t) - f_n(t)| \geq \epsilon\} = 0,$$

where m is Lebesgue measure on $[0, 1]$.

We claim the topology on $L_0(0, 1)$ is metrizable and is induced by the metric

$$d(f, g) = \int_0^1 \frac{|f(t) - g(t)|}{1 + |f(t) - g(t)|} \, dt,$$

where f and g are measurable functions. The only property of a metric that is not immediate is the triangle inequality. In order to verify this, it suffices to show that the function $\phi(x) = x/(1 + x)$ is a nondecreasing subadditive function on $[0, \infty)$.

A simple application of the quotient rule reveals that $\phi'(x) = 1/(1 + x)^2$, and so ϕ is strictly increasing for all $x \geq 0$. To see that ϕ is subadditive on $[0, \infty)$, observe that

$$\phi(x+y) = \frac{x + y}{1 + x + y} = \frac{x}{1 + x + y} + \frac{y}{1 + x + y} \leq \frac{x}{1 + x} + \frac{y}{1 + y} = \phi(x) + \phi(y),$$

because $x \geq 0$ and $y \geq 0$. Given these properties, the triangle inequality follows readily from the fact that $d(f, g) = \int_0^1 \phi(|f(t) - g(t)|) \, dt$.

To see that the topology on $L_0(0, 1)$ coincides with that induced by the metric d, it suffices to show that the same sequences converge in each topology (since both spaces are sequential spaces). Suppose the sequence $(f_n)_{n=1}^{\infty}$ of measurable functions converges in the metric d to a measurable function f. Then $d(f, f_n) \to 0$ as $n \to \infty$. Therefore, $\frac{|f - f_n|}{1 + |f - f_n|} \to 0$ in the L_1-norm, and hence in measure. It follows that $f_n \to f$ in measure, as required.

The reverse implication, that convergence in measure implies convergence in d, is true by the Lebesgue Dominated Convergence Theorem. (We state the Lebesgue Dominated Convergence Theorem in Theorem A.17 for almost everywhere convergence, but it remains valid for sequences that converge in measure on a σ-finite measure space.)

It remains to show that the metric d is complete. Let

$$\|f\|_0 = \int_0^1 \frac{|f(t)|}{1 + |f(t)|} \, dt, \quad f \in L_0(0, 1).$$

Observe that $d(f, g) = \|f - g\|_0$ for all measurable functions f and g on $[0, 1]$. Certainly, $\| \cdot \|_0$ is not a norm (it is not homogeneous), but it does satisfy the triangle

inequality, because $\|f\|_0 = \int_0^1 \phi(|f(t)|)\,dt$ and ϕ is subadditive on $[0, \infty)$. Consequently, we may use Lemma 2.24 (the Cauchy Summability Criterion) to prove that d is a complete metric space (because the proof of Lemma 2.24 does not require homogeneity of the norm).

Suppose $(f_n)_{n=1}^\infty$ is a sequence of measurable functions such that $\sum_{n=1}^\infty \|f_n\|_0 < \infty$. Then, by Fubini's Theorem,

$$\sum_{n=1}^\infty \|f_n\|_0 = \sum_{n=1}^\infty \int_0^1 \frac{|f_n(t)|}{1 + |f_n(t)|}\,dt = \int_0^1 \left(\sum_{n=1}^\infty \frac{|f_n(t)|}{1 + |f_n(t)|} \right) dt < \infty.$$

It follows that, for almost every $t \in [0, 1]$, there exists some $M_t > 0$ such that $\sum_{n=1}^\infty \frac{|f_n(t)|}{1+|f_n(t)|} \leq M_t$. Consequently, by the subadditivity of ϕ, for every $N \in \mathbb{N}$,

$$\phi\left(\sum_{n=1}^N |f_n(t)| \right) \leq \sum_{n=1}^N \phi(|f(t)|) \leq M_t < \infty \quad \text{a.e.}(t).$$

Since ϕ is a strictly increasing function on the interval $[0, \infty)$, we conclude that, for almost every t, the sequence $(\sum_{n=1}^N |f_n(t)|)_{N=1}^\infty$ converges. Therefore, $(\sum_{n=1}^N f_n)_{N=1}^\infty$ converges almost everywhere, and hence in measure. Therefore, $L_0(0, 1)$ is a complete metric space.

As was the case in Example A (where $0 < p < 1$), the dual space of $L_0(0, 1)$ is trivial; that is, $L_0(0, 1)^* = \{0\}$. We leave the verification of this fact as an exercise. (See Exercise 5.14.)

Example C: $\omega = \mathbb{R}^\mathbb{N}$

Let J be a (possibly uncountable) index set. Let $\mathbb{R}^J$ denote the product space $\prod_{j \in J} \mathbb{R}_j$, where $\mathbb{R}_j = \mathbb{R}$ for each $j \in J$. An element x in $\mathbb{R}^J$ is a function $x : J \to \mathbb{R}$, where $x(j) \in \mathbb{R}(= \mathbb{R}_j)$ for each $j \in J$.

When the space $\mathbb{R}^J$ is equipped with the product topology, it becomes a topological vector space. The vector space operations are done pointwise; that is, if x and y are elements in $\mathbb{R}^J$, then $(x + y)(j) = x(j) + y(j)$ for each $j \in J$. Convergence, too, is pointwise: $x_n \to x$ in $\mathbb{R}^J$ as $n \to \infty$ if $x_n(j) \to x(j)$ in $\mathbb{R}_j$ as $n \to \infty$ for each $j \in J$.

If J is an *uncountable* index set, then $\mathbb{R}^J$ is not metrizable, since $\mathbb{R}^J$ with the product topology is not first countable; i.e., it does not have a countable local base at 0. (See Example 5.4.)

In this example, we are interested in *countable* index sets, and so we let $J = \mathbb{N}$. We denote $\mathbb{R}^\mathbb{N}$ by the Greek letter ω. Generally, we think of ω as the collection of all sequences in $\mathbb{R}$. If $\xi \in \omega$, we let $\xi_k = \xi(k)$ for each $k \in \mathbb{N}$, and we write $\xi = (\xi_k)_{k=1}^\infty$. In this context, the vector space operations are done coordinate-wise. Convergence is also now viewed coordinate-wise, so that $\xi^{(n)} \to \xi$ in ω as $n \to \infty$ if $\xi_k^{(n)} \to \xi_k$ in $\mathbb{R}$ as $n \to \infty$ for each $k \in \mathbb{N}$.

Unlike $\mathbb{R}^J$ when J is uncountable, the space ω is first countable. A base of neighborhoods at the origin is formed by sets of the type

$$(-\epsilon_1, \epsilon_1) \times \cdots \times (-\epsilon_n, \epsilon_n) \times \mathbb{R} \times \mathbb{R} \times \cdots, \tag{5.1}$$

where $n \in \mathbb{N}$ and $\epsilon_i > 0$ for each $i \in \{1, \ldots, n\}$. If we denote elements of ω by $\xi = (\xi_k)_{k=1}^\infty$, then the set in (5.3.1) can be written

$$\{\xi : |\xi_1| < \epsilon_1, \cdots, |\xi_n| < \epsilon_n\}.$$

To identify a countable base, consider the sets with $\epsilon_i = \frac{1}{k}$, where $k \in \mathbb{N}$, for all $i \in \{1, \ldots, n\}$ and $n \in \mathbb{N}$.

Not only is ω first countable, but it is also metrizable. Recall that ω was defined to be $\prod_{k=1}^\infty \mathbb{R}_k$. Denote the metric on $\mathbb{R}_k$ by d_k. We define a metric d on ω by

$$d(\xi, \eta) = \sum_{k=1}^\infty \frac{1}{2^k} \frac{d_k(\xi_k, \eta_k)}{1 + d_k(\xi_k, \eta_k)},$$

where $\xi = (\xi_k)_{k=1}^\infty$ and $\eta = (\eta_k)_{k=1}^\infty$.

We now wish to identify the space dual to ω. To that end, we prove the following proposition.

Proposition 5.13 *Let X be a topological vector space and let $\mathbb{K}$ be the field of scalars. A linear functional $f : X \to \mathbb{K}$ is continuous if and only if there exists a neighborhood V of 0 such that the set $f(V)$ is bounded in $\mathbb{K}$.*

Proof Let $U_\mathbb{K}$ be the open unit ball in $\mathbb{K}$. If f is continuous, then $f^{-1}(U_\mathbb{K})$ is an open neighborhood of 0, and $f(f^{-1}(U_\mathbb{K})) \subseteq B_\mathbb{K}$ is bounded in $\mathbb{K}$.

Now suppose V is a neighborhood of 0 such that $f(V)$ is bounded in $\mathbb{K}$. By definition, there is some $M > 0$ such that $f(V) \subseteq M U_\mathbb{K}$. Let $\epsilon > 0$ be given. Then $f(\frac{\epsilon}{M} V) \subseteq \epsilon U_\mathbb{K}$. Therefore, $|f(x)| < \epsilon$ whenever $x \in \frac{\epsilon}{M} V$. In other words, f is continuous at zero. Continuity then follows from the linearity of f. □

We can now use the preceding proposition to identify the continuous linear functionals on ω. Let $f \in \omega^*$. By Proposition 5.13, there must be some neighborhood V of 0 such that $f(V)$ is bounded in $\mathbb{R}$. Without loss of generality, we may assume V is a basic set, say $V = \{\xi : |\xi_1| < \epsilon_1, \cdots, |\xi_n| < \epsilon_n\}$ for some $n \in \mathbb{N}$.

The set $f(V)$ is bounded, and so there exists some $M > 0$ such that $|f(\xi)| \leq M$ for any $\xi \in V$. Let $\xi = (0, \ldots, 0, \xi_{n+1}, \ldots)$. Then $\xi \in V$, and so too is any constant multiple of ξ. Therefore, for any $K > 0$, we have that $|f(K\xi)| \leq M$, and hence $|f(\xi)| \leq M/K$, by the linearity of f. Since this inequality holds for all $K > 0$, it must be that $f(\xi) = 0$. This is true for any $\xi \in V$ having $\xi_i = 0$ for all $i \in \{1, \ldots, n\}$. Thus, because f is linear, it follows that $f(\xi) = f(\xi')$ for any ξ and ξ' that agree on the first n coordinates.

Define a function $g : \mathbb{R}^n \to \mathbb{R}$ by $g(\xi_1, \ldots, \xi_n) = f(\xi_1, \ldots, \xi_n, 0, \ldots)$. Since f is linear and continuous, it follows that $g \in (\mathbb{R}^n)^*$. Consequently, there exists $\alpha_k \in \mathbb{R}$

for each $k \in \{1, \dots, n\}$ such that

$$g(\xi_1, \dots, \xi_n) = \sum_{k=1}^{n} \alpha_k \, \xi_k, \quad (\xi_k)_{k=1}^{n} \in \mathbb{R}^n.$$

Since the value of $f(\xi)$ depends only on the first n coordinates of ξ, we conclude that

$$f(\xi) = \sum_{k=1}^{n} \alpha_k \, \xi_k, \quad \xi = (\xi_k)_{k=1}^{\infty} \in \omega.$$

5.4 The Geometric Hahn–Banach Theorem

In this section, we will meet the Hahn–Banach Theorem without the advantages of a norm structure. The key property a space must have, we shall see, is *local convexity*.

Definition 5.14 Let X be a real or complex vector space. A subset V of X is called *convex* if given any x and y in V, we have $(1 - t)x + ty \in V$ for all $t \in [0, 1]$. That is, if two points are in V, then the line segment joining them is also in V. A balanced convex set is called *absolutely convex*.

Lemma 5.15 *Let X be a real or complex vector space. A subset V of X is absolutely convex if and only if $\alpha x + \beta y \in V$ whenever x and y are in V and α and β are scalars such that $|\alpha| + |\beta| \leq 1$.*

Proof We first observe that V is absolutely convex if the latter condition holds: to show balance, take $\beta = 0$; to show convexity, let $\alpha = t - 1$ and $\beta = t$.

Now suppose V is absolutely convex. Let x and y be in V and suppose α and β are scalars such that $|\alpha| + |\beta| \leq 1$. We wish to show $\alpha x + \beta y \in V$. Observe that

$$\alpha x + \beta y = \frac{\alpha}{\alpha + \beta}(\alpha + \beta)x + \frac{\beta}{\alpha + \beta}(\alpha + \beta)y.$$

Since V is balanced, $x' = (\alpha + \beta)x$ and $y' = (\alpha + \beta)y$ are both elements of V. Thus, by convexity,

$$\alpha x + \beta y = \frac{\alpha}{\alpha + \beta}x' + \frac{\beta}{\alpha + \beta}y' \in V.$$

This completes the proof. □

Definition 5.16 A topological vector space is *locally convex* if there is a base of neighborhoods of 0 consisting of convex sets.

By Proposition 5.9, we can always take the elements of a base in a locally convex topological vector space to be balanced, and hence absolutely convex.

Example 5.17 Any normed space is locally convex. It is easy to see that balls with center at the origin are convex.
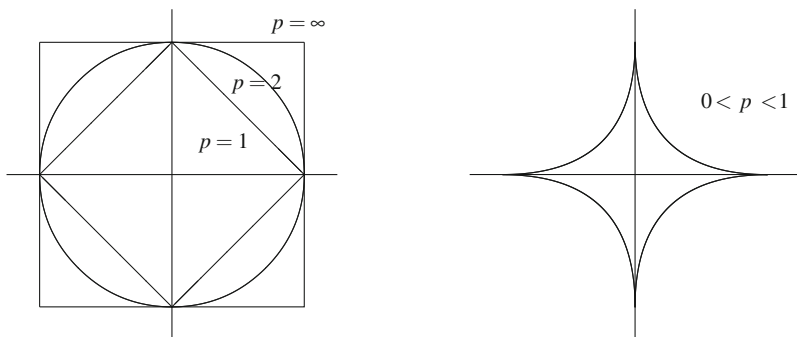
Fig. 5.1 Closed unit balls in ℓ_p^2 for various values of p

Example 5.18 Consider the space ℓ_p^2 of ordered pairs in the $\|\cdot\|_p$ norm for $p > 0$. (We use the term "norm" here even though it is not a norm when $0 < p < 1$.) If $p \geq 1$, then the unit ball is convex and balanced; however, if $0 < p < 1$, then the unit ball is balanced, but not convex. (See Fig. 5.1.)

Example 5.19 Let $X = L_p(0,1)$ for $0 < p < 1$. (See Example A in Sect. 5.3.) We claim that the only nonempty open convex subset of X is X. To show this, let V be a nonempty open convex subset of X. Without loss of generality, assume $0 \in V$. Then there exists some $\delta > 0$ such that $\delta B \subseteq V$, where $B = \{f : \|f\|_p < 1\}$. (We remind the reader that $\|\cdot\|_p$ is not a norm in this case.)

Choose any $f \in X$. Because $p < 1$, there is some $n \in \mathbb{N}$ such that $n^{p-1}\|f\|_p^p < \delta$. Pick real numbers $\{t_0, t_1, \ldots, t_n\}$ so that $0 = t_0 < t_1 < \cdots < t_n = 1$ and such that

$$\int_{t_{k-1}}^{t_k} |f(s)|^p \, ds = \frac{1}{n}\|f\|_p^p, \quad k \in \{1, \ldots, n\}.$$

For each $k \in \{1, \ldots, n\}$, let $g_k = n\,f\,\chi_{(t_{k-1}, t_k]}$. Then $\|g_k\|_p^p = n^{p-1}\|f\|_p^p < \delta$, and therefore $g_k \in \delta B \subseteq V$. This is true for each $k \in \{1, \ldots, n\}$, and so $\{g_1, \ldots, g_n\} \subseteq V$. Observe that $f = \frac{1}{n}(g_1 + \cdots + g_n)$. Since V is convex, it follows that $f \in V$. The choice of $f \in X$ was arbitrary, and so $V = X$, as required.

Note that this argument implies that $L_p(0,1)^* = \{0\}$, a fact we first observed in Example A in Sect. 5.3.

