IV: Topological Spaces

Everyone knows what a curve is, until he has studied enough mathematics to become confused through the countless number of possible exceptions. F. Klein

IV.1 General notions

The abstract notions of limit and convergence are the bread and butter of functional analysis. The purely metric space formulations that we have used thus far are sadly lacking in some cases, so it is necessary to introduce more general concepts. It is possible to describe what is known as a topological space purely in terms of convergence, but it is very awkward. Instead, one usually defines a topological space by abstracting the notion of open sets in metric spaces. Convergence then becomes a derived concept. We discuss convergence in Section IV.2.

This section consists primarily of definitions as we introduce an extensive language needed to describe topological notions. We urge the reader to learn the language by returning to this section when necessary rather than by brute memorization.

Definition A **topological space** is a set S with a distinguished family of subsets $\mathscr{T}$ called **open sets** with the properties:

(i) $\mathscr{T}$ is closed under finite intersections, that is, if A, $B \in \mathscr{T}$, then $A \cap B \in \mathscr{T}$.

(ii) $\mathscr{T}$ is closed under *arbitrary* unions, that is, if $A_\alpha \in \mathscr{T}$ for all α in some index set I, then $\bigcup_{\alpha \in I} A_\alpha \in \mathscr{T}$.
(iii) $\emptyset \in \mathscr{T}$ and $S \in \mathscr{T}$.

$\mathscr{T}$ is called a **topology** for S. We will occasionally write $\langle S, \mathscr{T} \rangle$ for a topological space.

In contradistinction to Borel structure, topological structures are *not* symmetric between intersection and union and involve not merely countable operations but arbitrary operations.

The prime example of a topological space is a metric space. The open sets, $\mathscr{T}$, are those sets, $M \subset S$, with the property $(\forall x \in M)(\exists r > 0)$ $\{y \mid \rho(x, y) < r\} \subset M$. After discussing continuous functions, we will describe another family of examples. We first mention, however, two trivial examples: Given a set S, the family of all subsets of S is a topology; it is called the discrete topology. $\mathscr{T} = \{\emptyset, S\}$ is also a topology; it is called the indiscrete topology.

The family of all topologies on a set S is ordered in a natural way $\mathscr{T}_1 \prec \mathscr{T}_2$ if $\mathscr{T}_1 \subset \mathscr{T}_2$ in the sense of set-theoretic inclusion. If $\mathscr{T}_1 \prec \mathscr{T}_2$, we say $\mathscr{T}_1$ is a **weaker** topology than $\mathscr{T}_2$. (The term weaker comes from the fact that more sequences converge in $\mathscr{T}_1$ than in $\mathscr{T}_2$; so $\mathscr{T}_1$ convergence is a weaker notion than $\mathscr{T}_2$ convergence.)

Definition A family $\mathscr{B} \subset \mathscr{T}$ is called a **base** if and only if any $T \in \mathscr{T}$ is of the form $T = \bigcup_\alpha B_\alpha$ for some family $\{B_\alpha\} \subset \mathscr{B}$.

For example, the balls in a metric space are always a base. We now take a whole family of definitions directly from metric spaces:

Definition A set N is called a **neighborhood** of a point $x \in S$, a topological space, if there exists an open set U with $x \in U \subset N$.

A family $\mathscr{N}$ of subsets of S, a topological space, is called a **neighborhood base at** x if each $N \in \mathscr{N}$ is a neighborhood of x and if given any neighborhood M of x, there is an $N \in \mathscr{N}$ with $N \subset M$. Equivalently, $\mathscr{N}$ is a neighborhood base at x if and only if $\{M \mid N \subset M$ for some $N \in \mathscr{N}\}$ is the family of all neighborhoods of x. For example, if $\mathscr{B}$ is a base for $\mathscr{T}$, $\{N \in \mathscr{B} \mid x \in N\}$ is a neighborhood base at x. We emphasize that neighborhoods need not be open. In a metric space, the *closed* balls of radius greater than zero are a neighborhood base.

Definition A set $C \subset S$, a topological space, is called **closed** if and only if it is the complement of an open set.

The properties of the family of all closed sets can be read off from the properties of $\mathscr{T}$.

Definition Let $A \subset S$, a topological space. The **closure of** A, $\bar{A}$, is the smallest closed set containing A. The **interior of** A, $A°$, is the largest open set contained in A. The **boundary** of A is the set $\bar{A} \backslash A° \equiv \bar{A} \cap \overline{[S \backslash A]}$.

That a smallest closed set containing A exists follows from the fact that $\mathscr{T}$ is closed under arbitrary unions.

As examples, we consider several topologies on $\mathbb{R}^2$:

Example 1 The ordinary metric topology.

Example 2 Consider the family of sets of the form $\{\langle x, y \rangle \mid x \in O\}$ where y is *fixed* and O is an open set of $\mathbb{R}$ in the usual topology. This family of sets is the base for a topology whose open sets are the sets C such that for each $y \in \mathbb{R}$, $\{x \mid \langle x, y \rangle \in C\}$ is open in $\mathbb{R}$ in the usual topology. In an intuitive sense, which we shall shortly make precise, this topology is the "product" of the usual topology in one factor and the discrete topology in the other factor.

Example 3 Let $\mathscr{T}$ consist of the empty set and all sets containing $\langle 0, 0 \rangle$. A neighborhood base for $\langle x, y \rangle$ in this topology is the single (!) set $\{\langle 0, 0 \rangle, \langle x, y \rangle\}$.

Our experience with metric spaces suggests that continuous functions will play a major role.

Definition Let $\langle S, \mathscr{T} \rangle$ and $\langle T, \mathscr{U} \rangle$ be two topological spaces. A function $f: S \to T$ is called **continuous** if $f^{-1}[A] \in \mathscr{T}$ for every $A \in \mathscr{U}$; that is, if the inverse image of any open set is open. f is called **open** if $f[B]$ is open for each $B \in \mathscr{T}$. If f is open and continuous, it is called **bicontinuous**. A bicontinuous bijection is called a **homeomorphism**.

Homeomorphisms are the "isomorphisms" of topological spaces. A topological notion is some notion (or object) invariant under homeomorphism. As an example, the intervals $(-\infty, \infty)$ and $(-1, 1)$ are homeomorphic under the homeomorphism $x \mapsto x/(1 + x^2)$. They are not isometric in the usual metric; in fact, only one of them is complete. This demonstrates that completeness is not a topological notion. However, most metric space notions that are useful in analysis are topological notions.

Continuity is often used to define topologies:

Definition Let $\mathscr{K}$ be a family of functions from a set S to a topological space $\langle T, \mathscr{U}\rangle$. The **$\mathscr{K}$-weak** (or simply **weak**) topology on S is the weakest topology for which all the functions $f \in \mathscr{K}$ are continuous.

To construct the $\mathscr{K}$-weak topology, take the family of all finite intersections of sets of the form $f^{-1}[U]$ where $f \in \mathscr{K}$ and $U \in \mathscr{U}$. These sets form a base for the $\mathscr{K}$-weak topology. If $\mathscr{K}$ is a family of functions on a set S but with values in different topological spaces, we define the $\mathscr{K}$-weak topology in the obvious way.