The next theorem is a geometric version of the Hahn–Banach Theorem. This version of the theorem is not set in the context of a complete normed space, but in that of a locally convex topological vector space.

Theorem 5.20 (Hahn–Banach Separation Theorem) *Let E be a real locally convex topological vector space. Let K be a closed nonempty convex subset of E. If $x_0 \notin K$, then there exists a continuous linear functional f on E such that*

$$f(x_0) > \sup_{y \in K} f(y).$$

Proof Without loss of generality, we may assume $0 \in K$. (If not, use a translation.) Since K is closed and $x_0 \notin K$, there exists some open neighborhood N of x_0 such that $N \cap K = \emptyset$. It follows that there exists an absolutely convex open neighborhood W of 0 such that $(x_0 + W) \cap K = \emptyset$. This implies that $x_0 \notin K + W$, for otherwise there would exist some $k \in K$ and $w \in W$ such that $x_0 - w = k$, contradicting the fact that the intersection of $x_0 + W$ with K is empty. (Here we use the fact that $W = -W$, because W is balanced.)

Let $V = K + \frac{1}{2}W$. Then V is a convex neighborhood of 0. Define a function $p : E \to \mathbb{R}$ by

$$p(x) = \inf\{\lambda > 0 : x \in \lambda V\}, \quad x \in E.$$

Recall that every neighborhood of 0 is absorbent. In particular V is absorbent, and so $p(x) < \infty$ for all $x \in E$. We claim p is sublinear. For any $x \in E$ and $\alpha \geq 0$,

$$p(\alpha x) = \inf\{\lambda > 0 : \alpha x \in \lambda V\} = \alpha \inf\left\{\frac{\lambda}{\alpha} > 0 : x \in \frac{\lambda}{\alpha}V\right\} = \alpha p(x).$$

This proves positive homogeneity. It remains to show that p is subadditive.

Let x and y be in E and let $\epsilon > 0$. Because $p(x)$ and $p(y)$ are infima, there exist real numbers $\lambda > 0$ and $\mu > 0$ such that $p(x) < \lambda < p(x) + \frac{\epsilon}{2}$ and $p(y) < \mu < p(y) + \frac{\epsilon}{2}$. By the definition of p, we have that $\frac{x}{\lambda} \in V$ and $\frac{y}{\mu} \in V$. By the convexity of V,

$$\frac{x+y}{\lambda + \mu} = \frac{\lambda}{\lambda + \mu}\left(\frac{x}{\lambda}\right) + \frac{\mu}{\lambda + \mu}\left(\frac{y}{\mu}\right) \in V.$$

Therefore,

$$p(x + y) \leq \lambda + \mu < p(x) + p(y) + \epsilon.$$

The choice of ϵ was arbitrary, and so $p(x+y) \leq p(x) + p(y)$, as required. Therefore p is sublinear.

By definition, $p(x) \leq 1$ for all $x \in V$. We now show $p(x_0) > 1$. Suppose to the contrary that $p(x_0) \leq 1$. It follows that $\frac{x_0}{\lambda} \in V$ for all $\lambda \geq 1$. Since $\frac{x_0}{\lambda} \to x_0$ as $\lambda \to 1$, we conclude that $x_0 \in \overline{V}$, and consequently $(x_0 + \frac{1}{2}W) \cap V \neq \emptyset$.

Recall that $V = K + \frac{1}{2}W$. Thus, $(x_0 + \frac{1}{2}W) \cap (K + \frac{1}{2}W) \neq \emptyset$, and so there exists an element $k \in K$ and elements w_1 and w_2 in W such that $x_0 + \frac{1}{2}w_1 = k + \frac{1}{2}w_2$. Hence,

$$x_0 = k + \frac{1}{2}w_2 - \frac{1}{2}w_1 \in K + \frac{1}{2}W - \frac{1}{2}W.$$

Because W is absolutely convex, we have that $\frac{1}{2}W - \frac{1}{2}W \subseteq W$. From this we conclude that $x_0 \in K + W$. This is a contradiction, and so it must be that $p(x_0) > 1$.

We now make use of Exercise 3.9. There exists a linear functional f on E such that $f \leq p$ and $f(x_0) > 1$. Because $K \subseteq V$, and because $p(x) \leq 1$ for all $x \in V$, we have

$$\sup_{y \in K} f(y) \leq 1 < f(x_0).$$

It remains to show that f is continuous. We will demonstrate this by showing that f is bounded on some neighborhood of zero and applying Proposition 5.13. Since

$0 \in K$, we have that $\frac{1}{2} W \subseteq K + \frac{1}{2} W = V$. By construction, $f(x) \leq p(x) \leq 1$ for all $x \in V$, and hence $f(x) \leq 1$ for all $x \in \frac{1}{2} W$. The set W is balanced, and thus $|f(x)| \leq 1$ for all $x \in \frac{1}{2} W$. Therefore, we have demonstrated that $f(\frac{1}{2} W) \subseteq [-1, 1]$. Consequently, the linear functional f is continuous, by Proposition 5.13.   □

Example 5.21 Suppose E is a real locally convex topological vector space and K is a closed linear subspace of E. If $x_0 \notin K$, then, by Theorem 5.20, there exists a continuous linear functional f on E such that $f(K) = 0$ and $f(x_0) > 0$. (See Exercise 5.20.)

There is also a version of Theorem 5.20 for complex topological vector spaces.

Theorem 5.22 *Let E be a complex locally convex topological vector space. Let K be a closed nonempty convex subset of E. If $x_0 \notin K$, then there exists a continuous linear functional f on E such that*

$$\Re(f(x_0)) > \sup_{x \in K} \Re(f(x)).$$

Proof Ignoring multiplication by complex scalars, we may treat E as a vector space over $\mathbb{R}$. Therefore, by Theorem 5.20, there exists a real linear functional g on E such that $g(x_0) > \sup_{x \in K} g(x)$. Now, define a complex linear functional on E by $f(x) = g(x) - i g(ix)$ for all $x \in E$. The functional f is the desired continuous linear functional on E.   □

Definition 5.23 Let X be a vector space and let $\mathbb{K}$ denote the scalar field. A function $p : X \to \mathbb{R}$ is called a *semi-norm* if the following three conditions are satisfied:

 (i) $p(x) \geq 0$ for all $x \in X$,
 (ii) $p(x + y) \leq p(x) + p(y)$ for all $\{x, y\} \subseteq X$, and
 (iii) $p(\alpha x) = |\alpha| \, p(x)$ for all $\alpha \in \mathbb{K}$ and $x \in X$.

What distinguishes a semi-norm from a norm is that a semi-norm p may satisfy $p(x) = 0$ even when $x \neq 0$. As in the case of a norm, we call the property in *(ii) subadditivity* (or the *triangle inequality*) and we call the property in *(iii) homogeneity*.

Theorem 5.24 *Suppose $\{p_\alpha\}_{\alpha \in A}$ is a family of semi-norms on a vector space X.*
   Let
$$V(\alpha, n) = \{x : p_\alpha(x) < 1/n\}, \quad \alpha \in A, \ n \in \mathbb{N}.$$

If η is the collection of all finite intersections of the sets $V(\alpha, n)$, where $\alpha \in A$ and $n \in \mathbb{N}$, then η determines a locally convex vector topology on X in which the elements of η form an absolutely convex base of neighborhoods at 0.

Proof We define a topology on X by declaring a set $E \subseteq X$ to be open if and only if E is a (possibly empty) union of translates of elements in η. This defines a topology for which all members of η are absolutely convex (that is, convex and balanced).

It remains to show that addition and scalar multiplication are continuous. Let U be an open neighborhood of 0 in X. Without loss of generality, we may assume U is

an element of η. Thus,

$$U = V(\alpha_1, n_1) \cap \cdots \cap V(\alpha_k, n_k) \tag{5.2}$$

for $\{\alpha_1, \ldots, \alpha_k\} \subseteq A$ and $\{n_1, \ldots, n_k\} \subseteq \mathbb{N}$. If $V = V(\alpha_1, 2n_1) \cap \cdots \cap V(\alpha_k, 2n_k)$, then $V + V \subseteq U$ (because p_α is subadditive for every $\alpha \in A$). Therefore, addition is continuous.

Now, let $x \in X$ and $\kappa \in \mathbb{K}$, where $\mathbb{K}$ is the scalar field. A basic open neighborhood of κx can be written as $\kappa x + U$, where U is written as in (5.2). We will show there exists an open neighborhood W of x and a $\delta > 0$ such that $\lambda W \subseteq \kappa x + U$ for all $|\kappa - \lambda| < \delta$.

Let $V = V(\alpha_1, 2n_1) \cap \cdots \cap V(\alpha_k, 2n_k)$, as above. Since V is an open neighborhood of 0, it is absorbent. Thus, there exists some $n \in \mathbb{N}$ such that $x \in nV$. Let

$$\delta = \frac{1}{n} \quad \text{and} \quad W = x + \frac{n}{1 + |\kappa|n} V.$$

Suppose $w \in W$ and $\lambda \in B(\kappa, \delta)$. Then

$$\kappa x - \lambda w = (\kappa - \lambda)x + \lambda(x - w).$$

Observe that $x = nv_1$ and $w - x = \frac{n}{1+|\kappa|n} v_2$ for some choice of v_1 and v_2 in V. Hence,

$$\kappa x - \lambda w = (\kappa - \lambda)nv_1 - \frac{\lambda n}{1 + |\kappa|n} v_2.$$

Therefore, because V is balanced,

$$\kappa x - \lambda w \in |\kappa - \lambda| nV + \frac{|\lambda|n}{1 + |\kappa|n} V \subseteq V + V \subseteq U.$$

It follows that scalar multiplication is continuous, and so the proof is complete. $\square$

Definition 5.25 Suppose X is a topological vector space and let U be an absorbent subset of X. The *Minkowski functional* of U on X is the function $p_U : X \to \mathbb{R}$ defined by

$$p_U(x) = \inf\{\lambda > 0 : x \in \lambda U\}, \quad x \in X.$$

Note that $p_U(x) < \infty$ for all $x \in X$, because U is absorbent.

Suppose that X is a locally convex topological vector space. Then X has a base of neighborhoods of 0 that are absolutely convex. Such sets are absorbent, and so each such set will give rise to a well-defined Minkowski functional.

Proposition 5.26 *Let X be a topological vector space and let U be an absorbent absolutely convex subset of X. The Minkowski functional p_U is a semi-norm on X.*

Proof Certainly $p_U(x) \geq 0$ for each $x \in X$, by the definition of p_U.

To show the subadditivity of p_U, we will use the convexity of U. Let x and y be elements in X and let $\epsilon > 0$. By the definition of p_U, there exist numbers $\lambda_1 > 0$

and $\lambda_2 > 0$ such that $p_U(x) < \lambda_1 < p_U(x) + \frac{\epsilon}{2}$ and $p_U(y) < \lambda_2 < p_U(y) + \frac{\epsilon}{2}$. It is necessarily the case that x/λ_1 and y/λ_2 are both in U. By the convexity of U,

$$\frac{x+y}{\lambda_1 + \lambda_2} = \frac{\lambda_1}{\lambda_1 + \lambda_2}\left(\frac{x}{\lambda_1}\right) + \frac{\lambda_2}{\lambda_1 + \lambda_2}\left(\frac{y}{\lambda_2}\right) \in U.$$

Therefore,

$$p_U(x+y) \leq \lambda_1 + \lambda_2 < p_U(x) + p_U(y) + \epsilon.$$

The choice of ϵ was arbitrary, and so $p_U(x+y) \leq p_U(x) + p_U(y)$. (Compare to the proof of Theorem 5.20.)

Finally, we show homogeneity. Let $\alpha \in \mathbb{K}$, where $\mathbb{K}$ is the field of scalars. Computing directly, we have

$$p_U(\alpha x) = \inf\{\lambda > 0 : \alpha x \in \lambda U\} = |\alpha| \inf\left\{\frac{\lambda}{|\alpha|} > 0 \; : \; x \in \frac{\lambda}{|\alpha|} \cdot \text{sign}(\alpha)U\right\}.$$

Since U is balanced, $\text{sign}(\alpha)U = U$. Letting $\lambda' = \lambda/|\alpha|$,

$$p_U(\alpha x) = |\alpha| \inf\left\{\lambda' > 0 : x \in \lambda' U\right\} = |\alpha|\, p_U(x).$$

Therefore, p_U is a semi-norm on X, as claimed.   □

If X is a locally convex topological vector space, then there exists a base of absolutely convex neigborhoods of 0, say η. By Proposition 5.26, the Minkowski functional p_U is a semi-norm on X for each $U \in \eta$. By Theorem 5.24, the family of semi-norms $\{p_U\}_{U \in \eta}$ generates a locally convex vector topology on X. We leave it as an exercise to show that the topology generated by $\{p_U\}_{U \in \eta}$ is, in fact, the original topology. (See Exercise 5.17.)

So far, we have considered general topological vector spaces. We now focus our attention on topological vector spaces that have a complete norm structure—that is, Banach spaces. We have already said much about the norm topology of a Banach space X. We now consider a new topology on X, the so-called *weak topology*.

Definition 5.27 Let X be a topological vector space. The *weak topology* on X (or the *w-topology*) is defined by a base of neighborhoods at 0 of the form

$$W(x_1^*, \ldots, x_n^*; \epsilon) = \{x : |x_i^*(x)| < \epsilon, \; 1 \leq i \leq n\},$$

where $\epsilon > 0$ and $\{x_1^*, \ldots, x_n^*\} \subseteq X^*$ for $n \in \mathbb{N}$.

The weak topology on X is the topology it inherits as a subspace of the space $\mathbb{K}^{X^*}$ with the product topology. The space $\mathbb{K}^{X^*}$ is the collection of all functions from X^* into the scalar field $\mathbb{K}$, and we identify X with a subspace of $\mathbb{K}^{X^*}$ by identifying $x \in X$ with $\hat{x} \in \mathbb{K}^{X^*}$ via the relationship $\hat{x}(x^*) = x^*(x)$ for all $x^* \in X^*$.

To distinguish between the norm and weak topologies on X, we will frequently denote X with the norm topology by $(X, \|\cdot\|)$ and X with the weak topology by (X, w). The weak and norm topologies are generally quite different. Any weakly open set is necessarily open in the norm topology (the basic sets are intersections of

preimages of open sets under continuous maps), but not every set open in the norm topology will be weakly open. (We will demonstrate this shortly.)

The weak topology on X generally has fewer open sets, and so it is "harder" for a function on (X, w) to be continuous than a function on $(X, \|\cdot\|)$. For example, consider the identity map Id_X on X. The map $\mathrm{Id}_X : (X, \|\cdot\|) \to (X, w)$ is always continuous, but $\mathrm{Id}_X : (X, w) \to (X, \|\cdot\|)$ need not be. Indeed, if both maps are continuous, then the topologies must coincide, and then X must be finite-dimensional. (See Proposition 5.30.)

Let us consider which sequences converge in X with the weak topology. Without loss of generality, we may consider only those sequences converging to 0. If a sequence $(x_n)_{n=1}^\infty$ converges to 0 in the weak topology on X, we say that $(x_n)_{n=1}^\infty$ *converges weakly* to 0 (or $x_n \to 0$ *weakly*). The sequence $(x_n)_{n=1}^\infty$ converges weakly to 0 precisely when it converges coordinate-wise to 0 in $\mathbb{K}^{X^*}$. That is to say, $x_n \to 0$ weakly if and only if

$$\lim_{n \to \infty} x^*(x_n) = 0, \quad \text{for all } x^* \in X^*.$$

In other words, x_n converges to 0 in the weak topology if and only if every weak neighborhood of the origin eventually contains the sequence $(x_n)_{n=1}^\infty$.

A sequence converges to 0 in the norm topology if and only if every "strong" neighborhood of the origin eventually contains the sequence. However, the norm topology has more open neighborhoods about 0 than the weak topology. Consequently, it is more "difficult" for a sequence to converge in the norm topology than to converge in the weak topology.

Example 5.28 Consider ℓ_p for $1 \le p < \infty$. For each $n \in \mathbb{N}$, let e_n be the sequence with 1 in the nth coordinate, and 0 elsewhere. If m and n are elements of $\mathbb{N}$ such that $m \neq n$, then $\|e_m - e_n\|_{\ell_p} = 2^{1/p}$. Consequently, the sequence $(e_n)_{n=1}^\infty$ does not converge in the norm topology. On the other hand, if $x^* = (x_n^*)_{n=1}^\infty$ is a sequence in $(\ell_p)^* = \ell_q$, where $p > 1$ and q is the exponent conjugate to p, then

$$\lim_{n \to \infty} x^*(e_n) = x_n^* = 0.$$

Since this is true for all $x^* \in \ell_q$, we conclude that $e_n \to 0$ weakly.

The above conclusion does not remain true when $p = 1$. In this case, $q = \infty$. Let $e = (1, 1, 1, \dots)$ be the constant sequence with all terms equal to 1. This sequence is bounded, and so $e \in \ell_\infty = (\ell_1)^*$. For each $n \in \mathbb{N}$, we have that $e(e_n) = 1$, and so $e_n \not\to 0$ in the weak topology in this case.

Example 5.29 Consider the Banach space $L_p(\mathbb{T})$ of p-integrable complex-valued functions on the torus $\mathbb{T} = [0, 2\pi)$, where $1 \le p < \infty$. For each $n \in \mathbb{N}$, define a function $f_n : \mathbb{T} \to \mathbb{C}$ by $f_n(\theta) = e^{in\theta}$, where $\theta \in \mathbb{T}$. Let $\Lambda \in L_p(\mathbb{T})^*$. By duality, there exists some $g \in L_q(\mathbb{T})$, where $1/p + 1/q = 1$, such that

$$\Lambda(f) = \int_\mathbb{T} f(\theta)\, g(\theta)\, \frac{d\theta}{2\pi}, \quad f \in L_p(\mathbb{T}).$$

Therefore,

$$\lim_{n\to\infty} \Lambda(f_n) = \lim_{n\to\infty} \left(\int_{\mathbb{T}} e^{in\theta} \, g(\theta) \, \frac{d\theta}{2\pi} \right) = \lim_{n\to\infty} \hat{g}(-n) = 0.$$

This last equality follows from the Riemann–Lebesgue Lemma (Theorem 4.37). Therefore, $\lim_{n\to\infty} \Lambda(f_n) = 0$ for all $\Lambda \in L_q(\mathbb{T})$, and so $f_n \to 0$ weakly. However, $\|f_n\|_{L_p(\mathbb{T})} = 1$ for all $n \in \mathbb{N}$, and so $f_n \nrightarrow 0$ in the norm topology.

If X is a finite-dimensional Banach space, then all linear functionals are continuous.

Proposition 5.30 *Let X be a Banach space. The following are equivalent:*

  (i) *$\dim(X) < \infty$,*
 (ii) *the weak topology on X coincides with the norm topology on X, and*
(iii) *the weak topology on X is metrizable.*

Proof The implications *(i)* $\Rightarrow$ *(ii)* $\Rightarrow$ *(iii)* are clear. It remains to show *(iii)* $\Rightarrow$ *(i)*. Assume the weak topology on X is metrizable.

Then (X, w) is first countable, and so there exists a weak base of neighborhoods $(W_n)_{n=1}^\infty$ at the origin of the form

$$W_n = \{x : |x_{n,j}^*(x)| \le \epsilon_n, \ \ 1 \le j \le N_n\},$$

where $x_{n,j}^* \in X^*$, $\epsilon_n > 0$, and $N_n \in \mathbb{N}$, for all $n \in \mathbb{N}$ and all $j \in \{1, \dots, N_n\}$.

For each $n \in \mathbb{N}$, define

$$E_n = \mathrm{span}\{x_{n,j}^* \ : \ 1 \le j \le N_n\}.$$

Fix some $x^* \in X^*$. The set $\{x : |x^*(x)| \le 1\}$ is a weak neighborhood of 0 in X, and consequently must contain W_n for some $n \in \mathbb{N}$. For this fixed n, define a linear map $T : X \to \mathbb{K}^{N_n}$, where $\mathbb{K}$ is the scalar field, by

$$T(x) = (x_{n,1}^*(x), \dots, x_{n,N_n}^*(x)), \quad x \in X.$$

We claim $x^* \in (\ker T)^\perp$. (Recall Definition 3.50.) To verify this, suppose $y \in \ker(T)$. By the definition of T, we have that $x_{n,j}^*(y) = 0$ for all $j \in \{1, \dots, N_n\}$. Naturally, if $\lambda \in \mathbb{K}$, then it follows that $x_{n,j}^*(\lambda y) = 0$ for all $j \in \{1, \dots, N_n\}$. Consequently, $\lambda y \in W_n$ for all $\lambda \in \mathbb{K}$. By design, $W_n \subseteq \{x : |x^*(x)| \le 1\}$, and so $|x^*(\lambda y)| \le 1$ for all $\lambda \in \mathbb{K}$. This can occur only if $|x^*(y)| \le 1/\lambda$ for all $\lambda \in \mathbb{K}$, and thus $x^*(y) = 0$. This remains true for any $y \in \ker(T)$, and so we have that $x^* \in (\ker T)^\perp$.

By Lemma 4.33, there then exists some $f \in (\mathbb{K}^{N_n})^*$ such that $x^*(x) = (f \circ T)(x)$ for all $x \in X$. Since $\mathbb{K}^{N_n}$ is finite-dimensional, there exists a finite sequence $(a_j)_{j=1}^{N_n}$

such that

$$f(\xi_1, \ldots, \xi_{N_n}) = \sum_{j=1}^{N_n} a_j \xi_j, \quad (\xi_j)_{j=1}^{N_n} \in \mathbb{K}^{N_n}.$$

Therefore,

$$x^*(x) = f(x_{n,1}^*(x), \ldots, x_{n,N_n}^*(x)) = \sum_{j=1}^{N_n} a_j x_{n,j}^*(x), \quad x \in X,$$

and so $x^* \in E_n$.

We have shown that each $x^* \in X^*$ is in E_n for some $n \in \mathbb{N}$. We therefore conclude that $X^* = \bigcup_{n=1}^{\infty} E_n$. For each $n \in \mathbb{N}$, the space E_n is finite-dimensional, and so is closed. Therefore, by Theorem 4.7 (the complementary version of the Baire Category Theorem), there exists some $n \in \mathbb{N}$ such that $\mathrm{int}(E_n) \neq \emptyset$. We conclude that E_n is an open neighborhood of the origin in X^*, and consequently is absorbent. Therefore, $X^* = \bigcup_{k=1}^{\infty} k E_n = E_n$. Thus, the space X^* is finite-dimensional, and so X is finite-dimensional, as well. $\qquad\square$

Proposition 5.31 *Let X be a Banach space. Then:*

 (i) *The weak topology on X is a Hausdorff topology.*
(ii) *A linear functional is continuous in the weak topology if and only if it is continuous in the norm topology.*

Proof (i) Assume x_1 and x_2 are elements in X such that $x_1 \neq x_2$. By the Hahn–Banach Separation Theorem (Theorem 5.20), there exists an $x^* \in X^*$ such that $\epsilon = x^*(x_2 - x_1) > 0$. Therefore, the set $\{x : |x^*(x) - x^*(x_1)| < \epsilon/2\}$ is a weak neighborhood of x_1, the set $\{x : |x^*(x) - x^*(x_2)| < \epsilon/2\}$ is a weak neighborhood of x_2, and these two neighborhoods are disjoint. Hence, (X, w) is a Hausdorff topological space.

(ii) If a linear functional f is continuous in the weak topology on X, then $f^{-1}(V)$ is a weakly open set whenever V is an open set in the scalar field. But the norm topology contains all of the weakly open sets, so $f^{-1}(V)$ is open in the norm topology. Therefore, f is continuous in the norm topology on X. (The idea is that it is "easier" to be continuous in the norm topology, because there are more open sets.)

Now, suppose f is a norm continuous linear functional. Then $f \in X^*$, and so the set $\{x : |f(x)| < \epsilon\}$ is a weak neighborhood of 0 (by the definition of the weak topology). Thus, f is continuous in the weak topology on X. $\qquad\square$

The weak topology on X is the *weakest* topology on X such that all norm continuous linear functionals remain continuous. When we say a topology is *weaker*, we mean that it contains fewer open sets. The norm topology on X is *stronger* than the weak topology on X, because it contains more open sets. Weakly open sets are open in the norm topology, but the converse need not be true. A function is continuous if the preimage of any open set is open. The stronger the topology on the domain, the easier it is for a function to be continuous, because with more open sets, it is more likely that a given preimage is open.

Definition 5.32 Let X be a topological vector space. The *weak* topology* on X^* (or the *w^*-topology*) is defined by a base of neighborhoods at 0 of the form

$$W^*(x_1, \dots, x_n; \epsilon) = \{x^* : |x^*(x_i)| < \epsilon, \ 1 \le i \le n\},$$

where $\epsilon > 0$ and $\{x_1, \dots, x_n\} \subseteq X$ for $n \in \mathbb{N}$.

The weak* topology on X^* is the topology inherited from viewing X^* as a subspace of $\mathbb{K}^X$, the space of all scalar-valued functions on X. As before, we endow $\mathbb{K}^X$ with the product topology. We use (X^*, w^*) to denote X^* with the weak* topology.

Observe that any $x \in X$ can be thought of as a linear functional on X^* via the mapping $x \mapsto \phi_x$, where $\phi_x(x^*) = x^*(x)$ for all $x^* \in X^*$. The weak* topology on X^* is the weakest topology on X^* for which the linear functionals ϕ_x are continuous for all $x \in X$.

The Banach space X^* has also a weak topology that is induced by it's dual space $(X^*)^* = X^{**}$ (the *bidual* of X). The weak* topology on X^* is weaker than the weak topology on X^*, because it requires fewer members in X^{**} to be continuous. (Only those coming from X.)

Example 5.33 Consider the sequence space ℓ_1. In Example 5.28, we saw that ℓ_1 had a weak topology induced upon it by $\ell_1^* = \ell_\infty$. In this weak topology, we saw that the sequence $(e_n)_{n=1}^\infty$ did not converge to 0 (because $e(e_n) = 1$ for all $n \in \mathbb{N}$, where $e = (1, 1, \dots)$ is the constant sequence with all terms equal to 1). The space ℓ_1 can also be given a weak* topology as the dual space of c_0.

Suppose $\xi = (\xi_k)_{k=1}^\infty$ is an element of c_0. Since c_0 consists of sequences that converge to 0, it follows that $e_n(\xi) = \xi_n \to 0$ as $n \to \infty$. This is true for every $\xi \in c_0$, and so the sequence $(e_n)_{n=1}^\infty$ converges to 0 in the weak* topology on ℓ_1.

In this example we have found a sequence which converges in the weak* topology on ℓ_1, but not in the weak topology on ℓ_1. This happens because the weak* topology has fewer open sets than the weak topology. (That is to say, the weak* topology is weaker than the weak topology).

Proposition 5.34 *Let X be a Banach space. Then:*

 (i) *The weak* topology on X^* is a Hausdorff topology.*
 (ii) *A linear functional f on X^* is weak*-continuous if and only if there exists some $x \in X$ such that $f(x^*) = x^*(x)$ for all $x^* \in X^*$. (In other words, $(X^*, w^*)^* = X$.)*

Proof (i) Let x_1^* and x_2^* be elements in X^* such that $x_1^* \neq x_2^*$. Then there exists some $x \in X$ such that $x^*(x) \neq x_2^*(x)$. (Otherwise they would be the same as linear functionals on X.) If $\epsilon = |(x_1^* - x_2^*)(x)|$, then the sets $\{x^* : |x^*(x) - x_1^*(x)| < \epsilon/2\}$ and $\{x^* : |x^*(x) - x_2^*(x)| < \epsilon/2\}$ are disjoint weak*-open sets containing x_1^* and x_2^*, respectively.

(ii) Certainly, if $f(x^*) = x^*(x)$ for all $x^* \in X^*$, then f is continuous in the weak* topology. It remains only to show that any weak*-continuous linear functional on X^* can be achieved in this way.

   Assume f is a continuous linear functional on (X^*, w^*). By Proposition 5.13, there exists a basic neighborhood of (X^*, w^*) on which f is bounded. Thus, there exists a real number $\epsilon > 0$ and a finite set $\{x_1, \ldots, x_n\} \subseteq X$ such that $|f(x^*)| \leq 1$ for all $x^* \in W^*(x_1, \ldots, x_n; \epsilon)$.

   Define a map $T : X^* \to \mathbb{K}^n$, where $\mathbb{K}$ is the scalar field, by

$$T(x^*) = (x^*(x_1), \ldots, x^*(x_n)), \quad x^* \in X^*.$$

Suppose $x^* \in \ker(T)$. Then $x^*(x_j) = 0$ for each $j \in \{1, \ldots, n\}$. Thus, for any $\lambda \in \mathbb{K}$, we have that $x^*(\lambda x_j) = 0$ for $j \in \{1, \ldots, n\}$, and so $(\lambda x^*) \in W^*(x_1, \ldots, x_n; \epsilon)$. It follows that $|f(\lambda x^*)| \leq 1$, and consequently $|f(x^*)| \leq 1/|\lambda|$ for all $\lambda \neq 0$. From this we conclude that $f(x^*) = 0$, and hence $f \in (\ker T)^\perp$. By Lemma 4.33, then, there exists a bounded linear functional $\phi : \mathbb{K}^n \to \mathbb{K}$ such that $f = \phi \circ T$. Therefore, there exists a finite collection of scalars $\{a_1, \ldots, a_n\} \subseteq \mathbb{K}$ such that

$$f(x^*) = \phi(Tx) = \phi(x^*(x_1), \ldots, x^*(x_n)) = \sum_{j=1}^{n} a_j\, x^*(x_j), \quad x^* \in X^*.$$

The desired element of X is $x = \displaystyle\sum_{j=1}^{n} a_j\, x_j$. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remark 5.3 In Example 5.33 we saw that the weak* topology may be strictly weaker than the weak topology. If X is a *reflexive* space (recall Definition 3.33), however, then the weak and weak* topologies coincide.

   Shortly, we will prove Proposition 5.37 which (in some sense) demonstrates that it is "hard" to be compact in a normed space. Before we state and prove this proposition, however, we need a lemma, which is of independent interest.

Lemma 5.36 *All norms on a finite-dimensional vector space are equivalent.*

Proof Let X be a finite-dimensional vector space over the scalar field $\mathbb{K}$. Choose $x_1, \ldots, x_n$ in X so that $X = \mathrm{span}\{x_1, \ldots, x_n\}$. We recall that each element of X has a unique representation of the form $\displaystyle\sum_{i=1}^{n} \alpha_i x_i$, where $\alpha_i \in \mathbb{K}$ for each $i \in \{1, \ldots, n\}$. Define a norm $|||\cdot|||$ on X as follows:

$$\left|\left|\left|\sum_{i=1}^{n} \alpha_i x_i\right|\right|\right| = \sum_{i=1}^{n} |\alpha_i|.$$

It is straightforward to show that this does indeed define a norm on X.

   Now, let $\|\cdot\|$ be another norm on X. We will find positive constants c and C such that $c|||x||| \leq \|x\| \leq C|||x|||$ for all $x \in X$. By the triangle inequality,

$$\left\|\sum_{i=1}^{n} \alpha_i x_i\right\| \leq \sum_{i=1}^{n} |\alpha_i| \cdot \|x_i\| \leq (\max_i \|x_i\|)\left(\sum_{i=1}^{n} |\alpha_i|\right) = (\max_i \|x_i\|)\left|\left|\left|\sum_{i=1}^{n} \alpha_i x_i\right|\right|\right|.$$

$$(5.3)$$

Thus, we may choose $C = \max_i \|x_i\|$.