Example 4 Consider $C[a, b]$, the continuous functions on $[a, b]$. The **topology of pointwise convergence** on $C[a, b]$ is the weak topology given by the family of functions $f \mapsto f(x)$. That is, for each $x \in [a, b]$, let $E_x(f) = f(x)$ so the $E_x(\cdot)$ are maps of $C[a, b]$ to $\mathbb{R}$. As we will see, the topology of pointwise convergence is the topology on $C[a, b]$ for which $f_n \to f$ if and only if $f_n(x) \to f(x)$ for each x.

Example 5 Let $\mathscr{H}$ be a Hilbert space. The "weak topology" is the weakest topology making $\varphi \mapsto (\psi, \varphi)_{\mathscr{H}}$ continuous for each ψ in $\mathscr{H}$. A neighborhood base for 0 is given explicitly by the sets

$$N(\psi_1, \ldots, \psi_n; \varepsilon_1, \ldots, \varepsilon_n) = \{\varphi \,|\, |(\psi_i, \varphi)| < \varepsilon_i, \quad i = 1, \ldots, n\}$$

where $\varepsilon_i > 0$, $\psi_1, \ldots, \psi_n$ are arbitrary, and $n = 1, 2, \ldots$. Thus, the neighborhoods in the weak topology are cylinders in all but finitely many dimensions. That is, there is a subspace M (the orthogonal complement of $\psi_1, \ldots, \psi_n$) whose complement, $M^{\perp}$, is finite dimensional and so that $\varphi \in N, \eta \in M$ implies $\varphi + \eta \in N$.

Example 6 On $\mathbb{R}^2$ consider the maps π_1, π_2 given by $\pi_1(x, y) = x$; $\pi_2(x, y) = y$. The weak topology defined by π_1 and π_2 and the usual topology on $\mathbb{R}$ has rectangles $(a, b) \times (c, d)$ as a base for its open sets and thus the weak topology is the "usual" topology on $\mathbb{R}^2$.

Example 7 The weak topology can be used to topologize Cartesian products. Recall if $\{S_\alpha\}_{\alpha \in I}$ is a family of sets, $S = \times_{\alpha \in I} S_\alpha$ is the family of all $\{x_\alpha\}_{\alpha \in I}$ with $x_\alpha \in S_\alpha$. For each α, we have a map $\pi_\alpha \colon S \to S_\alpha$ given by $\pi_\alpha(\{x_\beta\}_{\beta \in I}) = x_\alpha$. If each S_α has a topology $\mathcal{T}_\alpha$, we define the **product topology**, $\times_{\alpha \in I} \mathcal{T}_\alpha$ as the weak topology generated by the projections π_α.

We now return to our listing of definitions by classifying spaces by how well open sets separate points and closed sets:

Definition

(a) A topological space is called a **T_1 space** if and only if for all x and y, $x \neq y$, there is an open set O with $y \in O$, $x \notin O$. Equivalently, a space is T_1 if and only if $\{x\}$ is closed for each x.

(b) A topological space is called **Hausdorff** (or T_2) if and only if for all x and y, $x \neq y$, there are open sets O_1, O_2 such that $x \in O_1$, $y \in O_2$, and $O_1 \cap O_2 = \emptyset$.

(c) A topological space is called **regular** (or T_3) if and only if it is T_1 and for all x and C, closed, with $x \notin C$, there are open sets O_1, O_2 such that $x_1 \in O_1$, $C \subset O_2$, and $O_1 \cap O_2 = \emptyset$. Equivalently, a space is T_3 if the closed neighborhoods of any point are a neighborhood base.

(d) A topological space is called **normal** (or T_4) if and only if it is T_1 and for all C_1, C_2, closed, with $C_1 \cap C_2 = \emptyset$, there are open sets O_1, O_2 with $C_1 \subset O_1$, $C_2 \subset O_2$, and $O_1 \cap O_2 = \emptyset$.

Obviously:

Proposition $T_4 \Rightarrow T_3 \Rightarrow T_2 \Rightarrow T_1$

We remark that the two most important notions are Hausdorff and normal. At this time, we avoid discussing another way of separating sets, namely with continuous functions. Urysohn's lemma (Theorem IV.7) deals with this question.

We next consider various countability criteria:

Definition

(i) A topological space S is called **separable** if and only if it has a countable dense set.

(ii) A topological space S is called **first countable** if and only if each point $x \in S$ has a countable neighborhood base.

(iii) A topological space S is called **second countable** if and only if S has a countable base.

The relation between these topological notions and metric spaces is set forward in the elementary:

Proposition (a) Every metric space is first countable.
(b) A metric space is second countable if and only if it is separable.
(c) Any second countable topological space is separable.

Warning There are separable spaces that are not second countable (see Problem 7). To add to the confusion, some authors use "separable" to mean second countable. *By separable we always mean that there exists a countable dense set.*

The geometric idea of being connected has a topological formulation:

Definition A topological space S is called **disconnected** if and only if it contains a nonempty proper subset, C, which is both open and closed; equivalently, S is disconnected if and only if it can be written as the union of two disjoint nonempty closed sets. If S is not disconnected, it is called **connected.**

We examine connectivity in Problems 3 and 6. As a final topological notion, we consider restricting topologies to subsets.

Definition Let $\langle S, \mathcal{T} \rangle$ be a topological space and let $A \subset S$. The **relative topology** on A is the family of sets $\mathcal{T}_A = \{O \cap A \mid O \in \mathcal{T}\}$. A subset $B \subset A$ is called **relatively open** if $B \in \mathcal{T}_A$ and **relatively closed** if $A \backslash B \in \mathcal{T}_A$.

IV.2 Nets and convergence

In this section we introduce new objects, called nets, in order to handle limit operations in general topological spaces. Although nets seem on first acquaintance to be bizarre, the propositions in this section show how natural they are.

Definition A **directed system** is an index set I together with an ordering $\prec$ which satisfies:

(i) If $\alpha, \beta \in I$, then there exists $\gamma \in I$ so that $\gamma \succ \alpha$ and $\gamma \succ \beta$.
(ii) $\prec$ is a partial ordering.

Definition A **net** in a topological space S is a mapping from a directed system I to S; we denote it by $\{x_\alpha\}_{\alpha \in I}$.

If we choose the positive integers with the usual order as a directed system, the nets on that directed system are just sequences in S, so nets are a generalization of the notion of sequence. If $P(\alpha)$ is a proposition depending on an index α in a directed set I we say $P(\alpha)$ is **eventually true** if there is a β in I with $P(\alpha)$ true if $\alpha \succ \beta$. We say $P(\alpha)$ is **frequently true** if it is not eventually false, that is, if for any β there is an $\alpha \succ \beta$ with $P(\alpha)$ true.

Definition A net $\{x_\alpha\}_{\alpha \in I}$ in a topological space S is said to **converge** to a point $x \in S$ (written $x_\alpha \to x$) if for any neighborhood N of x, there is a $\beta \in I$ so that $x_\alpha \in N$ if $\alpha \succ \beta$.

Thus $x_\alpha \to x$ if and only if x_α is eventually in any neighborhood of x. If x_α is frequently in any neighborhood of x, we say that x is a **cluster point** of $\{x_\alpha\}$. Notice that the notions of limit and cluster point generalize the same notions for sequences in a metric space.

Theorem IV.1 Let A be a set in a topological space S. Then, a point x is in the closure of A if and only if there is a net $\{x_\alpha\}_{\alpha \in I}$ with $x_\alpha \in A$, so that $x_\alpha \to x$.