Next, define a set

$$S = \Big\{(\alpha_1, \ldots, \alpha_n) : \sum_{i=1}^{n} |\alpha_i| = 1\Big\}.$$

Observe that S is a closed and bounded subset of $\mathbb{K}^n$. Therefore, S is compact by the Heine–Borel Theorem. Define a function $f : S \to \mathbb{R}^+$ by

$$f(\alpha_1, \ldots, \alpha_n) = \Big\| \sum_{i=1}^{n} \alpha_i x_i \Big\|.$$

We claim that the function f is continuous. To see this, observe that

$$\Big| f(\alpha_1, \ldots, \alpha_n) - f(\beta_1, \ldots, \beta_n) \Big| = \Big| \|\sum_{i=1}^{n} \alpha_i x_i\| - \|\sum_{i=1}^{n} \beta_i x_i\| \Big| \le \Big\| \sum_{i=1}^{n} \alpha_i x_i - \sum_{i=1}^{n} \beta_i x_i \Big\|$$

$$= \Big\| \sum_{i=1}^{n} (\alpha_i - \beta_i) x_i \Big\| \le \sum_{i=1}^{n} |\alpha_i - \beta_i| \|x_i\| \le \Big(\sum_{i=1}^{n} |\alpha_i - \beta_i|^2 \Big)^{1/2} \Big(\sum_{i=1}^{n} \|x_i\|^2 \Big)^{1/2}.$$

The last inequality follows from the Cauchy–Schwarz Inequality. From this, it follows that f is continuous.

By the Extreme Value Theorem, since f is continuous on a compact set, the function f attains a minimum value on the set S. Let c be that minimum value. Then $f(\alpha_1, \ldots, \alpha_n) \ge c$ for all $(\alpha_1, \ldots, \alpha_n)$ in S. This means that $\Big\| \sum_{i=1}^{n} \alpha_i x_i \Big\| \ge c$ for all $(\alpha_1, \ldots, \alpha_n)$ in $\mathbb{K}^n$ such that $\sum_{i=1}^{n} |\alpha_i| = 1$. Alternately, for any $(\alpha_1, \ldots, \alpha_n)$ in $\mathbb{K}^n$,

$$\Big\| \sum_{i=1}^{n} \alpha_i x_i \Big\| \ge c \sum_{i=1}^{n} |\alpha_i| = c \Big| \Big\| \sum_{i=1}^{n} \alpha_i x_i \Big\| \Big|. \tag{5.4}$$

Combining (5.3) and (5.4), we conclude that the two norms are equivalent. Since the norm $\| \cdot \|$ was arbitrary, it follows that all norms on X are equivalent. $\qquad \square$

Proposition 5.37 *Suppose X is a Banach space (or just a normed linear space). Then B_X is compact in the norm topology on X if and only if $\dim(X) < \infty$.*

Proof Suppse X is a finite-dimensional normed vector space. By Lemma 5.36, X is homeomorphic to $\mathbb{K}^n$ with the Euclidean norm. Therefore, B_X is compact by the Heine–Borel Theorem.

Next, suppose B_X is compact in the norm topology on X. Denote by $B(x,r)$ the open ball of radius r centered at $x \in X$. Since B_X is compact, there exists a finite sequence $\{x_1, \ldots, x_k\}$ of elements in X, such that

$$B_X \subseteq \bigcup_{j=1}^{k} B\left(x_j, \frac{1}{2}\right) = \bigcup_{j=1}^{k} \left(x_j + \frac{1}{2} B_X\right). \tag{5.5}$$

Let $F = \mathrm{span}\{x_1, \ldots, x_k\}$. Then (5.5) implies that $B_X \subseteq F + \frac{1}{2} B_X$. This is a recursive statement, and so we apply it to itself to get

$$B_X \subseteq F + \frac{1}{2}\left(F + \frac{1}{2} B_X\right) = F + \frac{1}{2} F + \frac{1}{4} B_X = F + \frac{1}{4} B_X.$$

Continuing recursively, we have $B_X \subseteq F + \frac{1}{2^n} B_X$ for all $n \in \mathbb{N}$. Therefore,

$$B_X \subseteq \bigcap_{n=1}^{\infty} \left(F + \frac{1}{2^n} B_X\right).$$

However, F is closed, as a consequence of Lemma 5.36 (because F is finite-dimensional). Thus,

$$\bigcap_{n=1}^{\infty} \left(F + \frac{1}{2^n} B_X\right) = F,$$

and so $B_X \subseteq F$. Since B_X is absorbent, we have $X = \bigcup_{n=1}^{\infty} n B_X \subseteq \bigcup_{n=1}^{\infty} n F$. But F is a vector space, and so $X \subseteq F$. Therefore, $X = F$, as required. □

While the unit ball in a Banach space can be compact in the norm topology only if the space is finite-dimensional, the unit ball in the weak* topology will always be compact. Before proving this statement, known as the Banach-Alaoglu Theorem, let us recall a theorem from general topology.

Theorem 5.38 (Tychonoff's Theorem) *Let I be an arbitrary index set. If $\{K_i\}_{i \in I}$ is a collection of compact topological spaces, then $\prod_{i \in I} K_i$ is compact in the product topology.*

We will not prove this theorem, but we do wish to point out it relies on the Axiom of Choice. We are now ready to state and prove the Banach-Alaoglu Theorem.

Theorem 5.39   (Banach-Alaoglu Theorem) *If X is a Banach space, then B_{X^*} is compact in the weak* topology on X^*.*

Proof Let X be a Banach space over the scalar field $\mathbb{K}$. Recall that X^* in the weak* topology is achieved by viewing X^* as a subspace of $\mathbb{K}^X = \prod_{x \in X} \mathbb{K}$ in the product topology. We make this explicit by defining the map $\phi : X^* \to \mathbb{K}^X$ by

$$\phi(x^*) = (x^*(x))_{x \in X}, \quad x^* \in X^*.$$

If $x^* \in B_{X^*}$, then for each $x \in X$, we have $|x^*(x)| \leq \|x\|$. Consequently,

$$\phi(B_{X^*}) \subseteq \prod_{x \in X} \|x\| B_{\mathbb{K}},$$

where $B_{\mathbb{K}}$ is the closed unit ball in the scalar field $\mathbb{K}$ and $\|x\|B_{\mathbb{K}}$ is the closed ball of radius $\|x\|$ centered at the origin. The product $A = \prod_{x \in X} \|x\| B_{\mathbb{K}}$ is compact, by Tychonoff's Theorem. There is no reason the image of B_{X^*} would be all of A, but it is a closed subspace. Indeed, the image is precisely the collection of elements in the following set:

$$\bigcap_{\substack{\{\alpha_1, \alpha_2\} \subseteq \mathbb{R} \\ \{x_1, x_2\} \subseteq X}} \left\{ f : f(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 f(x_1) + \alpha_2 f(x_2) \right\} \bigcap \prod_{x \in X} \|x\| B_{\mathbb{K}}.$$

(The first set of relations ensures $f \in \mathbb{K}^X$ is linear, while the second ensures it is bounded.) Therefore, $\phi(B_{X^*})$ is a closed subset of the compact set A, and hence $\phi(B_{X^*})$ is compact in the product topology on $\mathbb{K}^X$. It follows that B_{X^*} is compact in the weak* topology on X^*, as required.   □

The Banach–Alaoglu Theorem as given here is due to Leonidas Alaoglu [1], although the result was known to Banach. Banach did not have the notions of general topology available to him, and so he could not formulate it in this way.

5.5 Goldstine's Theorem

Let X be a Banach space. Recall that X can be thought of as a subspace of it's bidual X^{**}. The space X^{**} is the dual space for X^*, and as such can be given a weak* topology. The weak* topology on X^{**} is the weakest topology under which elements of X^* define continuous functions on X^{**}. If we restrict to the subspace X, then the weakest topology under which elements of X^* are continuous is the weak topology on X. Therefore

$$(X^{**}, w^*)|_X = (X, w).$$

In other words, the restriction of the weak* topology on X^{**} to X is the weak topology on X.

Theorem 5.40 (Goldstine's Theorem) *If X is a Banach space, then B_X is weak*-dense in $B_{X^{**}}$.*

Proof Let X be a Banach space. For simplicity, we will assume X is real. (If X is complex, the argument is similar.) Denote the closure of B_X in the weak* topology on X^{**} by $\overline{B_X}^{(w^*)}$. Our goal is to show the equality $\overline{B_X}^{(w^*)} = B_{X^{**}}$.

By the Banach–Alaoglu Theorem (Theorem 5.39), the set $B_{X^{**}}$ is a compact (and hence closed) set in the weak* topology on X^{**}. Therefore, since $X \subseteq X^{**}$, we see that $\overline{B_X}^{(w^*)} \subseteq B_{X^{**}}$.

Suppose $x_0^{**} \in B_{X^{**}} \setminus \overline{B_X}^{(w^*)}$. By the Hahn–Banach Separation Theorem (Theorem 5.20), there exists a weak*-continuous linear functional f on X^{**} such that

$$f(x_0^{**}) > \sup \{ f(u^{**}) : u^{**} \in \overline{B_X}^{(w^*)} \}. \tag{5.6}$$

By Proposition 5.34, since f is continuous in the weak* topology on X^{**}, there exists an $x^* \in X^*$ such that $f(x^{**}) = x^{**}(x^*)$ for all $x^{**} \in X^{**}$. Therefore, (5.5.1) becomes

$$x_0^{**}(x^*) > \sup \{u^{**}(x^*) : u^{**} \in \overline{B_X}^{(w^*)}\} \geq \sup\{x^*(x) : x \in B_X\} = \|x^*\|.$$

This implies $\|x_0^{**}\| > 1$, contradicting the assumption that $x_0^{**} \in B_{X^{**}}$. The result follows. $\square$

In the proof of Goldstine's Theorem, we assumed that X was a real Banach space for the sake of simplicity. The argument is similar when the Banach space is complex, but instead of Theorem 5.20, which is the Hahn–Banach Separation Theorem for real spaces, we use Theorem 5.22, which is the Hahn–Banach Separation Theorem for complex spaces, and we replace f with $\Re(f)$.

Theorem 5.41 *A Banach space X is reflexive if and only if the closed unit ball B_X is weakly compact.*

Proof Assume first that X is reflexive. Then $B_X = B_{X^{**}}$. By the Banach–Alaoglu Theorem (Theorem 5.39), the set $B_{X^{**}}$ is compact in the weak* topology on X^{**}. Since X is reflexive, the weak* topology on X^{**} coincides with the weak topology on X. Therefore, B_X is compact in the weak topology on X.

Now, assume instead that B_X is weakly compact. The weak topology on X is the restriction of the weak* topology on X^{**}, and so B_X is compact (and hence closed) in the weak* topology on X^{**}. By Goldstine's Theorem (Theorem 5.40), we conclude that $B_X = B_{X^{**}}$, since B_X is closed and dense in $B_{X^{**}}$. Therefore, X is reflexive. $\square$

Proposition 5.42 *Suppose X and Y are Banach spaces (or simply normed linear spaces). If $T : X \to Y$ is a linear map, then the following are equivalent:*

 (i) *T is bounded (i.e., norm-to-norm continuous).*
 (ii) *T is $(X, \| \cdot \|)$ to (Y, w) continuous.*
(iii) *T is (X, w) to (Y, w) continuous.*

Proof Certainly *(iii)* implies *(ii)*. We will show that *(ii)* implies *(i)*, and then *(i)* implies *(iii)*.

Assume *(ii)*. We wish to show that $T(B_X)$ is bounded in the norm topology on Y. Let $y^* \in Y^*$. Then y^* is continuous in the weak topology on Y. Consequently, since T is norm-to-weak continuous, the functional $y^* \circ T$ is continuous in the norm topology on X. Thus, $y^* \circ T \in X^*$, and so

$$\sup_{\|x\| \leq 1} |y^*(Tx)| = \sup_{\|x\| \leq 1} |(y^* \circ T)(x)| < \infty. \tag{5.7}$$

Since (5.5.2) holds for each $y^* \in Y^*$, we conclude that the set $T(B_X)$ is weakly bounded in Y. Therefore, $T(B_X)$ is bounded in the norm topology, by Theorem 4.12.

Now assume *(i)*. Consider a weak neighborhood in Y, say

$$W_Y = W_Y(y_1^*, \ldots, y_n^*; \epsilon) = \{y : |y_j^*(y)| < \epsilon, \ 1 \leq j \leq n\},$$

for $\{y_1^*, \ldots, y_n^*\} \subseteq Y^*$ and $\epsilon > 0$. Suppose $x \in X$ is such that $Tx \in W_Y$. Then for each $j \in \{1, \ldots, n\}$, we have $|y_j^*(Tx)| < \epsilon$. Recall that the adjoint operator T^* was defined so that $T^* \circ y^* = y^* \circ T$. Therefore, $|T^* y_j^*(x)| < \epsilon$ for all $j \in \{1, \ldots, n\}$, and so it follows that $x \in W_X = W_X(T^* y_1^*, \ldots, T^* y_n^*; \epsilon)$, a weak neighborhood of X. We conclude that $T^{-1}(W_Y) \subseteq W_X$. Equality is obtained by running through the same argument in reverse, and so T is weak-to-weak continuous, as required.   □

Suppose that $T : X \to Y$ is a bounded linear mapping between real Banach spaces. If X is reflexive, then $T(B_X)$ is weakly compact, and hence norm-closed in Y. This is not true in general (i.e., for non-reflexive spaces X). Consider any $x^* \in X^*$ with $\|x^*\| = 1$. Then $x^*(B_X)$ could be either $(-1, 1)$ or $[-1, 1]$. If X is reflexive, then the second interval (the closed one) is the only option.

Example 5.43 Consider the real Banach space $X = \ell_1$. Recall that $\ell_1^* = \ell_\infty$. Let ξ in ℓ_∞ be the bounded sequence $\xi = (1 - 1/n)_{n=1}^\infty$. Now suppose $x = (x_n)_{n=1}^\infty$ is any element in B_{ℓ_1}, so that $\sum_{n=1}^\infty |x_n| \le 1$. Then

$$|\xi(x)| = \Big| \sum_{n=1}^\infty \xi_n \, x_n \Big| \le \sum_{n=1}^\infty \xi_n \, |x_n| < 1.$$

Since $\|\xi\|_{\ell_\infty} = 1$, we have a norm-one element $\xi \in \ell_1^*$ such that $\xi(B_{\ell_1}) = (-1, 1)$.

We see that a linear functional on a reflexive Banach space attains its maximum value on the closed unit ball. It was a long standing question whether or not this property characterized reflexive spaces. In 1964, R.C. James showed that it did when he proved the statement: *If every bounded linear functional on X attains its maximum value on the closed unit ball, then X is reflexive* [19].

We conclude this section with a result about the adjoint operator.

Proposition 5.44 *If $T : X \to Y$ is a bounded linear map between Banach spaces, then $T^* : Y^* \to X^*$ is weak*-to-weak* continuous.*

Proof The proof is very similar to the proof that *(i)* implies *(iii)* in Proposition 5.42. Consider a weak* neighborhood in X^*, say

$$W_{X^*} = W_{X^*}(x_1, \ldots, x_n; \epsilon) = \{x^* : |x^*(x_j)| < \epsilon, \ 1 \le j \le n\},$$

for $\{x_1, \ldots, x_n\} \subseteq X$ and $\epsilon > 0$. Suppose $y^* \in Y^*$ is such that $T^* y^* \in W_{X^*}$. Then $|T^* y^*(x_j)| < \epsilon$ for all $j \in \{1, \ldots, n\}$. But $T^* y^*(x) = y^*(Tx)$ for all $x \in X$, and so $|y^*(Tx_j)| < \epsilon$ for all $j \in \{1, \ldots, n\}$. Then $y^* \in W_{Y^*} = W_{Y^*}(Tx_1, \ldots, Tx_n; \epsilon)$, which is a weak* neighborhood of Y^*. This implies that $(T^*)^{-1}(W_{X^*}) \subseteq W_{Y^*}$. Similarly, $W_{Y^*} \subseteq (T^*)^{-1}(W_{X^*})$, and so T^* is weak*-to-weak* continuous, as required.   □

5.6 Mazur's Theorem

In this section we explore the consequences of convexity in weak topologies.