Proof We first observe that $\bar{A}$ is just the set of points x such that any neighborhood of x contains a point of A. This set certainly contains A and its complement is the largest open set not containing any points of A. Now suppose $x_\alpha \to x$ where each $x_\alpha \in A$. Then any neighborhood of x contains some x_α and hence some points of A, that is, x is a limit point of A, so $x \in \bar{A}$.

Conversely, suppose $x \in \bar{A}$. Let I be the collection of neighborhoods of x with the ordering $N_1 \prec N_2$ if $N_2 \subset N_1$. For each $N \in I$, let x_N be a point in $A \cap N$. Then $\{x_N\}_{N \in I}$ is a net and $x_N \to x$. ∎

In spaces that are first countable, we can construct the closures of sets by using only sequences. Such is the case in metric spaces. The following example is a case where sequences are not enough:

Example Let $S = [0, 1]$; the nonempty open sets will be the subsets of $[0, 1]$ whose complements contain at most a countable infinity of points. Let $A = [0, 1)$. Then $\bar{A} = S$ since $\{1\}$ is not open. But, let $\{x_n\}_{n=1}^{\infty}$ be any sequence of points of $[0, 1)$. $\{x_n\}_{n=1}^{\infty}$ cannot converge to 1 since the complement of the points $\{x_n\}_{n=1}^{\infty}$ is an open set containing 1.

Although the above example seems artificial, spaces that are not first countable play a large role in functional analysis. Usually, they arise when dual spaces of Banach spaces are considered with topologies weaker than the norm topology (Section IV.5).

We state two facts about nets whose proofs are not difficult and are left as problems:

Theorem IV.2 (a) A function f from a topological space S to a topological space T is continuous if and only if for every convergent net $\{x_\alpha\}_{\alpha \in I}$ in S, with $x_\alpha \to x$, the net $\{f(x_\alpha)\}_{\alpha \in I}$ converges in T to $f(x)$.

(b) Let S be a Hausdorff space. Then a net $\{x_\alpha\}_{\alpha \in I}$ in S can have at most one limit; that is, if $x_\alpha \to x$ and $x_\alpha \to y$, then $x = y$.

Analogous to the concept of a subsequence we have the following definition:

Definition A net $\{x_\alpha\}_{\alpha \in I}$ is a **subnet** of a net $\{y_\beta\}_{\beta \in J}$ if and only if there is a function $F: I \to J$ such that

(i) $x_\alpha = y_{F(\alpha)}$ for each $\alpha \in I$.
(ii) For all $\beta' \in J$, there is an $\alpha' \in I$ such that $\alpha \succ \alpha'$ implies $F(\alpha) \succ \beta'$ (that is, $F(\alpha)$ is eventually larger than any fixed $\beta \in J$).

We then have the following proposition which shows that the above definition is the right one.

Proposition A point x in a topological space S is a cluster point of a net $\{x_\alpha\}$ if and only if some subnet of $\{x_\alpha\}$ converges to x.

Of course, subsequences are subnets of sequences. But it is also possible for a sequence in a topological space to have no convergent subsequences but to have convergent subnets (see Problem 12).

IV.3 Compactness †

The reader no doubt remembers the special role that closed bounded subsets of $\mathbb{R}^n$ played in elementary analysis. In this section we will study the topological abstraction of this concept:

† A supplement to this section begins on p. 351.

Definition We say a topological space $\langle S, \mathcal{T} \rangle$ is **compact** if any open cover of S has a finite subcover, that is, if for any family $\mathcal{U} \subset \mathcal{T}$ with $S = \bigcup_{U \in \mathcal{U}} U$, there is a finite subset $\{U_1, \ldots, U_n\} \subset \mathcal{U}$ with $S = \bigcup_{i=1}^n U_i$. A subset of a topological space is called a **compact set** if it is a compact space in the relative topology.

Henceforth in our discussion we will always suppose that all compact spaces are Hausdorff, although occasionally we will repeat this condition for emphasis.

Since we have a considerable amount of material to discuss, it is perhaps useful to describe briefly the contents of the next two sections. After studying some equivalent formulations of compactness and some elementary properties of compact spaces, we turn to some of the pillars of functional analysis. We first state and discuss Tychonoff's theorem. We then turn to the study of continuous functions on compact sets. After showing that a compact Hausdorff space X has lots of continuous functions (Urysohn's lemma), we discuss the Banach space $C(X)$ of continuous functions. We state the Stone–Weierstrass theorem but defer its instructive proof to an appendix. In the next section, we determine the dual of $C(X)$. Using the Riesz–Markov theorem, we will prove that $C(X)^*$ is identical with $\mathcal{M}(X)$, the family of signed measures on X.

We first reformulate the notion of compactness by taking complements of open sets:

Definition A topological space S is said to have the **finite intersection property** (f.i.p.) if and only if any family of closed sets $\mathcal{F}$ with $\bigcap_{i=1}^n F_i \neq \emptyset$ for any finite subfamily $\{F_i\}_{i=1}^n \subset \mathcal{F}$ satisfies $\bigcap_{F \in \mathcal{F}} F \neq \emptyset$.

Proposition (f.i.p. criterion) S is compact if and only if S has the f.i.p.

Proof Let $\mathcal{F}$ be given and let $\mathcal{U} = \{S \backslash F \mid F \in \mathcal{F}\}$. Then $\mathcal{F}$ has the property that $\bigcap_{i=1}^n F_i \neq \emptyset$ if and only if $\mathcal{U}$ has no finite subcover and the property that $\bigcap_{F \in \mathcal{F}} F \neq \emptyset$ if and only if $\mathcal{U}$ is not a cover. The reader is invited to wend his way through the double negatives to complete the proof. ▌

A somewhat deeper reformulation is:

Theorem IV.3 (The Bolzano–Weierstrass theorem) A space S is compact if and only if every net in S has a convergent subnet.

Proof Suppose that every net has a convergent subnet and let $\mathcal{U}$ be an open cover. Let us suppose that $\mathcal{U}$ has no finite subcover and derive a contradiction.

Order the finite subfamilies $\mathfrak{S}$ of $\mathscr{U}$ by inclusion; $\mathfrak{S}$ is thereby a directed set. For each $\mathscr{F} \equiv \{F_1, \ldots, F_m\} \in \mathfrak{S}$, pick $x_{\mathscr{F}} \notin \bigcup_{i=1}^m F_i$. By assumption, the net $x_{\mathscr{F}}$ has a cluster point x. Since $\mathscr{U}$ is a cover, we can find $U \in \mathscr{U}$ with $x \in U$. Since $x_{\mathscr{F}}$ is frequently in U we can find a finite subfamily $\mathscr{G} \in \mathfrak{S}$ so that $\{U\} \prec \mathscr{G}$ and $x_{\mathscr{G}} \in U$. Since $\{U\} \prec \mathscr{G}$, $U \subset \bigcup_{G \in \mathscr{G}} G$, and so $x_{\mathscr{G}} \in \bigcup_{G \in \mathscr{G}} G$, which is a contradiction.