Theorem 5.45 (Mazur's Theorem) *Let X be a locally convex topological vector space. A convex subset of X is closed if and only if it is weakly closed.*

Proof Without loss of generality, assume X is a real topological vector space. A weakly closed set is always strongly closed, regardless of convexity. Suppose K is closed in the original topology, and let $\overline{K}^{(w)}$ denote the closure of K in the weak topology. Assume $x_0 \in \overline{K}^{(w)} \setminus K$. Then, by the Hahn–Banach Separation Theorem (Theorem 5.20), there is an $x^* \in X^*$ such that

$$x^*(x_0) > \sup\{x^*(x) : x \in K\}.$$

This contradicts the assumption x_0 is in the weak closure of K, and so $\overline{K}^{(w)} = K$.   □

Example 5.46 Consider the real sequence space ℓ_p, where $1 < p < \infty$. As usual, for each $n \in \mathbb{N}$, let e_n be the sequence with 1 in the nth coordinate, and 0 elsewhere. We know that $e_n \to 0$ weakly as $n \to \infty$ (see Example 5.28), and so 0 is in the weak closure of the set $E = \{e_n : n \in \mathbb{N}\}$. Let $\mathrm{co}(E)$ denote the set of convex linear combinations of elements in E:

$$\mathrm{co}(E) = \Big\{ \sum_{j=1}^{m} \lambda_j e_j \ : \ \lambda_j \geq 0, \ \sum_{j=1}^{m} \lambda_j = 1, \ m \in \mathbb{N} \Big\}.$$

We denote the closure of $\mathrm{co}(E)$ in the norm topology by $\overline{\mathrm{co}}(E)$. This set is convex and closed in the norm topology, and hence weakly closed by Mazur's Theorem (Theorem 5.45). It follows that $\overline{\mathrm{co}}(E)$ contains 0, since 0 is a weak limit point of E. We conclude that 0 can be approximated (in norm) by convex linear combinations of elements in $E = \{e_n : n \in \mathbb{N}\}$. In fact, if $1 < p < \infty$, then

$$\Big\| \frac{1}{n}(e_1 + \cdots + e_n) \Big\|_{\ell_p} = \frac{1}{n} \Big(\sum_{j=1}^{n} 1^p \Big)^{1/p} = n^{\frac{1}{p}-1} \xrightarrow[n \to \infty]{} 0.$$

The same cannot be said of ℓ_1—in this case, there exists no convex combination of elements in $\{e_n : n \in \mathbb{N}\}$ that will approximate 0. To see this, let $\lambda_1 e_1 + \cdots + \lambda_n e_n$ be any convex combination of elements from $\{e_n : n \in \mathbb{N}\}$. Then

$$\|\lambda_1 e_1 + \cdots + \lambda_n e_n\|_{\ell_1} = \|(\lambda_1, \lambda_2, \cdots, \lambda_n, 0, \dots)\|_{\ell_1} = \sum_{j=1}^{n} \lambda_j = 1.$$

In the above example, we introduced the notation $\mathrm{co}(E)$ to denote the set of convex linear combinations of elements in the set E. This idea will prove important later, and so we give the following definition.

Definition 5.47 Let X be a topological vector space and let A be any subset of X. The *convex hull* of A is the smallest convex subset of X that contains A. The convex hull of A is denoted by $\mathrm{co}(A)$ and consists of all convex linear combinations of elements in A; that is,

$$\mathrm{co}(A) = \Big\{ \sum_{j=1}^{m} \lambda_j a_j \ : \ a_j \in A, \ \lambda_j > 0, \ \sum_{j=1}^{m} \lambda_j = 1, \ m \in \mathbb{N} \Big\}.$$

The *closed convex hull* of A is the closure of the convex hull and is denoted $\overline{\mathrm{co}}(A)$.

Example 5.48 Let $X = L_p(0, 1)$, where $0 < p < 1$. (See Example A in Sect. 5.3.)
In Example 5.19, we saw that the only nonempty open convex subset of X is X itself.
Consequently, if $B_X = \{f : \|f\|_p \leq 1\}$, then $\mathrm{co}(B_X) = X$.

Let W be a subset of a topological vector space X. If f is a linear functional
on X such that $|f(x)| \leq M$ for all $x \in W$, then, by the definition of the convex
hull, it must be that $|f(x)| \leq M$ for all $x \in \mathrm{co}(W)$. In the case of $L_p(0, 1)$, where
$0 < p < 1$, from the example above, the convex hull of the unit ball is the entire
space. Since no nonzero linear functionals are bounded on $L_p(0, 1)$ (see Example A
in Sect. 5.3), it follows that there are no nonzero linear functionals bounded on the
unit ball of $L_p(0, 1)$. Indeed, there are no nonzero linear functionals bounded on any
nonempty open subsets of $L_p(0, 1)$ if $0 < p < 1$.

We now give an example of how convexity can help to solve optimization
problems. We begin by making a definition.

Definition 5.49 Let X be a vector space. A function $f : X \to \mathbb{R}$ is called *convex* if

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y),$$

for all $\{x, y\} \subseteq X$ and $t \in [0, 1]$.

Our problem is as follows: Suppose K is a closed bounded convex set in a Banach
space X and let $f : X \to \mathbb{R}$ be a continuous convex function. Does there exist some
$x_0 \in K$ such that

$$f(x_0) = \min\{f(x) : x \in K\}?$$

There is no reason we should assume this minimum exists, as there is no compactness
assumption made on K.

Suppose for a moment that X is reflexive. Then B_X is weakly compact, by The-
orem 5.41. Since K is closed in norm, K is weakly closed, by Mazur's Theorem. It
follows that K is weakly compact. (Here, we use the fact that B_X is weakly com-
pact and absorbent.) We now have continuity and compactness, but not in the same
topology: f is continuous in the norm topology, and K is compact in the weak
topology.

Despite this, we claim that if X is a reflexive Banach space and if f is a convex
function, then f does attain its minimum value on K. Let $\alpha = \inf\{f(x) : x \in K\}$.
Our goal is to show that $\alpha > -\infty$ and that $f(x_0) = \alpha$ for some $x_0 \in K$.

Suppose $\alpha = -\infty$ and, for each $n \in \mathbb{N}$, define $K_n = \{x \in K : f(x) \leq -n\}$. For
each $n \in \mathbb{N}$, the set K_n is closed (and hence weakly closed by Mazur's Theorem),
convex, and nonempty (since $\alpha = -\infty$). Therefore, $(K_n)_{n=1}^\infty$ forms a nested sequence
of weakly compact nonempty sets. By the Nested Interval Property (Corollary B.7),
it must be that $\bigcap_{n=1}^\infty K_n \neq \emptyset$. But this implies that there is some $x_0 \in K$ such that
$f(x_0) \leq -n$ for all $n \in \mathbb{N}$, an impossibility. Consequently, $\alpha > -\infty$.

Define for $n \in \mathbb{N}$ a sequence of sets $K_n' = \{x \in K : f(x) \leq \alpha + 1/n\}$. As
before, $(K_n')_{n=1}^\infty$ is a nested sequence of weakly compact nonempty sets, and so
$\bigcap_{n=1}^\infty K_n' \neq \emptyset$. If $x_0 \in \bigcap_{n=1}^\infty K_n'$, then $x_0 \in K$ and $f(x_0) = \alpha$, as required.

We summarize in the following proposition.
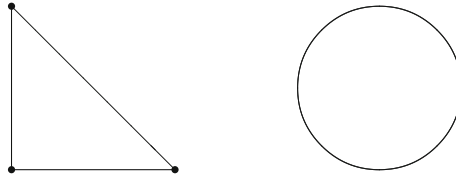
Fig. 5.2 Some elementary convex objects

Proposition 5.50 *Suppose K is a nonempty closed bounded convex set in a Banach space X and let $f : X \to \mathbb{R}$ be a continuous convex function. If X is reflexive, then there exists an $x_0 \in K$ such that $f(x_0) = \min\{f(x) : x \in K\}$.*

Proof See the discussion preceding the statement of the proposition.   □

A special case of the above is $f(x) = \|u - x\|$, the function representing the distance between $x \in X$ and a fixed point $u \in X$. If X is a reflexive Banach space, and if K is a closed and bounded convex set in X (not containing u), then there exists a $x_0 \in K$ such that

$$\|u - x_0\| = \min_{x \in K} \|u - x\|.$$

That is, there exists some point $x_0 \in K$ which is closest to u. (Actually, the boundedness assumption on K is not needed for this statement to be true.)

5.7 Extreme Points

In this section, we consider sets K that are convex in some vector space X.

Definition 5.51 Let X be a vector space and suppose K is a convex subset of X. A point $x \in K$ is an *extreme point* of K if it does not lie on a line segment in K. That is, x is an extreme point of K provided that the following is true: If u and v are elements of K such that $x = (1 - t)u + tv$ for some $t \in (0, 1)$, then $x = u = v$.

The set of extreme points of K is denoted $\mathrm{ex}(K)$.

For example, a triangle has an extreme point at each vertex, while any boundary point of a circle is an extreme point. (See Fig. 5.2.)

Example 5.52 We now determine the extreme points for the unit ball B_X in several cases where X is a real Banach space. Note that no point of the interior of B_X can be extreme, and so we must consider only points on the boundary ∂B_X.

(i) $X = \ell_2$. Denote the inner product on ℓ_2 by $\langle \cdot, \cdot \rangle$. Suppose $x \in \ell_2$ is such that $\|x\| = 1$. Now let $\{u, v\} \subseteq B_{\ell_2}$ and suppose $x = (1 - t)u + tv$ for some $t \in (0, 1)$. By the triangle inequality, $\|u\| = \|v\| = 1$ (otherwise $\|x\| < 1$). Since $\|x\| = 1$, we have

$$1 = \langle x, x \rangle = (1 - t)\langle u, x \rangle + t\langle v, x \rangle. \tag{5.8}$$

By assumption, $0 < t < 1$, and so (5.7.1) implies that $\langle u, x \rangle = \langle v, x \rangle = 1$ (again using the triangle inequality). Observe that $\langle u, x \rangle = 1 = \|u\|\|x\|$, and thus we have equality in the Cauchy–Schwarz Inequality. This can only happen if $u = x$. Similarly, $v = x$. Therefore, x is an extreme point of B_{ℓ_2} whenever x is on the boundary ∂B_{ℓ_2}.

(ii) $X = L_p(0, 1)$, $1 < p < \infty$. We will take our cue from the preceding example. Suppose $f \in L_p(0, 1)$ is such that $\|f\|_p = 1$. Let $\{g, h\} \subseteq L_p(0, 1)$ and suppose that $f = (1 - t)g + th$ for some $t \in (0, 1)$. Then $\|g\|_p = \|h\|_p = 1$, by the triangle inequality (as in *(i)*). By the Hahn–Banach Theorem, there exists a linear functional ϕ in $L_p(0, 1)^*$ such that $\|\phi\| = 1$ and $\phi(f) = 1$. In fact, in this case we can write ϕ explicitly:

$$\phi(k) = \int_0^1 |f(x)|^{p-1}(\operatorname{sign} f(x)) \, k(x) \, dx, \quad k \in L_p(0, 1).$$

Since $\phi(f) = \|f\|_p^p$, we have

$$1 = \phi(f) = (1 - t)\phi(g) + t\phi(h).$$

Again using the triangle inequality, we have that $\phi(g) = \phi(h) = 1$.
By Hölder's Inequality,

$$|\phi(g)| \leq \int |f(x)|^{p-1} |g(x)| \, dx \leq \left(\int |f(x)|^{(p-1)q} \, dx \right)^{1/q} \|g\|_p = \|f\|_p^{p/q} \|g\|_p,$$

where $\frac{1}{p} + \frac{1}{q} = 1$ (and so $q = \frac{p}{p-1}$). Since the left and right sides of the above inequality are both equal to 1, we have equality in Hölder's Inequality. This happens only if there are positive constants a and b such that $a(|f|^{p-1})^q = b|g|^p$ (as members of $L_p(0, 1)$). Because $q = \frac{p}{p-1}$, this equality (which is valid almost everywhere) becomes $a|f|^p = b|g|^p$, which is equivalent to $a^{1/p}|f| = b^{1/p}|g|$. Since $\|f\|_p = \|g\|_p$, this can only happen if $|f| = |g|$ in $L_p(0, 1)$. A similar argument shows $|f| = |h|$ in $L_p(0, 1)$. From these equalities, together with the assumption that f is a convex combination of g and h, we conclude that $f = g$ and $f = h$. It follows that f is an extreme point of $B_{L_p(0,1)}$. The choice of f was arbitrary in $\partial B_{L_p(0,1)}$, and therefore any f on the boundary of the unit ball is an extreme point of the unit ball $B_{L_p(0,1)}$.

A similar argument shows that the extreme points of the unit ball in ℓ_p, where $1 < p < \infty$, are the elements of the boundary ∂B_{ℓ_p}.

(iii) $X = L_1(0, 1)$. If f is an extreme point of $B_{L_1(0,1)}$, then f is on the boundary of $B_{L_1(0,1)}$. Let f be a function in $L_1(0, 1)$ such that $\|f\|_1 = \int_0^1 |f(s)| \, ds = 1$. Define a new function F on $[0, 1]$ by

$$F(t) = \int_0^t |f(s)| \, ds, \quad t \in [0, 1].$$

The function F is continuous with $F(0) = 0$ and $F(1) = 1$. By the Intermediate Value Theorem, there exists a $\tau \in (0, 1)$ such that $F(\tau) = 1/2$.

Let $g = 2f \, \chi_{(0,\tau)}$ and $h = 2f \, \chi_{(\tau,1)}$. By the choice of τ, we have $\|g\|_1 = 1$ and $\|h\|_1 = 1$. We have found distinct functions g and h in $B_{L_1(0,1)}$ such that $f = \frac{1}{2}g + \frac{1}{2}h$. Therefore, f is not an extreme point of the unit ball of $L_1(0,1)$. Since f was an arbitrary element of $\partial B_{L_1(0,1)}$, we conclude that the unit ball in $L_1(0,1)$ has no extreme points.

(iv) $X = c_0$. We will show there are no extreme points in B_{c_0}. Suppose $x = (x_k)_{k=1}^{\infty}$ is a sequence in B_{c_0}. Then $\lim_{k \to \infty} x_k = 0$, and so there exists some $n \in N$ such that $|x_n| < 1/2$. We will define sequences $y = (y_k)_{k=1}^{\infty}$ and $z = (z_k)_{k=1}^{\infty}$ in c_0 so that $x = \frac{1}{2}y + \frac{1}{2}z$. If $x_n \neq 0$, then define y and z as follows:

$$y_k = \begin{cases} x_k & \text{if } k \neq n, \\ 0 & \text{if } k = n \end{cases} \quad \text{and} \quad z_k = \begin{cases} x_k & \text{if } k \neq n, \\ 2x_n & \text{if } k = n. \end{cases}$$

We have $x = \frac{1}{2}y + \frac{1}{2}z$ and, because $|x_n| < 1/2$, the sequences y and z are in B_{c_0}. The previous sequences work only if $x_n \neq 0$. If instead $x_n = 0$, then define y and z so that:

$$y_k = \begin{cases} x_k & \text{if } k \neq n, \\ \frac{1}{2} & \text{if } k = n \end{cases} \quad \text{and} \quad z_k = \begin{cases} x_k & \text{if } k \neq n, \\ -\frac{1}{2} & \text{if } k = n. \end{cases}$$

Once again, we have $x = \frac{1}{2}y + \frac{1}{2}z$, where the sequences y and z are in B_{c_0}. Therefore, x is not an extreme point of the set B_{c_0}.

(v) $X = C[0,1]$. In this case the extreme points of $B_{C[0,1]}$ are the two functions $\chi_{[0,1]}$ and $-\chi_{[0,1]}$. (That is, the constant functions 1 and -1.) If $f \in B_{C[0,1]}$ is a continuous function such that $|f(t)| < 1$ for some $t \in (0,1)$, then we may use an argument similar to the perturbation argument used in *(iv)*. (That is, we can put a small "wiggle" in the function.)

Similarly, if K is a compact Hausdorff space, then the extreme points of $B_{C(K)}$ are the two functions χ_K and $-\chi_K$, which are the constant functions 1 and -1 on K.

In the case that $C(K)$ is a *complex* Banach space, the extreme points of $B_{C(K)}$ are all functions $f \in C(K)$ for which $|f(s)| = 1$ for all $s \in K$.

(vi) $X = \ell_1$. The set of extreme points in B_{ℓ_1} is $\{\pm e_n : n \in \mathbb{N}\}$. First, let us show that each of these points is indeed extreme in B_{ℓ_1}. Fix some $n \in \mathbb{N}$. Let $y = (y_k)_{k=1}^{\infty}$ and $z = (z_k)_{k=1}^{\infty}$ be elements of B_{ℓ_1} such that $e_n = ay + bz$, where a and b are positive numbers such that $a + b = 1$. (Note that these conditions imply $y_n > 0$ and $z_n > 0$.) We have that $ay_n + bz_n = 1$ and $ay_k + bz_k = 0$ for all $k \neq n$. By assumption, we know that $a \neq 0$, and so $y_n = \frac{1}{a} - \frac{b}{a}z_n$ and $y_k = -\frac{b}{a}z_k$ for $k \neq n$. Once again making use of the triangle inequality, we see that $\|y\|_1 = 1$ and $\|z\|_1 = 1$. Computing $\|y\|_1$ directly, we have

$$\left(\sum_{k \neq n} |y_k|\right) + y_n = \left(\sum_{k \neq n} \frac{b}{a}|z_k|\right) + \left(\frac{1}{a} - \frac{b}{a}z_n\right) = \left(\frac{b}{a}\sum_{k=1}^{\infty}|z_k|\right) + \left(\frac{1}{a} - \frac{2b}{a}z_n\right).$$

Thus,

$$\|y\|_1 = \frac{b}{a}\|z\|_1 + \left(\frac{1}{a} - \frac{2b}{a}z_n\right).$$

But $\|y\|_1 = 1$ and $\|z\|_1 = 1$, and so

$$1 = \frac{b}{a}(1) + \left(\frac{1}{a} - \frac{2b}{a}z_n\right) = \frac{b + 1 - 2bz_n}{a}.$$

A little arithmetic (and the fact that $a + b = 1$) reveals that $z_n = 1$, and so z must in fact be e_n. Therefore, e_n is an extreme point. A similar argument shows that $-e_n$ is an extreme point for each $n \in \mathbb{N}$.

Now we show that no other element of ∂B_{ℓ_1} is an extreme point. Suppose $x \in B_{\ell_1}$ with $\|x\|_1 = 1$, but $x \neq \pm e_n$ for any $n \in \mathbb{N}$. Then there must be at least two non-zero entries, say x_{m_1} and x_{m_2}. Without loss of generality, we may assume both terms are positive. Choose some constant $\epsilon > 0$ such that $\epsilon < \min\{x_{m_1}, x_{m_2}, 1 - x_{m_1}, 1 - x_{m_2}\}$. Define $y = (y_k)_{k=1}^{\infty}$ and $z = (z_k)_{k=1}^{\infty}$ in ℓ_1 as follows:

$$y_k = \begin{cases} x_k & \text{if } k \notin \{m_1, m_2\}, \\ x_{m_1} + \epsilon & \text{if } k = m_1, \\ x_{m_2} - \epsilon & \text{if } k = m_2 \end{cases} \quad \text{and} \quad z_k = \begin{cases} x_k & \text{if } k \notin \{m_1, m_2\}, \\ x_{m_1} - \epsilon & \text{if } k = m_1, \\ x_{m_2} + \epsilon & \text{if } k = m_2. \end{cases}$$

It follows that $x = \frac{1}{2}y + \frac{1}{2}z$, where $\{y, z\} \subseteq B_{\ell_1}$. Therefore, x is not an extreme point.

The next theorem describes a condition under which the set of extreme points is never empty.

Theorem 5.53 (Krein–Milman Theorem) *Suppose E is a locally convex Hausdorff topological vector space. If K is a nonempty compact convex subset of E, then $K = \overline{\mathrm{co}}(\mathrm{ex}\,K)$. In particular, $\mathrm{ex}(K) \neq \emptyset$.*

Before proving Theorem 5.53, let us observe a consequence.

Corollary 5.54 *If X is a Banach space, then $B_{X^*} = \overline{\mathrm{co}}^{(w^*)}(\mathrm{ex}\,B_{X^*})$.*

Proof By Theorem 5.39, the set B_{X^*} is compact in the w^*-topology whenever X is a Banach space. The result then follows from the Krein–Milman Theorem. □

What makes this result so interesting is that we can now, courtesy of Example 5.52, conclude that neither c_0 nor $L_1(0, 1)$ is a dual space of a Banach space, since the unit balls in these spaces have no extreme points. While the unit ball of $C[0, 1]$ does have extreme points, it does not have enough (only two!) to construct the entire unit ball using only convex linear combinations. Therefore, $C[0, 1]$ cannot be the dual space of any Banach space, either.

In order to proceed with the proof of Theorem 5.53, we now introduce a definition and a lemma.

Definition 5.55 Let E be a topological vector space with nonempty subset K. A subset F of K is called *extremal* (in K) if F is a nonempty compact convex set such

that the following holds: If $\{u, v\} \subseteq K$ and $(1 - t)u + tv \in F$ for all $t \in (0, 1)$, then $\{u, v\} \subseteq F$.

Lemma 5.56 *Suppose K is a nonempty compact convex subset of a locally convex topological vector space. Every extremal subset of K contains an extreme point.*

Proof Let K be a nonempty compact convex subset of the locally convex topological vector space E and suppose F is an extremal set in K. Consider the partially ordered set of all subsets of F that are extremal in K, where $G \geq H$ whenever $G \subseteq H$. We wish to find a *maximal* element of this partially ordered set (which in turn will be a *minimal* extremal subset of F).

Suppose $\mathcal{C} = (G_i)_{i \in I}$ is a chain of subsets of F that are extremal in K. Let G be the intersection of all sets in $\mathcal{C}$, so $G = \bigcap_{i \in I} G_i$. Then G is a compact and convex subset of F. For any finite collection of indices $i_1, \ldots, i_n$ in I, we have $\bigcap_{k=1}^n G_{i_k} = G_{i_j}$ for some $j \in \{1, \ldots, n\}$, because $\mathcal{C}$ is a chain of subsets. Therefore, by the Finite Intersection Property, G is nonempty.

We claim that G is extremal in K. Suppose $\{u, v\} \subseteq K$ and $(1 - t)u + tv \in G$ for all $t \in (0, 1)$. For each $i \in I$, we have $G \subseteq G_i$, and so $(1 - t)u + tv \in G_i$ for all $t \in (0, 1)$. But G_i is extremal in K, and hence $\{u, v\} \subseteq G_i$. It follows that $\{u, v\} \subseteq \bigcap_{i \in I} G_i = G$. Thus, G is extremal in K.

By Zorn's Lemma, there exists a maximal element in the partially ordered set of subsets of F that are extremal in K. Hence, there is a minimal subset of F that is extremal in K. Denote this minimal extremal set by F_0. We will show that F_0 consists of only one element. Assume to the contrary that there exist distinct elements u and v in F_0. The space E is Hausdorff, and so $\{u\}$ and $\{v\}$ are closed sets. By the Hahn–Banach Separation Theorem (Theorem 5.20 for real spaces and Theorem 5.22 for complex spaces), there exists a continuous linear functional ϕ on E such that $\phi(u) \neq \phi(v)$. In particular, ϕ is not constant on F_0.

Without loss of generality, we may assume ϕ is real-valued. The functional ϕ is continuous on E, and therefore attains its maximum on the compact set F_0. Let

$$G_0 = \Big\{ x \in F_0 : \phi(x) = \max_{\xi \in F_0} \phi(\xi) \Big\}.$$

Let M denote the maximum value of ϕ on F_0, so that $\phi(x) = M$ for all $x \in G_0$.

The set G_0 is nonempty (by the continuity of ϕ), compact, and convex. It is also a proper subset of F_0, because ϕ is not constant. We claim that G_0 is extremal in K. Suppose $\{u, v\} \subseteq K$ and $(1 - t)u + tv \in G_0$ for all $t \in (0, 1)$. We know that $G_0 \subseteq F_0$, and we know that F_0 is extremal; hence $\{u, v\} \subseteq F_0$. For each $t \in (0, 1)$, we have that $(1 - t)u + tv \in G_0$, and so

$$M = \phi((1 - t)u + tv) = (1 - t)\phi(u) + t\phi(v),$$

for all $t \in (0, 1)$. From this we conclude that $\phi(u) = \phi(v) = M$, and consequently $\{u, v\} \subseteq G_0$. This implies G_0 is extremal, but this violates the minimality of F_0. We have derived a contradiction, and so it must be the case that F_0 contains only one element. The set F_0 is a single-point set and an extremal set. Therefore, the one element of F_0 is an extreme point.

We are now prepared to prove the Krein–Milman Theorem.

Proof of Theorem 5.53   Without loss of generality, we may assume E is a real topological vector space. The set K is extremal in itself, and so must contain an extreme point, by Lemma 5.56. Let $K_0 = \overline{\mathrm{co}}(\mathrm{ex}\,K)$, the closed convex hull of the set of extreme points in K. Suppose $x \in K \backslash K_0$. By the Hahn–Banach Separation Theorem (Theorem 5.20), there exists a continuous linear functional f on E such that

$$f(x) > \max_{y \in K_0} f(y). \tag{5.9}$$

Let

$$G_0 = \big\{z \in K : f(z) = \max_{y \in K} f(y)\big\}.$$

The set G_0 is nonempty because f is continuous and K is compact; it is also disjoint from K_0 because of (5.9). The set G_0 is extremal (see the argument in the proof of Lemma 5.56), and consequently contains an extreme point of K, by Lemma 5.56. This, however, is a contradiction, because K_0 contains all of the extreme points of K, and K_0 and G_0 are disjoint. Therefore, $K = K_0$, as required.

The Krein–Milman Theorem originally appeared in a work by Krein and Milman in 1940 [21]. The local convexity assumption on E was needed to invoke the Hahn–Banach Separation Theorem (Theorem 5.20). Local convexity is a necessary condition, a fact which was not shown until the 1970s [32]. The Krein–Milman Theorem has a deep relationship with the Axiom of Choice. (See [4].)

5.8   Milman's Theorem

Suppose K is a compact Hausdorff space. The Riesz Representation Theorem identifies the dual space of the space of continuous functions on K as the space of regular Borel measures on K; that is, $C(K)^* = M(K)$. (See Theorem A.35.) We recall that the norm on $M(K)$ is the total variation norm: $\|\mu\|_M = |\mu|(K)$ for all $\mu \in M(K)$. We define the *probability measures* on K to be elements in the set

$$\mathcal{P}(K) = \{\mu \in M(K) \,:\, \mu \geq 0, \ \|\mu\|_M = 1\}.$$

This set is convex. It is also closed in the w^*-topology, which can be seen from the equality $\mathcal{P}(K) = \{\mu \in B_{M(K)} : \int_K 1 \, d\mu = 1\}$. (See Exercise 5.27.)

By the Banach–Alaoglu Theorem (Theorem 5.39), the unit ball $B_{M(K)}$ is compact in the w^*-topology, and hence $\mathcal{P}(K)$ is w^*-compact as a w^*-closed subset. A simple computation shows that $\mathcal{P}(K)$ is an extremal set in the unit ball of $M(K)$. (Again, see Exercise 5.27.) Since $\mathcal{P}(K)$ is a w^*-compact convex extremal set, Lemma 5.56 assures us that $\mathcal{P}(K)$ must have at least one extreme point.

Proposition 5.57   *Let K be a compact Hausdorff space. A probability measure in $M(K)$ is an extreme point of $\mathcal{P}(K)$ if and only if it is a Dirac measure.*

Proof We first show that δ_s is an extreme point of $\mathcal{P}(K)$ for $s \in K$. Suppose there exist probability measures μ and ν such that $\delta_s = (1 - t)\mu + t\nu$ for some $t \in (0, 1)$. It follows that $\mu(\{s\}) = \nu(\{s\}) = 1$. Thus, $\mu = \nu = \delta_s$, as required.

Now suppose μ is an extreme point of $\mathcal{P}(K)$. We will show that $\mu = \delta_s$ for some $s \in K$. Let $\mathcal{U} = \{U \text{ open} : \mu(U) = 0\}$ and let $V = \bigcup_{U \in \mathcal{U}} U$. We claim $\mu(V) = 0$. Suppose E is a compact subset of V. The collection of sets $\mathcal{U}$ forms an open cover of E, and so by compactness there exists a finite subcover, say $E \subseteq U_1 \cup \cdots \cup U_n$. The measure μ is nonnegative, and so (by subadditivity)

$$\mu(E) \leq \mu(U_1) + \cdots + \mu(U_n) = 0.$$

Therefore, $\mu(E) = 0$ for all compact subsets of V. By the regularity of μ,

$$\mu(V) = \sup\{\mu(E) : E \text{ is a compact subset of } V\} = 0.$$

Let $F = K \backslash V$. Then $\mu(F) = 1$. We wish to show that F contains only one point. Assume to the contrary that F contains more than one point. Let $\{s, t\} \subseteq F$. By the Hausdorff property, there are open sets W_1 and W_2 in K such that $s \in W_1$, $t \in W_2$, and $W_1 \cap W_2 = \emptyset$. Since W_1 and W_2 are not subsets of V, they have non-zero μ-measure. Define measures μ_1 and μ_2 on K as follows:

$$\mu_1(B) = \frac{\mu(W_1 \cap B)}{\mu(W_1)} \quad \text{and} \quad \mu_2(B) = \frac{\mu((K \backslash W_1) \cap B)}{\mu(K \backslash W_1)},$$

where B is any measurable subset of K. We note that $\mu(K \backslash W_1) \neq 0$ since $\mu \geq 0$ and $W_2 \subseteq K \backslash W_1$.

Both μ_1 and μ_2 are probability measures, and

$$\mu(W_1)\,\mu_1 + \mu(K \backslash W_1)\,\mu_2 = \mu.$$

By assumption, μ is an extreme point of $\mathcal{P}(K)$. Thus, since $\mu(W_1) + \mu(K \backslash W_1) = 1$, it follows that $\mu = \mu_1 = \mu_2$. This is not possible, however, since $\mu_1(W_1) = 1$ and $\mu_2(W_1) = 0$. We have derived a contradiction. Therefore, there can be no more than one point in the set F, say s. Since $\mu(F) = 1$, it follows that $\mu = \delta_s$, as required. $\quad\blacksquare$

In the above proof, the set V is the maximal open set of μ-measure zero. The entire μ-mass of K is contained in $K \backslash V$. This motivates the next definition.

Definition 5.58 Suppose μ is a positive nonzero measure on K. If V is the maximal open set of μ-measure zero, then $K \backslash V$ is called the *support of μ*. If μ is a signed (or complex) measure, the support of μ is defined to be the support of $|\mu|$.

In the proof of Proposition 5.57, observe that the measures μ_1 and μ_2 were defined in such a way that they had disjoint supports. As a result, it was certainly the case that $\mu_1 \neq \mu_2$. (They "live" on different sets, so to speak.)

Theorem 5.59 (Milman's Theorem) *Suppose E is a locally convex Hausdorff topological vector space and let K be a compact subset of E. If D is a closed subset of K such that $K = \overline{co}(D)$, then $ex(K) \subseteq D$. Furthermore, for every $x \in K$, there*

exists a $\mu_x \in \mathcal{P}(D)$ *such that* $f(x) = \int_D f(y)\,\mu_x(dy)$ *for all linear functionals* $f \in E^*$.