Suppose that S is compact and let $\{y_\alpha\}_{\alpha \in I}$ be a net. If $\{y_\alpha\}$ has no cluster points, then for any $x \in S$, there is an open set U_x containing x and an $\alpha_x \in I$ with $y_\alpha \notin U_x$ if $\alpha \succ \alpha_x$. The family $\{U_x \mid x \in S\}$ is an open cover of S, so we can find $x_1, \ldots, x_n$ so that $\bigcup_{i=1}^n U_{x_i} = S$. Since I is directed, we can find $\alpha_0 \succ \alpha_{x_i}$ for $i = 1, \ldots, n$. But $y_{\alpha_0} \notin U_{x_i}$, $i = 1, \ldots, n$, which is impossible since $\bigcup_{i=1}^n U_{x_i} = S$. This contradiction establishes that $\{y_\alpha\}_{\alpha \in I}$ has a cluster point and thus a convergent subnet. ∎

Second countable spaces are compact if and only if every sequence has a convergent subsequence (this can be shown by mimicking the above proof).

Example 1 The unit ball in ℓ_2 is not compact in the *metric topology*. No subset of a sequence of orthonormal elements can converge.

Example 2 Let $S = \{\{a_n\} \in \ell_2 \mid |a_n| \le 1/n\}$. It is easy to see that a sequence of elements of S converges if and only if each component converges. Using the diagonalization trick, we conclude that every sequence has a convergent subsequence. Therefore, by the Bolzano–Weierstrass theorem, S is compact.

Warning Compact is not the same as closed and bounded in a general Banach space. In fact the unit ball in a Banach space is compact (in the norm topology) if and only if the space is finite dimensional (see Problem 4 of Chapter V).

We now mention two simple "hereditary" properties of compact spaces (see Problem 38):

Proposition (a) A closed subset of a compact space is compact in the relative topology.

(b) A continuous image of a compact space is compact.

Corollary Any continuous function on a compact space takes on its maximum and minimum values. That is, there are $x_\pm$ so that

$$f(x_+) = \sup_{x \in C} f(x) \qquad \text{and} \qquad f(x_-) = \inf_{x \in C} f(x)$$

The following theorem is often useful:

Theorem IV.4 Let S and T be compact Hausdorff spaces; let $f: S \to T$ be a continuous bijection. Then f is a homeomorphism.

We need the following lemma:

Lemma If T is Hausdorff and $S \subset T$ is compact, then S is closed.

Proof. Let $x \in \bar{S}$. We can find a net $\{x_\alpha\}_{\alpha \in I}$ in S with $x_\alpha \to x$. Since limits are unique in Hausdorff spaces, x is the only cluster point of the net. But since S is compact, the net has a cluster point in S, that is, $x \in S$. Thus $S = \bar{S}$. ∎

Proof of Theorem IV.4 We need only prove f is open or equivalently, since f is a bijection, that $f[C]$ is closed if C is closed. But if $C \subset S$ is closed, then C is compact. By the last proposition, $f[C]$ is compact. The result now follows from the lemma. ∎

Proposition If $\{A_i\}_{i=1}^n$ is a family of compact sets, then $\times_{i=1}^n A_i$ with the product topology is compact.

Proof Let $\{x_\alpha\}_{\alpha \in I}$ be a net in $A = \times_{i=1}^n A_i$, $x_\alpha = \langle x_\alpha^1, x_\alpha^2, \ldots, x_\alpha^n \rangle$. Since A_1 is compact, we can find a subnet $\{x_{\alpha(i)}\}_{i \in D_1}$ so that $\{x_{\alpha(i)}^1\}$ converges to an $x_1 \in A_1$. By a finite induction, we can find a subnet $\{x_{\alpha(i)}\}_{i \in D_n}$ so that $x_{\alpha(i)}^j$ converges to an $x_j \in A_j$ for each j. Then $\{x_{\alpha(i)}\}$ converges in A to $x = \langle x_1, \ldots, x_n \rangle$, so A is compact by the Bolzano–Weierstrass criterion. ∎

This last proposition is not deep; what is deep is that it remains true for an arbitrary product of compact spaces:

Theorem IV.5 (Tychonoff's theorem) Let $\{A_\alpha\}_{\alpha \in I}$ be a collection of compact spaces. Then $\times_{\alpha \in I} A_\alpha$ is compact in the product (that is weak) topology.

Since this theorem has a mildly complicated proof well-treated in the textbook literature, we refer the reader to the references given in the Notes. Let us, however, make several comments. We first remark that it is this theorem that supports the feeling that the weak topology is the "natural" topology for $\times_\alpha A_\alpha$. Another a priori candidate, the "box topology," which is generated by sets of the form $\times_\alpha U_\alpha$, where each U_α is open in A_α is not a topology for which Tychonoff's theorem holds. Secondly, we note that this theorem depends crucially on the axiom of choice (Zorn's lemma). In fact it is known that, set theoretically speaking, Tychonoff's theorem implies Zorn's lemma. Finally, we note that in the special case of countably many *metric* spaces, Theorem IV.5 can be proven by the method of the proposition and the diagonal trick of Section I.5.

Next, we would like to discuss functions on compact Hausdorff spaces. We first show that compact Hausdorff spaces have strong separation properties in the sense of separating closed sets with open sets. We then use these separation properties to construct continuous functions:

Theorem IV.6 Any compact Hausdorff space X is normal (T_4).

Proof We first prove X is regular (T_3). Let $p \in X$ and let $C \subset X$ be closed with $p \notin C$. Since X is Hausdorff, we can find, for any $y \in C$, open and disjoint sets, U_y and V_y, so that $y \in U_y$, and $p \in V_y$. The $\{U_y\}_{y \in C}$ cover C, which is compact. Thus $U_{y_1}, \ldots, U_{y_n}$ cover C. Let $U = \bigcup_{i=1}^n U_{y_i}$; $V = \bigcap_{i=1}^n V_{y_i}$. Then U and V are open and disjoint with $C \subset U$ and $p \in V$. This shows that X is regular. Now let C, D be closed and disjoint. By repeating the above argument with D replacing p and "since X is regular" replacing "since X is Hausdorff," we prove that X is normal. ∎

Normal spaces always have lots of continuous functions for:

Theorem IV.7 (Urysohn's lemma) Let C and D be closed disjoint sets in a normal space, X. Then, there is a continuous function from X to $\mathbb{R}$ with $0 \leq f(x) \leq 1$ for all x such that $f(x) = 0$ if $x \in C$ and $f(x) = 1$ if $x \in D$.

Sketch of proof Using the normality of X, one constructs by induction for each dyadic rational (that is, $r = k/2^n$, k, n integers, $0 \leq k \leq 2^n$) open sets, U_r, with $C \subset U_r \subset \overline{U}_r \subset U_s \subset \overline{U}_s \subset X \backslash D$ if $r < s$. One uses the U_r to define a function with $f(x) < r$ if and only if $x \in U_r$. f can be shown to be continuous. For details, see the references discussed in the notes. ∎

We will see below that one can prove even stronger function theoretic results (Theorem IV.11).

As a final result about the general properties of functions on X we will prove that certain families are dense in $C_{\mathbb{R}}(X)$, the family of all real-valued continuous functions on X. We first note that our proof in Section I.5 for $C[a, b]$ holds on any compact set:

Theorem IV.8 Let $C(X)$ be the family of all continuous complex-valued functions on a compact Hausdorff space, X, endowed with the norm $\|f\|_\infty = \sup_{x \in X} |f(x)|$. Let $C_{\mathbb{R}}(X) = \{f \in C(X) \mid f \text{ is real-valued}\}$. Then $C(X)$ is a complex Banach space and $C_{\mathbb{R}}(X)$ is a real Banach space.

The density theorem we state generalizes a classical theorem of Weierstrass which says that any real-valued continuous function on $[0, 1]$ is a uniform limit (on $[0, 1]$) of polynomials (see Problems 19 and 20). Note that $C_{\mathbb{R}}(X)$ has a natural multiplication given by $(fg)(x) = f(x)g(x)$. A **subalgebra** of $C_{\mathbb{R}}(X)$ is a subspace closed under multiplication:

Theorem IV.9 (Stone–Weierstrass theorem) Let B be a subalgebra of $C_{\mathbb{R}}(X)$ which is closed in $\|\cdot\|_\infty$. We say that B separates points if, given any $x, y \in X$, we can find $f \in B$ with $f(x) \neq f(y)$. If B separates points, then either $B = C_{\mathbb{R}}(X)$ or for some $x_0 \in X$, $B = \{f \in C_{\mathbb{R}}(X) \mid f(x_0) = 0\}$. If $1 \in B$, and B separates points, $B = C_{\mathbb{R}}(X)$.

We defer the instructive lattice-theoretic proof to an appendix.

The fact that we deal with $C_{\mathbb{R}}(X)$ and not $C(X)$ is crucial (see Problem 15), but, by adding an extra hypothesis we can easily extend Theorem IV.9 to the complex case.

Theorem IV.10 (complex Stone–Weierstrass theorem) Let B be a subalgebra of $C(X)$ with the property that if f is in B, then the complex conjugate, $\bar{f}$, is in B also. If B is closed and separates points, then $B = C(X)$ or $B = \{f \mid f(x) = 0\}$ for some fixed x.

The complex conjugate condition is crucial. For example, let D be the unit disc in the complex plane. The functions analytic in the interior of D, continuous on all of D, are a closed subalgebra of D containing 1 and separating points which is not $C(D)$. It is, however, not closed under complex conjugation.

As an example of how to use the Stone–Weierstrass theorem as well as an example of how several functional analytic theorems can combine in a very powerful way, we prove an extension theorem for functions in $C(Y)$ for $Y \subset X$ when X is compact and Y is closed. Actually, this theorem is true if X is merely normal (Problem 18):

Theorem IV.11 (Tietze extension theorem) Let X be a compact space and let $Y \subset X$ be closed. Let f be any continuous real-valued function on Y. Then there is an $\tilde{f} \in C_{\mathbb{R}}(X)$ so that $f(y) = \tilde{f}(y)$ for all $y \in Y$.

Proof Consider the map $\rho: C_{\mathbb{R}}(X) \to C_{\mathbb{R}}(Y)$ given by $\rho(f) = f \upharpoonright Y$. The theorem is equivalent to the statement that ρ is onto. Clearly, Ran ρ is a subalgebra of $C_{\mathbb{R}}(Y)$ and $1 \in \text{Ran}\,\rho$. Moreover, by Urysohn's lemma, Ran ρ separates points. If we can show that Ran ρ is closed in $\|\cdot\|_{C(Y)}$ we can complete the proof by using the Stone–Weierstrass theorem.

Let $I = \text{Ker}\,\rho$. Then I is clearly closed in $C_{\mathbb{R}}(X)$, so we can form the quotient Banach space $C_{\mathbb{R}}(X)/I$. By elementary algebra, ρ "lifts" to a bijection, $\tilde{\rho}: C_{\mathbb{R}}(X)/I \to \text{Ran}\,\rho$. If we can prove $\|\tilde{\rho}([f])\|_{C_{\mathbb{R}}(Y)} = \|[f]\|_{C_{\mathbb{R}}(X)/I}$, Ran ρ will be a Banach space and thus closed.

Clearly, $\|\rho(f)\|_{C_{\mathbb{R}}(Y)} \le \|f\|_{C_{\mathbb{R}}(X)}$, so $\|\tilde{\rho}([f])\|_{C_{\mathbb{R}}(Y)} \le \|[f]\|_{C_{\mathbb{R}}(X)/I}$. Thus, it is enough to show that given $g \in \text{Ran}\,\rho$, we can find $f \in C_{\mathbb{R}}(X)$ with $g = \rho(f)$ and $\|g\|_{C_{\mathbb{R}}(Y)} = \|f\|_{C_{\mathbb{R}}(X)}$ (remember the definition of quotient norm!). Since $g \in \text{Ran}\,\rho$, we know that $g = \rho(h_1)$ for some $h_1 \in C_{\mathbb{R}}(X)$. Let

$$h_2 = \min\{\|g\|_{C_{\mathbb{R}}(Y)}, h_1\}$$

so that $\rho(h_2) = g$ and $h_2(x) \le \|g\|_{C_{\mathbb{R}}(Y)}$ for all x. Let $h_3 = \max\{-\|g\|_{C_{\mathbb{R}}(X)}, h_2\}$. Then, $\|h_3\|_{C_{\mathbb{R}}(X)} = \|g\|_{C_{\mathbb{R}}(Y)}$ and $\rho(h_3) = g$. This completes the proof. ∎

Appendix to IV.3 The Stone–Weierstrass theorem

In this appendix we prove Theorem IV.9 in the case $1 \in B$. The general proof is left as an exercise. Interestingly enough, the first step in the proof is the proof of the classical Weierstrass theorem (which is a special case of the general theorem!)

Lemma 1 The polynomials are dense in $C_{\mathbb{R}}[a, b]$ for any finite real numbers a, b.

Proof See Problems 19 and 20.

This can now be used to prove that B is a lattice, where:

Definition A subset $S \subset C_{\mathbb{R}}(X)$ is called a lattice if for all $f, g \in S$, $f \wedge g = \min\{f, g\}$ and $f \vee g = \max\{f, g\}$ are in S.

Lemma 2 Any closed subalgebra B of $C_{\mathbb{R}}(X)$ with $1 \in B$ is a lattice.

Proof We show that if $f \in B$, then $|f| \in B$. The result then follows from the formulas: $f \vee g = \frac{1}{2}|f - g| + \frac{1}{2}(f + g)$, $f \wedge g = -[(-f) \vee (-g)]$. Without loss suppose that $\|f\|_\infty \leq 1$. By the classical Weierstrass theorem, we can find a sequence of polynomials $P_n(x)$ converging uniformly to $|x|$ on $[-1, 1]$, for example $\big| P_n(x) - |x| \big| < 1/n$ for all x in $[0, 1]$. Since $\|f\|_\infty \leq 1$, it follows that $\|P_n(f) - |f|\,\|_\infty < 1/n$, i.e. $|f| = \lim_{n \to \infty} P_n(f)$. Since B is an algebra with $1 \in B$, $P_n(f) \in B$. Since B is closed, $|f| \in B$. ∎

Finally, the full Stone–Weierstrass theorem is a consequence of Lemma 2 and the following theorem which is of some interest in itself:

Theorem IV.12 (Kakutani–Krein theorem) Let X be a compact Hausdorff space. Any lattice $\mathscr{L} \subset C_{\mathbb{R}}(X)$ which is a closed subspace containing 1 and which separates points is all of $C_{\mathbb{R}}(X)$.