Proof Observe that $E \subseteq \mathbb{K}^{E^*}$, the collection of all scalar-valued functions on E^*, where the superspace is equipped with the product topology. The inclusion is made explicit by the embedding $\rho : E \to \mathbb{K}^{E^*}$ defined by

$$\rho(e) = (f(e))_{f \in E^*}, \quad e \in E.$$

(We often make this identification implicitly, suppressing the letter ρ.)

Introduce a map $T : M(D) \to \mathbb{K}^{E^*}$, defined by

$$T(\mu) = \left(\int_D f(y)\,\mu(dy) \right)_{f \in E^*}, \quad \mu \in M(D).$$

Observe that T is continuous in the w^*-topology on $M(D)$.

For each $s \in D$, we have

$$T(\delta_s) = (f(s))_{f \in E^*} = s. \tag{5.10}$$

By the w^*-continuity of T, then, it follows that T maps $\overline{\mathrm{co}}^{(w^*)}(\{\delta_s\}_{s \in D})$ onto $\overline{\mathrm{co}}^{(w)}(D)$. We note that we have the weak closure of $\mathrm{co}(D)$ in E because the topology E inherits from $\mathbb{K}^{E^*}$ is the weak topology. (See the comments after Definition 5.27.)

By Theorem 5.53 (the Krein–Milman Theorem) and Proposition 5.57, we make the identification $\overline{\mathrm{co}}^{(w^*)}(\{\delta_s\}_{s \in D}) = \mathcal{P}(D)$. By assumption, $K = \overline{\mathrm{co}}(D)$, and so by Mazur's Theorem (Theorem 5.45), we have that $K = \overline{\mathrm{co}}^{(w)}(D)$. Therefore, the restriction $T|_{\mathcal{P}(D)} : \mathcal{P}(D) \to K$ is a surjection. This proves the second part of the theorem.

It remains to prove the first part of the theorem; that is, that the extreme points of K are in D. Suppose $x \in \mathrm{ex}(K)$. We claim that the set $T^{-1}(x) \subseteq \mathcal{P}(D)$ is an extremal set. Suppose $\{\mu, \nu\} \subseteq \mathcal{P}(D)$ and $(1 - t)\mu + t\nu \in T^{-1}(x)$ for all $t \in (0, 1)$. It follows that $(1 - t)T(\mu) + tT(\nu) = x$ for all $t \in (0, 1)$. By assumption, x is an extreme point in K, and consequently $T\mu = T\nu = x$. Therefore, $\{\mu, \nu\} \subseteq T^{-1}(x)$, and so $T^{-1}(x)$ is extremal.

By Lemma 5.56, $T^{-1}(x)$ contains an extreme point of $\mathcal{P}(D)$. Therefore, there exists some $s \in D$ such that $\delta_s \in T^{-1}(x)$, and consequently $T(\delta_s) = x$. By (5.10), however, $T(\delta_s) = s$, and so $x = s \in D$. The result follows. $\square$

5.9 Haar Measure on Compact Groups

We now turn our attention to topological groups. We saw in Sect. 3.4 that if G is a compact abelian metrizable group, then there exists a unique translation-invariant probability measure on G. We noted at the time that the metrizability assumption was not needed. Now, using the tools of the previous sections, we will extend this

result to include compact topological groups that are not abelian. Let us review some definitions.

Definition 5.60 A group G is called a *topological group* if the set G is endowed with a topology for which the group operations (multiplication and inversion)

$$(s, t) \mapsto s \cdot t \quad \text{and} \quad s \mapsto s^{-1}, \quad (s, t) \in G \times G,$$

are continuous. If G is compact in the given topology, then G is called a *compact group*.

Classical examples of topological groups include $\mathbb{R}^n$ (where the group multiplication is given by addition) and the set of orthogonal $n \times n$ matrices $\mathcal{O}_n$ (where the group multiplication is matrix multiplication). Multiplication in a group is usually denoted either with a dot $(\cdot)$ or by juxtaposition. When the group is abelian, however, it is traditional to use a plus symbol $(+)$, provided it will not result in any confusion.

When G is a compact group, we denote the space of continuous functions on G by $C(G)$. The σ-algebra on G is implicitly taken to be the Borel σ-algebra generated by the open sets in G. We denote the Borel σ-algebra on G by $\mathcal{B}$.

Definition 5.61 Let G be a compact group with Borel algebra $\mathcal{B}$. A measure μ on G is called *left-invariant* if $\mu(gB) = \mu(B)$ for all $B \in \mathcal{B}$ and $g \in G$. Correspondingly, the measure μ is called *right-invariant* if $\mu(Bg) = \mu(B)$ for all $B \in \mathcal{B}$ and $g \in G$.

Theorem 5.62 (Existence of Haar Measure) *Suppose G is a compact group. There exists a unique left-invariant probability measure on the Borel sets of G. Furthermore, this measure is also the unique right-invariant probability measure on the Borel sets of G.*

Proof First, we assume we can find a left-invariant probability measure λ and a right-invariant probability measure μ. We will show that $\lambda = \mu$. (Notice that this will imply uniqueness.)

Let $f \in C(G)$. By Fubini's Theorem,

$$\int_G \left(\int_G f(s \cdot t) \, \lambda(dt) \right) \mu(ds) = \int_G \left(\int_G f(s \cdot t) \, \mu(ds) \right) \lambda(dt). \qquad (5.11)$$

By the left-invariance of λ,

$$\int_G \left(\int_G f(s \cdot t) \, \lambda(dt) \right) \mu(ds) = \int_G \left(\int_G f(t) \, \lambda(dt) \right) \mu(ds) = \int_G f(t) \, \lambda(dt),$$

since $\mu(G) = 1$. Similarly, by the right-invariance of μ,

$$\int_G \left(\int_G f(s \cdot t) \, \mu(ds) \right) \lambda(dt) = \int_G \left(\int_G f(s) \, \mu(ds) \right) \lambda(dt) = \int_G f(s) \, \mu(ds),$$

since $\lambda(G) = 1$. Substituting into (5.11), we obtain

$$\int_G f(t) \, \lambda(dt) = \int_G f(s) \, \mu(ds).$$

The choice of $f \in C(G)$ was arbitrary, and so by duality we conclude $\lambda = \mu$.

It remains to prove the existence of a left-invariant measure on G. (A similar argument will produce a right-invariant measure.) We begin by defining for each $s \in G$ a left-multiplication operator $L_s : C(G) \to C(G)$ by

$$(L_s f)(t) = f(s \cdot t), \quad t \in G.$$

Observe that $\|L_s\| = 1$, $L_s^{-1} = L_{s^{-1}}$, and $L_u \circ L_s = L_{u \cdot s}$, whenever u and s are in G.

Claim 1 *Let $f \in C(G)$. The map $s \mapsto L_s f$ is continuous from G into $C(G)$.*

We wish to estimate, for s and s' in G, the quantity

$$\|L_s f - L_{s'} f\|_\infty = \sup_{t \in G} |f(s \cdot t) - f(s' \cdot t)|.$$

Multiplication in the group is continuous, and so the map $(s,t) \mapsto f(s \cdot t)$ is continuous on $G \times G$. Since the group G is compact, we conclude the map $(s,t) \mapsto f(s \cdot t)$ is in fact uniformly continuous. Therefore, for any given $\epsilon > 0$, there exists an open neighborhood V_ϵ of the identity such that $|f(s \cdot t) - f(s' \cdot t')| < \epsilon$ whenever $s' \cdot s^{-1} \in V_\epsilon$ and $t' \cdot t^{-1} \in V_\epsilon$. In this case, we have $t' = t$, and so if $s' \cdot s^{-1} \in V_\epsilon$, then

$$\|L_s f - L_{s'} f\|_\infty = \sup_{t \in G} |f(s \cdot t) - f(s' \cdot t)| < \epsilon.$$

This proves Claim 1.

Claim 2 *Let $\mu \in C(G)$. The map $s \mapsto L_s^* \mu$ is w^*-continuous from G into $M(G)$.*

Observe that $L_s^* : M(G) \to M(G)$. Then, for $f \in C(G)$, by the definition of the adjoint, $\int f \, dL_s^* \mu = \int L_s f \, d\mu$. If s and s' are in G, then

$$\left| \int_G f \, dL_s^* \mu - \int_G f \, dL_{s'}^* \mu \right| = \left| \int_G L_s f \, d\mu - \int_G L_{s'} f \, d\mu \right| \leq \|L_s f - L_{s'} f\|_\infty \|\mu\|_M.$$

The rest follows from Claim 1.

Claim 3 *If $s \in G$, then $L_s^*(\mathcal{P}(G)) \subseteq \mathcal{P}(G)$.*

If $f \geq 0$, then $\int f \, dL_s^* \mu = \int L_s f \, d\mu \geq 0$, whenever $\mu \geq 0$. Thus, $L_s^* \mu \geq 0$ for any $\mu \in \mathcal{P}(G)$. Furthermore,

$$L_s^* \mu(G) = \int_G 1 \, L_s^* \mu(dt) = \int_G L_s(1) \, \mu(dt) = \mu(G) = 1.$$

Thus, $L_s^* \mu$ is in $\mathcal{P}(G)$ whenever μ is a probability measure. This proves Claim 3.

The gist of Claim 3 is that the set $\mathcal{P}(G)$ is invariant under multiplication on the left; i.e., $\mathcal{P}(G)$ is *left-invariant*. We wish to find a set with this property that contains only one element. To that end, let $\mathcal{K}$ be the collection of all weak*-compact convex subsets of $\mathcal{P}(G)$ that are left-invariant; that is, all weak*-compact convex subsets K such that $L_s^* K \subseteq K$ for all $s \in G$. Define a partial order $\leq$ on $\mathcal{K}$ so that $A \leq B$ when $A \subseteq B$. We know that $\mathcal{K}$ is nonempty, because $\mathcal{P}(G) \in \mathcal{K}$. If $(C_i)_{i \in I}$ is a chain in $\mathcal{K}$,

then $C = \bigcap_{i \in I} C_i$ is nonempty, by the Finite Intersection Property. Furthermore, C is a lower bound for the chain $(C_i)_{i \in I}$. Therefore, by Zorn's Lemma, there exists a minimal element of $\mathcal{K}$, say K.

We wish to show that K is a single-point set. Assume to the contrary that μ_1 and μ_2 are distinct elements in K. Let $\nu = \frac{1}{2}(\mu_1 + \mu_2)$. Then $\nu \in K$, by convexity. Define a new set $E = \{L_s^* \nu : s \in G\}$. By Claim 2, the set E is weak*-compact (as the image of the compact set G under a weak*-continuous mapping). Furthermore, $E \subseteq K$, by the left-invariance of K.

For all u and s in G,

$$L_u^*(L_s^* \nu) = L_u^* L_s^* \nu = L_{s \cdot u}^* \nu \in E.$$

Thus $L_u^*(E) \subseteq E$, and so E is left-invariant.

Let $K_0 = \overline{\mathrm{co}}^{(w^*)}(E)$. The set K_0 is convex by construction. We also have that K_0 is weak*-compact, because it is weak*-closed in the weak*-compact set K. By construction, K_0 is left-invariant, and so $K_0 \in \mathcal{K}$. But $K_0 \subseteq K$ and K is minimal in $\mathcal{K}$. Therefore, $K = K_0$.

By the Krein–Milman Theorem (Theorem 5.53), there is some extreme point in K; and by Milman's Theorem (Theorem 5.59), every extreme point of K is in E. Therefore, there exists some $s \in G$ such that $L_s^* \nu$ is extreme in K. Recalling the definition of ν, we see that

$$L_s^* \nu = \frac{1}{2}(L_s^* \mu_1 + L_s^* \mu_2).$$

But $L_s^* \mu_1$ and $L_s^* \mu_2$ are in K, by left-invariance, and $L_s^* \nu$ is extreme in K. Therefore, $L_s^* \nu = L_s^* \mu_1 = L_s^* \mu_2$. If we multiply all sides of this equation by s^{-1} on the left (that is, apply $L_{s^{-1}}^*$ to all sides), we discover that $\nu = \mu_1 = \mu_2$. This violates the assumption that μ_1 and μ_2 are distinct. Thus, K contains only one element, say λ. The measure λ is the desired left-invariant probability measure on G.

Definition 5.63   Let G be a compact group. The unique left-invariant probability measure on the Borel subsets of G is called *Haar measure* on G.

5.10   The Banach–Stone Theorem

In this section, we prove a classical theorem about the structure of spaces of continuous functions. We recall that two Banach spaces X and Y are called *isometrically isomorphic* if there exists a continuous linear bijection that preserves norms. That is, if there exists some linear bijection $T : X \to Y$ such that $\|T\| = \|T^{-1}\| = 1$.

Theorem 5.64   (Banach–Stone Theorem) *Suppose K_1 and K_2 are compact Hausdorff spaces. If $C(K_1)$ and $C(K_2)$ are isometrically isomorphic, then K_1 and K_2 are homeomorphic. Furthermore, if $T : C(K_1) \to C(K_2)$ is an isometric isomorphism,*

then there exists some $u \in C(K_2)$ such that $|u(s)| = 1$ for all $s \in K_2$, and such that

$$Tf(s) = u(s) f(\phi(s)), \quad s \in K_2,$$

where $\phi : K_2 \to K_1$ is a homeomorphism.

Before proving the Banach–Stone Theorem, we will provide a simple lemma that will not only help us now, but will come in handy later, too. In order to prove this lemma, however, we need to make use of another result from general topology.

Theorem 5.65 (Urysohn's Lemma) *A topological space X is normal if and only if any two disjoint closed subsets A and B can be separated by a continuous function. That is, if there exists a continuous function $f : X \to [0, 1]$ such that $f|_A = 0$ and $f|_B = 1$.*

We recall that a topological space X is *normal* if for disjoint closed sets E and F, there exist disjoint open sets U and V such that $E \subseteq U$ and $F \subseteq V$. We will not prove Urysohn's Lemma; however, we will observe that, as a consequence, if K is a compact Hausdorff space, then $C(K)$ *separates the points* of K. That is, if a and b are distinct points in K, then there is a function $f \in C(K)$ such that $f(a) = 0$ and $f(b) = 1$. (See Exercise 5.7.)

Lemma 5.66 *Let K be a compact Hausdorff space. If $\Delta = \{\delta_s : s \in K\}$, then K is homeomorphic to Δ with the subspace topology inherited from $(M(K), w^*)$.*

Proof We remind the reader that for each $s \in K$, the Dirac measure at s is a measure δ_s defined so that $\int_K f \, d\delta_s = f(s)$ for all $f \in C(K)$. The set Δ is closed in the w^* topology and $\Delta \subseteq B_{M(K)}$. Therefore, Δ is w^*-compact.

Define a map $\psi : K \to \Delta$ by $\psi(s) = \delta_s$ for every $s \in K$. Clearly, ψ is a surjection. Suppose that s and t are two distinct points in K such that $\psi(s) = \psi(t)$. Then $\delta_s = \delta_t$. This means that $\delta_s(f) = \delta_t(f)$ for all $f \in C(K)$. Thus, $f(s) = f(t)$ for all $f \in C(K)$. This contradicts the fact that $C(K)$ separates the points of K. (See the comments before the statement of Lemma 5.66.) Therefore, $\psi(s) = \psi(t)$ only if $s = t$, and so ψ is an injection as well as a surjection.

We next show that ψ is a homeomorphism by showing that it is a continuous closed map (so that it maps closed sets to closed sets). Certainly, ψ is continuous in the w^* topology on Δ, since for every $f \in C(K)$,

$$\psi(s)(f) = \int_K f \, d\delta_s = f(s), \quad s \in K,$$

and because f is continuous (by assumption). Now let F be a closed set in K. Then F is compact, because it is a closed subset of the compact set K. Since ψ is continuous for Δ with the w^* topology, it follows that $\psi(F)$ is w^*-compact in Δ. Since $\psi(F)$ is a compact set in a Hausdorff topology, it must be closed (in that topology). Therefore, ψ is a closed map, and it follows that ψ is a homeomorphism. (See Exercise 5.4.)