Proof Let $h \in C_{\mathbb{R}}(X)$ and let ε be given. We seek $f \in \mathscr{L}$ with $\|h - f\| < \varepsilon$. Suppose we can show for any $x \in X$, there is $f_x \in \mathscr{L}$ with $f_x(x) = h(x)$ and $h \leq f_x + \varepsilon$. Then for each x, find U_x, an open neighborhood of x with $h(y) \geq f_x(y) - \varepsilon$ for all $y \in U_x$ (by the continuity of $h - f_x$). The U_x cover X so let $U_{x_1}, \ldots, U_{x_n}$ be a subcover. Then $f = f_{x_1} \wedge \cdots \wedge f_{x_m}$ obeys $f(y) + \varepsilon = \min_i\{f_{x_i}(y) + \varepsilon\} \geq h(y)$. Moreover, since any $y \in U_{x_i}$ for some i, $f(y) - \varepsilon \leq f_{x_i}(y) - \varepsilon \leq h(y)$. Thus $\|f - h\|_\infty < \varepsilon$.

It remains to find some f_x with the desired properties. Since $\mathscr{L}$ separates points and $1 \in \mathscr{L}$, for any x and y in X, we can find $f_{xy} \in \mathscr{L}$ with $f_{xy}(x) = h(x)$ and $f_{xy}(y) = h(y)$. For each y, we can find V_y, an open set about y with $f_{xy}(z) + \varepsilon \geq h(z)$ for $z \in V_y$. $V_{y_1}, \ldots, V_{y_n}$ will cover X for suitable $y_1, \ldots, y_n$. If we take $f_x = f_{xy_1} \vee \cdots \vee f_{xy_n}$, then $f_x(x) = h(x)$, and for any $z \in X$

$$f_x(z) + \varepsilon = \max_{i=1,\ldots,n} \{f_{xy_i}(z) + \varepsilon\} \geq h(z)$$

This completes the proof. ∎

IV.4 Measure theory on compact spaces †

In this section, we wish to discuss several aspects of measure theory which are special for compact spaces. In particular, we will see that the dual of $C(X)$ can be interpreted as a space of measures (the Riesz–Markov theorem). Since many of the measure-theoretic proofs are not enlightening, we will not prove all of the theorems.

† A supplement to this section begins on p. 353.

The first question that arises is what to pick as the σ-field of measurable sets. Let us begin with a minimal family. We clearly want to integrate continuous functions $f \in C(X)$. This might lead one to suspect that we want to allow all closed (and open) sets to be measurable but this is not necessary:

Definition A G_δ set is a set which is a countable intersection of open sets.

Proposition Let X be a compact Hausdorff space and let $f \in C_{\mathbb{R}}(X)$. Then $f^{-1}([a, \infty))$ is a compact G_δ set.

Proof $f^{-1}([a, \infty))$ is closed and thus compact. Since

$$f^{-1}([a, \infty)) = \bigcap_{n=1}^{\infty} f^{-1}((a - 1/n, \infty))$$

it is a G_δ. ∎

Thus, to integrate continuous functions, we need only have compact G_δ's in our σ-field.

Definition The σ-field generated by the compact G_δ's in a compact space X is called the family of **Baire sets**. The functions $f\colon X \to \mathbb{R}$ (or $\mathbb{C}$) measurable relative to this σ-field are called **Baire functions**. A measure on the Baire sets is called a **Baire measure** if in addition it is finite, that is $\mu(X) < \infty$.

As in the case of the finite intervals of the real line and Lebesgue measure:

Theorem IV.13 If μ is a Baire measure, then $C(X) \subset L^p(X, d\mu)$ for all p and $C(X)$ is dense in $L^1(X, d\mu)$ or any L^p space for $p < \infty$ (but *not* L^∞ except in pathological cases where $C(X)$ is already all of L^∞!).

Despite the fact that Baire sets are all that are needed, the reader no doubt wants to repress G_δ's and consider all **Borel** sets, i.e. the σ-field generated by all open sets. The question of extending Baire measures to Borel measures, that is, measures on all Borel sets, is answered by the following remarks:

(1) Every Baire measure is automatically **regular**, that is,

$$\begin{aligned}\mu(Y) &= \inf\{\mu(O) \mid Y \subset O,\ O \text{ open and Baire}\}\\ &= \sup\{\mu(C) \mid C \subset Y,\ C \text{ compact and Baire}\}\end{aligned}$$

(2) In general, a Baire measure has many extensions to all Borel sets but there is exactly one regular extension to a Borel measure. A Borel measure is called regular if

$$\mu(Y) = \inf\{\mu(O) \mid Y \subset O,\ O \text{ open}\}$$
$$= \sup\{\mu(C) \mid C \subset O,\ C \text{ compact and } \textit{Borel}\}$$

Thus there is a one-one correspondence between Baire measures and *regular* Borel measures.

(3) If μ is a Borel measure, then $C(X)$ is dense in $L^1(X, d\mu)$ if and only if μ is regular. If μ is regular, every Borel set is almost everywhere a Baire set in the sense that given a Borel set Y, there is a a Baire set $\tilde{Y}$ with

$$\int |\chi_Y - \chi_{\tilde{Y}}|\ d\mu \equiv \mu(Y\backslash\tilde{Y}) + \mu(\tilde{Y}\backslash Y) = 0$$

In addition, every Borel function is equal, after a change on a Borel set of measure zero, to a Baire function.

(4) In certain cases, every compact set is a G_δ, so the Baire and Borel sets are identical. This is the case if X is a compact metric space (see Problem 30).

Henceforth, we will use the word measure in the context of a compact set, X, to mean Baire (or equivalently regular Borel) measure unless we specifically indicate otherwise.

Now, let X be compact and let μ be a measure on X. Consider the map $C(X) \to \mathbb{C}$ given by $f \mapsto \ell_\mu(f) \equiv \int f\, d\mu$. ℓ_μ is clearly linear and

$$|\ell_\mu(f)| \leq \int |f|\ d\mu \leq \|f\|_\infty \mu(X)$$

so ℓ_μ is a continuous linear functional on $C(X)$. In fact, $\|\ell_\mu\|_{C(X)^*} \equiv \mu(X)$, for take $f = 1$. Moreover, ℓ_μ is positive in the sense:

Definition A **positive linear functional** on $C(X)$ is a (not necessarily a priori continuous) linear functional ℓ with $\ell(f) \geq 0$ for all f with $f \geq 0$ pointwise.

In the more general context of C^*-algebras, positive linear functionals will again arise; see Chapter XVII. They have the following nice property (for other properties of positivity, see Problem 37):

Proposition Let ℓ be a positive linear functional. Then ℓ is continuous and $\|\ell\|_{C(X)^*} = \ell(1)$.