We are now prepared to prove the Banach–Stone Theorem.

Proof of Theorem 5.64 Let K_1 and K_2 be compact Hausdorff spaces and suppose $T : C(K_1) \to C(K_2)$ is an isometric isomorphism. We wish to show that K_1 and K_2

are homeomorphic. Observe that T maps extreme points of $B_{C(K_1)}$ to extreme points of $B_{C(K_2)}$. To see this, let f be an extreme point in $B_{C(K_1)}$ and suppose that g and h are functions in $B_{C(K_2)}$ such that $Tf = \frac{1}{2}(g + h)$. Then $f = \frac{1}{2}(T^{-1}g + T^{-1}h)$, and from this we deduce that $T^{-1}g = T^{-1}h = f$ (because f is an extreme point). It follows that $g = h = Tf$, and so Tf is an extreme point in $B_{C(K_2)}$ whenever f is an extreme point in $B_{C(K_1)}$. In particular, since χ_{K_1} (which is identically equal to 1 on K_1) is extreme in $B_{C(K_1)}$, its image $T(\chi_{K_1})$ is extreme in $B_{C(K_2)}$. Consequently, we have that $|T(\chi_{K_1})(s)| = 1$ for all $s \in K_2$.

If we define an operator $S : C(K_1) \to C(K_2)$ by $Sf = (Tf)/T(\chi_{K_1})$ for all functions $f \in C(K_1)$, then S is an isometry such that $S(\chi_{K_1}) = \chi_{K_2}$. Therefore, we may assume without loss of generality that $T(\chi_{K_1}) = \chi_{K_2}$.

The adjoint $T^* : M(K_2) \to M(K_1)$ is also an isometry because $(T^*)^{-1} = (T^{-1})^*$, and so $T^*(B_{M(K_2)}) = B_{M(K_1)}$. Suppose $\mu \in \mathcal{P}(K_2)$. Then

$$(T^*\mu)(K_1) = \int_{K_1} 1 \, d\, T^*\mu = \int_{K_2} T(1) \, d\mu = \int_{K_2} 1 \, d\mu = \mu(K_2) = 1.$$

(Here we use the fact that $\chi_{K_1} = 1$ on K_1 and $\chi_{K_2} = 1$ on K_2.) The measure $T^*\mu$ is then an element of $B_{M(K_1)}$ such that $T^*\mu(K_1) = 1$. It can be shown that these two facts imply that $T^*\mu \in \mathcal{P}(K_1)$. (See Exercise 5.28.)

As an isometry, T^* will map extreme points to extreme points. By Proposition 5.57, the extreme points in $M(K_1)$ and $M(K_2)$ are the Dirac masses, and so for each $s \in K_2$, there must be some $t_s \in K_1$ (depending on s) such that $T^*\delta_s = \delta_{t_s}$. Define a map $\phi : K_2 \to K_1$ by $\phi(s) = t_s$ for each $s \in K_2$. Then,

$$T^*\delta_s = \delta_{\phi(s)}, \quad s \in K_2.$$

The map ϕ is a bijection from K_2 onto K_1, which follows from the fact that T^* is a bijection from $M(K_2)$ onto $M(K_1)$. We wish to show that ϕ is a homeomorphism, and as such we must show that both ϕ and ϕ^{-1} are continuous.

Let $\psi : K_1 \to \{\delta_t : t \in K_1\}$ and $\varphi : K_2 \to \{\delta_s : s \in K_2\}$ be defined by $\psi(t) = \delta_t$ and $\varphi(s) = \delta_s$ for all $t \in K_1$ and $s \in K_2$. By Lemma 5.66, the maps ψ and φ are homeomorphisms. Observe that, for all $s \in K_2$,

$$\phi(s) = \psi^{-1}(\delta_{\phi(s)}) = \psi^{-1}(T^*\delta_s) = (\psi^{-1} \circ T^* \circ \varphi)(s).$$

We assumed T was continuous, and hence $T^* : M(K_2) \to M(K_1)$ is weak*-to-weak* continuous, by Proposition 5.44. Therefore, ϕ is continuous. Similarly, we can show that $\phi^{-1} = \varphi^{-1} \circ (T^{-1})^* \circ \psi$, and so ϕ^{-1} is continuous. Thus, ϕ is a homeomorphism.

Finally, we have

$$Tf(s) = \int_{K_2} (Tf) \, d\delta_s = \int_{K_1} f \, d(T^*\delta_s) = \int_{K_1} f \, d\delta_{\phi(s)} = f(\phi(s)).$$

The factor u appearing in the statement of the theorem does not appear now because of the normalization we made at the beginning of the proof. Had we not assumed

$T(\chi_{K_1}) = \chi_{K_2}$, then we would have $u = T(\chi_{K_1})$, which is a function with the property that $|T(\chi_{K_1})(s)| = 1$ for all $s \in K_2$.

Remark T here are other ways to prove that K_1 and K_2 are homeomorphic using properties of $C(K_1)$ and $C(K_2)$. For example, it is possible to prove that K_1 and K_2 are homeomorphic using only the fact that $C(K_1)$ and $C(K_2)$ are isomorphic as rings, so that no norm structure is required in the proof. (See [15] for more.)

Exercises

Exercise 5.1 Let a and b be real numbers such that $a < b$. Show explicitly that (a, b) is not a compact set in $\mathbb{R}$ by finding an open cover with no finite subcover. (Use the standard topology on $\mathbb{R}$.)

Exercise 5.2 Prove the following theorems:

(a) Every closed subset of a compact space is compact.
(b) The image of a compact space under a continuous map is compact.
(c) Every compact subset of a Hausdorff space is closed.
(d) Let X be a compact space and Y be a Hausdorff space. If $f : X \to Y$ is continuous, then f is a *closed map*. (That is, $f(C)$ is closed in Y whenever C is closed in X.)

Exercise 5.3 Let K be a compact topological space and suppose $\phi : K \to E$ is a continuous one-to-one map, where E is a Hausdorff topological space. Show that ϕ is a homeomorphism onto its image $\phi(K)$. (*Hint:* See Exercise 5.2.)

Exercise 5.4 Let X and Y be topological spaces. If $\phi : X \to Y$ is a continuous closed bijection, show that ϕ is a homeomorphism.

Exercise 5.5 Let X be a Hausdorff space. If A and B are disjoint compact subsets of X, then show there exist disjoint open sets U and V such that $A \subseteq U$ and $B \subseteq V$.

Exercise 5.6 Suppose X is a compact Hausdorff space. Show that X is a *normal* space. That is, if E and F are disjoint closed subsets of X, show there exist disjoint open sets U and V such that $E \subseteq U$ and $F \subseteq V$. (*Hint:* Use Exercise 5.5.)

Exercise 5.7 Suppose X is a compact Hausdorff space. Use Urysohn's Lemma (Theorem 5.65) and Exercise 5.6 to show that $C(X)$ separates the points of X. That is, if a and b are distinct points in X, show there is a function $f \in C(X)$ such that $f(a) = 0$ and $f(b) = 1$.

Exercise 5.8 Show that limits in a Hausdorff space are unique. That is, if X is a Hausdorff space, show that a sequence $(x_n)_{n=1}^{\infty}$ in X cannot converge to two distinct limits x and $\tilde{x}$.

Exercise 5.9 Prove a metric space is second countable if and only if it is separable.

Exercise 5.10 (a) Suppose (M, d) is a metric space and A is a set. If $f : A \to M$ is an injective function, show that $d_A(x, y) = d(f(x), f(y))$ for $(x, y) \in A \times A$ defines a metric on A.
(b) Show that $\rho(x, y) = |\log(y/x)|$ defines a metric on the set $\mathbb{R}^+ = (0, \infty)$.

Exercise 5.11 Let $d(x, y) = |\phi(x) - \phi(y)|$, where $\phi(x) = x/(1 + |x|)$. Show that d is a metric on $\mathbb{R}$ that is not complete.

Exercise 5.12 Show that the space $L_p(0, 1)$, where $0 < p < 1$, is a complete metric space with metric $d(f, g) = \|f - g\|_p^p$, where $\|f\|_p^p = \int_0^1 |f(t)|^p \, dt$. (*Hint:* Use Lemma 2..4, which still applies in this case, despite the fact that $\|\cdot\|_p^p$ is not a norm.)

Exercise 5.13 Let $L_0(0, 1)$ denote the space of all (equivalence classes of) Lebesgue measurable functions on $[0, 1]$. Define

$$d(f, g) = \int_0^1 \min\left(1, |f(s) - g(s)|\right) ds, \quad \{f, g\} \subseteq L_0(0, 1).$$

Prove that d is a metric on $L_0(0, 1)$. Furthermore, show that $d(f_n, f) \to 0$ if and only if $f \to 0$ in measure. Conclude that $L_0(0, 1)$ is a topological vector space (i.e., show that addition and scalar multiplication are continuous).

Exercise 5.14 Show that any continuous linear functional on $L_0(0, 1)$ is identically zero.

Exercise 5.15 Suppose (Ω, μ) is a positive measure space such that $\mu(\Omega) = 1$.

(a) If $0 < p < q \le 1$, then show $\|f\|_p \le \|f\|_q$ for all measurable functions f.
(b) Assume that f is a measurable function such that $\|f\|_r < \infty$ for some $r \le 1$.
   Prove that
   $$\lim_{p \to 0^+} \|f\|_p = \exp\left(\int_\Omega \log|f(\omega)| \, \mu(d\omega)\right),$$

   where we adopt the convention that $e^{-\infty} = 0$.

(Compare to Exercise 2.13.)

Exercise 5.16 Suppose (Ω, μ) is a positive measure space such that $\mu(\Omega) = 1$. Let $L_0(\mu)$ denote the space of all (equivalence classes of) μ-measurable functions on Ω. For any measurable function f, define

$$\|f\|_0 = \exp\left(\int_\Omega \log|f(\omega)| \, \mu(d\omega)\right).$$

(See Exercise 5.15.) If $d(f, g) = \|f - g\|_0$ for all measurable functions f and g in $L_0(\mu)$, does d define a metric on $L_0(\mu)$?

Exercise 5.17 Let X be a locally convex topological vector space with η a base of absolutely convex neighborhoods of 0. Verify that the topology on X is generated by the family of Minkowski functionals $\{p_U\}_{U \in \eta}$. Deduce that $x_n \to 0$ in X if and only if $p_U(x_n) \to 0$ for all $U \in \eta$.

Exercise 5.18 Consider the set $\partial B_{\ell_2} = \{x \in \ell_2 : \|x\|_2 = 1\}$. Show that ∂B_{ℓ_2} is closed in the norm topology, but not the weak topology on ℓ_2. (This example shows that the convexity assumption cannot be omitted from Mazur's Theorem.)

Exercise 5.19 Let K be a compact subset of a Hausdorff topological vector space E, and suppose C is a closed subset of E. Show that $C - K = \{x - y : x \in C, y \in K\}$ is a closed subset of E.

Exercise 5.20 Let E be a locally convex topological vector space and suppose K is a closed linear subspace of E. If $x_0 \notin K$, show that there exists a continuous linear functional $f \in E^*$ such that $f(x_0) = 1$, but $f(x) = 0$ for all $x \in K$.

Exercise 5.21 Let E be a real locally convex topological vector space. Suppose K is a nonempty compact convex subset of E, and C is a nonempty closed convex subset of E, and that $K \cap C = \emptyset$. Show there is a continuous linear functional ϕ on E such that

$$\inf_{x \in C} \phi(x) > \sup_{y \in K} \phi(y).$$

(We say ϕ *separates* K and C.)

Exercise 5.22 Let X be a real Banach space and let E be a weak*-closed subspace of X^*. If ϕ is a weak* continuous linear functional on E with $\|\phi\| = 1$, show for any $\epsilon > 0$ there exists an $x \in X$ with $\|x\| < 1 + \epsilon$ such that $\phi(e^*) = e^*(x)$ for all $e^* \in E$. (*Hint:* Consider the sets $C = \{e^* \in E : \phi(e^*) = 1\}$ and $K = (1+\epsilon)^{-1} B_{X^*}$.)

Exercise 5.23 Let $(X, \|\cdot\|)$ be a real reflexive Banach space and let $\phi \in X^*$. Define a map $f : X \to \mathbb{R}$ by $f(x) = \frac{1}{2}\|x\|^2 - \phi(x)$ for all $x \in X$. Show that f attains a minimum value.

Exercise 5.24 Let $(f_n)_{n=1}^\infty$ be a bounded sequence in $C[0, 1]$. Show that $f_n(s) \to 0$ for every $s \in [0, 1]$ if and only if $f_n \to 0$ weakly.

Exercise 5.25 Let $p \in [1, \infty)$ and for each $n \in \mathbb{N}$ let e_n be the sequence with 1 in the nth coordinate, and 0 elsewhere. Show that the sequence $(ne_n)_{n=1}^\infty$ does not converge weakly to 0 in ℓ_p. (Compare to Example 5.28.)

Exercise 5.26 Let $p \in (1, \infty)$ and for each $n \in \mathbb{N}$ define a function $f_n : [0, 1] \to \mathbb{R}$ by $f_n(x) = n^{1/p} \chi_{[0,1/n]}(x)$ for all $x \in [0, 1]$. Show that $f_n \to 0$ weakly in $L_p(0, 1)$, but not in norm. (Recall that χ_A is the *characteristic function* of the measurable set A.)

Exercise 5.27 Suppose K is a real compact Hausdorff space. Show that the set $\mathcal{P}(K)$ of regular Borel probability measures on K is a convex and w^*-closed subset of $M(K)$, the set of regular Borel measures on K. Show that $\mathcal{P}(K)$ is an extremal set in the unit ball of $M(K)$. (See Sect. 5.8.)

Exercise 5.28 Let K be a compact Hausdorff space and let ν be a Borel measure on K so that $\|\nu\|_{M(K)} \le 1$ and $\nu(K) = 1$. Show that ν is a probability measure.

Exercise 5.29 Let G be a group that is also a topological space. Show that G is a topological group if and only if the map $g : G \times G \to G$ defined by $g(x, y) = x^{-1}y$ is continuous.

Exercise 5.30 Let X be a real separable Banach space. Show that B_{X^*} is metrizable in the weak* topology. (*Hint:* Let $(x_n)_{n=1}^{\infty}$ be a countable dense subset in X and define $\phi(x^*) = (x^*(x_n))_{n=1}^{\infty} \in \mathbb{R}^{\mathbb{N}}$.)

Exercise 5.31 Let X be a Banach space. If $x \in X$, use the Banach–Alaoglu Theorem to prove that there exists an element $x^* \in X^*$ such that $\|x^*\| = 1$ and $x^*(x) = \|x\|$. (*Note:* We proved this in Proposition 3.29 using the Hahn–Banach Theorem.)

Exercise 5.32 A subset E of a topological vector space X is called *bounded* if for every open neighborhood V of 0, there exists an $n \in \mathbb{N}$ such that $E \subseteq nV$. Show that any compact subset of a topological vector space is bounded.

Exercise 5.33 A topological vector space X has the *Heine–Borel property* if every closed and bounded subset of X is compact. (See Exercise 5.32 for the definition of a bounded set in a topological vector space.)

(a) Show that a Banach space has the Heine–Borel property if and only if it is finite-dimensional. (*Hint:* Use Lemma 5.36.)
(b) Show that (X^*, w^*) has the Heine–Borel property if X is a Banach space.

Exercise 5.34 Show that $C[0, 1]$ is not reflexive by showing that $B_{C[0,1]}$ is not compact in the weak topology. (*Hint:* Find a $\Lambda \in C[0, 1]^*$ such that $\Lambda(B_{C[0,1]})$ is open.)

Exercise 5.35 Let X be an infinite-dimensional Banach space. Show that (X^*, w^*) is of the first category in itself.