Proof Suppose first that f is real. Since $-\|f\|_\infty \le f \le \|f\|_\infty$, we have $-\ell(1)\|f\|_\infty \le \ell(f) \le \ell(1)\|f\|_\infty$; that is, $|\ell(f)| \le \|f\|_\infty \ell(1)$. If f is arbitrary, $\ell(f) = e^{i\phi}r$ with r real and positive, so

$$|\ell(f)| = \ell(\mathrm{Re}[e^{-i\phi}f]) \le \|\mathrm{Re}(e^{-i\phi}f)\|_\infty \ell(1) \le \ell(1)\,\|f\|_\infty \quad \blacksquare$$

We have seen that any Baire measure provides an example of a positive linear functional on $C(X)$; that these are the only examples is the content of:

Theorem IV.14 (the Riesz–Markov theorem) Let X be a compact Hausdorff space. For any positive linear functional ℓ on $C(X)$ there is a unique Baire measure μ on X with

$$\ell(f) = \int f d\mu$$

While we will not give a detailed proof, let us show how μ may be recovered from ℓ_μ. A similar process allows one to construct a measure from any positive linear functional, even if one does not know a priori that it is of the form ℓ_μ. Since μ is inner regular (that is, $\mu(Y) = \sup\{\mu(C) \mid C \subset Y,\ C$ compact$\}$), we need only find $\mu(C)$ for C compact to "recover" μ. We claim $\mu(C) = \inf\{\ell_\mu(f) \mid f \in C(X), f \ge \chi_C\}$. Since μ is positive, it is clear that $\mu(C) \le \ell_\mu(f)$ if $f \ge \chi_C$; thus, we need only show that, given ε, we can find $f \in C(X)$ with $\chi_C \le f$ and $\ell_\mu(f) \le \mu(C) + \varepsilon$. Since μ is outer regular, given ε we can find O open with $\mu(O \backslash C) < \varepsilon$ and $C \subset O$. By Urysohn's lemma, we can find $f \in C(X)$ with $0 \le f \le 1$, $f(x) = 1$ if $x \in C$ and $f(x) = 0$ if $x \in X \backslash O$. Thus $\ell(f) \le \mu(O) < \mu(C) + \varepsilon$. This shows μ can be recovered from ℓ_μ, and so it is not too surprising that a measure can be constructed from an arbitrary ℓ.

The Riesz-Markov theorem is the usual way that measures arise in functional analysis. For example, we have already intimated that measures on $\mathbb{R}$ are associated with quantum mechanical Hamiltonians and they, in turn, arise from certain positive linear functions and the use of the Riesz–Markov theorem (or rather its extension to locally compact spaces which we will discuss shortly).

In general, a pointwise limit of a net of Baire functions is not a Baire or even a Borel function (Problem 13). However, if $\{f_\alpha\}_{\alpha \in I}$ is a net of functions with each f_α continuous and $\{f_\alpha\}$ is increasing in the sense that $f_\alpha \ge f_\beta$ if $\alpha \succ \beta$, then $f = \lim_\alpha f_\alpha = \sup_\alpha f_\alpha$ is a Borel function because

$$f^{-1}[(a, \infty)] = \bigcup_\alpha f_\alpha^{-1}[(a, \infty)]$$

is open. The monotone convergence theorem has the following net generalization:

Theorem IV.15 (monotone convergence theorem for nets) Let μ be a regular Borel measure on a compact Hausdorff space X. Let $\{f_\alpha\}_{\alpha \in I}$ be an increasing net of continuous functions. Then $f = \lim_\alpha f_\alpha \in L^1(X, d\mu)$ if and only if $\sup_\alpha \|f_\alpha\|_1 < \infty$ and in that case $\lim_\alpha \|f - f_\alpha\|_1 = 0$.

Before leaving measure theory on compact spaces, we should identify the dual space of $C(X)$. Of course, not every continuous linear functional on $C(X)$ is a positive linear functional, but the major result we are heading toward is that any $\ell \in C_{\mathbb{R}}(X)^*$ is the difference of two positive linear functionals. This depends on a simple "lattice-theoretic" result about $C_{\mathbb{R}}(X)$:

Lemma Let $f, g \in C_{\mathbb{R}}(X)$ with $f, g \geq 0$. Suppose $h \in C_{\mathbb{R}}(X)$ and $0 \leq h \leq f + g$. Then, we can write $h = h_1 + h_2$ with $0 \leq h_1 \leq f$, $0 \leq h_2 \leq g$, $h_1, h_2 \in C_{\mathbb{R}}(X)$.

Proof Let $h_1 = \min\{f, h\}$. Then $0 \leq h_1 \leq f$ and if $h_2 \equiv h - h_1$, then $h_2 \geq 0$. Moreover, if $h_1(x) = h(x)$, then $h_2(x) = 0 \leq g(x)$ and if $h_1(x) = f(x)$, then $h_2(x) = h(x) - f(x) \leq f(x) + g(x) - f(x) = g(x)$, so $h_2 \leq g$. ∎

Theorem IV.16 Let $\ell \in C_{\mathbb{R}}(X)^*$. Then ℓ can be written $\ell = \ell_+ - \ell_-$ with ℓ_+ and ℓ_- positive linear functionals. Moreover, $\ell_+(1) + \ell_-(1) = \|\ell\|$ and this uniquely determines ℓ_+ and ℓ_-.

Proof For $f \in C(X)_+ \equiv \{f \in C(X) \mid f \geq 0\}$, define $\ell_+(f) = \sup\{\ell(h) \mid h \in C(X);\ 0 \leq h \leq f\}$. Since $|\ell(h)| \leq \|\ell\| \|h\|_\infty \leq \|\ell\| \|f\|_\infty$, this supremum is finite. Clearly $\ell_+(tf) = t\ell_+(f)$ for any scalar $t > 0$ and $\ell_+(f) \geq \ell(0) = 0$ for all $f \in C(X)_+$. Let $f, g \in C(X)_+$. Then, by the lemma:

$$\begin{aligned} \ell_+(f+g) &= \sup\{\ell(h) \mid 0 \leq h \leq f+g\} \\ &= \sup\{\ell(h_1) + \ell(h_2) \mid 0 \leq h_1 \leq f,\quad 0 \leq h_2 \leq g\} \\ &= \ell_+(f) + \ell_+(g) \end{aligned}$$

For any $f \in C(X)$, define $f_+ = \max\{f, 0\}$ and $f_- = -\min\{f, 0\}$, so $f = f_+ - f_-$. Define $\ell_+(f) = \ell_+(f_+) - \ell_+(f_-)$. It is then easy to show ℓ_+ is linear on $C(X)$. By definition $\ell_+(f) \geq \ell(f)$ if $f \geq 0$ so $\ell_-(f) \equiv \ell_+(f) - \ell(f)$ is a positive linear functional. We have thus written $\ell = \ell_+ - \ell_-$ as the difference of positive linear functionals.

To prove $\ell_+(1) + \ell_-(1) = \|\ell\|$, we note first $\|\ell\| \leq \|\ell_+\| + \|\ell_-\| = \ell_+(1) + \ell_-(1)$. For the inequality in the other direction, we first rewrite ℓ_- in a way symmetric to ℓ_+. For $f \geq 0$

$$\begin{aligned} \ell_-(f) &= \sup\{\ell(h) - \ell(f) \mid 0 \leq h \leq f\} \\ &= \sup\{\ell(k) \mid -f \leq k \leq 0\} \end{aligned}$$

where $k = h - f$. Thus:

$$\begin{aligned}\ell_+(1) + \ell_-(1) &= \sup\{\ell(h) \mid 0 \leq h \leq 1\} + \sup\{\ell(k) \mid -1 \leq k \leq 0\}\\ &= \sup\{\ell(g) \mid -1 \leq g \leq 1\}\\ &\leq \|\ell\| \sup\{\|g\|_\infty \mid -1 \leq g \leq 1\}\\ &= \|\ell\|\end{aligned}$$

The proof of uniqueness is left to the reader (Problem 31). ▌

Definition A **complex Baire measure** is a finite linear complex combination of Baire measures.

An easy consequence of Theorem IV.14 and Theorem IV.16 is:

Theorem IV.17 Let X be a compact space. Then the dual $C(X)^*$ of $C(X)$ is the space of all complex Baire measures.

Definition We write $\mathcal{M}(X) = C(X)^*$; $\mathcal{M}_+(X) = \{\ell \in \mathcal{M}(X) \mid \ell$ is a positive linear functional$\}$ and $\mathcal{M}_{+,1}(X) = \{\ell \in \mathcal{M}_+ \mid \|\ell\| = 1\}$.

In some cases, it is important to think of measures not merely as individual objects but instead as elements of $\mathcal{M}(X)$, so that we can employ geometric ideas. To give the reader a feel for this sort of reasoning we conclude our discussion of $\mathcal{M}(X)$ by a simple convexity theorem.

Definition A set A in a vector space Y is called **convex** if x and $y \in A$ and $0 \leq t \leq 1$ implies $tx + (1-t)y \in A$. Thus A is convex if the line segment between x and y is in A whenever x and y are in A (Figure IV.1). A is called a **cone** if $x \in A$ implies $tx \in A$ for all $t > 0$. If A is convex and a cone, it is called a **convex cone**.

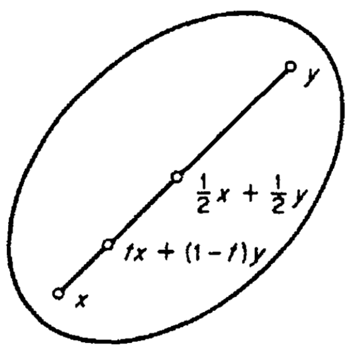

FIGURE IV.1 A convex set.

Proposition Let X be a compact Hausdorff space. Then $\mathcal{M}_{+,1}(X)$ is convex and $\mathcal{M}_{+}(X)$ is a convex cone.

Proof A convex combination of positive linear functionals is clearly a positive linear functional. Moreover, $\|\ell\| = \ell(1)$ if ℓ is a positive linear functional so $\|t\ell_1 + (1-t)\ell_2\| = 1$, if $\ell_1, \ell_2 \in \mathcal{M}_{+,1}$. ∎

At first sight, this geometric fact may appear a little strange since the reader is used to thinking of the unit sphere, $\{x \mid \|x\| = 1\}$ as "round" and here we are saying a piece of it is absolutely flat! The moral is that every norm is not the Euclidean norm (the parallelogram law implies that in a Hilbert space, if $\|x\| = \|y\| = 1$, and $x \neq y$, then $\|tx + (1-t)y\| < 1$). In fact, $\mathbb{R}^n$ with the norm $\|\langle x_1, \ldots, x_n \rangle\| = \sum_{i=1}^n |x_i|$ has a unit sphere with flat faces, see Figure IV.2. This is not a coincidence; $\{1, \ldots, n\}$ is a compact set when given the discrete topology, and $\mathbb{R}^n$ with the norm considered is precisely $\mathcal{M}(\{1, \ldots, n\})$.

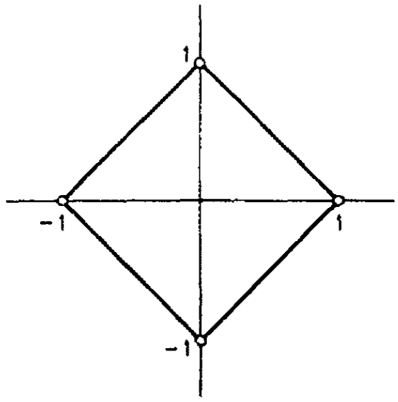

FIGURE IV.2 The unit sphere in R^2 when $\|\langle x, y \rangle\| = |x| + |y|$.

Now, we want to extend "topological measure theory" to a larger class of spaces:

Definition A topological space, X is called **locally compact** if and only if every point $p \in X$, has a compact neighborhood.

By thinking of Lebesgue measure on $\mathbb{R}$, we realize that we want to relax the condition $\mu(X) < \infty$ which we required when X was compact. We first define the Baire sets in X, a locally compact space, to be the σ-ring $\mathcal{B}$ generated by compact G_δ sets. Note that, in general, X may not be a set of $\mathcal{B}$. However, if X is σ-compact, that is, a countable union of compact sets, X is in $\mathcal{B}$.

Definition A **Baire measure** on X, a locally compact space, is a measure on the Baire sets for which $\mu(C) < \infty$ for any compact Baire set C.

Given any Baire measure μ on X, and given a compact G_δ set $C \subset X$, there is induced by restriction a Baire measure μ_C on C. Conversely, it is easy to see that a family of measures $\{\mu_C\}$, one for each compact G_δ set, with the property that $\mu_C(Y) = \mu_D(Y)$ if $Y \subset C \cap D$, defines a Baire measure. This association allows us to prove theorems in the locally compact case from their compact case analogues.

Definition Let X be a locally compact space. $\kappa(X)$, the algebra of **continuous functions of compact support,** is the set of functions that vanish outside some compact set. $C_\infty(X)$, the algebra of continuous functions vanishing at ∞, is the set of $f \in C(X)$ with the property that for any $\varepsilon > 0$, there is a compact set $D_\varepsilon \subset X$ such that $|f(x)| < \varepsilon$ if $x \notin D_\varepsilon$. Thus

$$\kappa(X) \subset C_\infty(X) \subset C(X)$$

With this definition, Theorem IV.14 implies

Theorem IV. 18 (Riesz–Markov) Let X be a locally compact space. A positive linear functional on $\kappa(X)$ is of the form $\ell(f) = \int f\, d\mu$ for some Baire measure, μ. A positive linear functional on $C_\infty(X)$ comes from a measure μ with total finite mass, that is, $\sup_{A \in \mathscr{B}} \mu(A) < \infty$.

In the next chapter, we will find a topology on $\kappa(X)$ for which the dual is just the complex Baire measures. Notice that this topology is *not* given by $\|\cdot\|_\infty$. $\kappa(X)$ is not complete in the norm $\|\cdot\|_\infty$; its completion is $C_\infty(X)$ and its dual in $\|\cdot\|_\infty$ is the *finite* measures.

IV.5 Weak topologies on Banach spaces †

Definition Let X be a Banach space with dual space X^*. The **weak topology on** X is the weakest topology on X in which each functional ℓ in X^* is continuous.

Thus a neighborhood base at zero for the weak topology is given by the sets of the form

$$N(\ell_1, \ldots, \ell_n; \varepsilon) = \{x \mid |\ell_i(x)| < \varepsilon; \quad i = 1, \ldots, n\}$$

† A supplement to this section begins on p. 354.

that is, neighborhoods of zero contain cylinders with finite-dimensional open bases. A net $\{x_\alpha\}$ converges weakly to x, written $x_\alpha \xrightarrow{w} x$, if and only if $\ell(x_\alpha) \to \ell(x)$ for all $\ell \in X^*$.

For infinite dimensional Banach spaces, the weak topology does not arise from a metric. This is one of the main reasons we have introduced topological spaces. Before considering examples, let us note three elementary properties of the weak topology:

Proposition (a) The weak topology is weaker than the norm topology, that is, every weakly open set is norm open.

(b) Every weakly convergent *sequence* is norm bounded.

(c) The weak topology is a Hausdorff topology.

Proof (a) follows from $|\ell(x)| \leq \|\ell\| \, \|x\|$; (b) is a consequence of the uniform boundedness principle; and (c) follows from the Hahn–Banach theorem. We leave the details to the reader. ∎

We emphasize that (b) is only true for sequences. In Problem 39, the reader is asked to construct a counterexample to the analogous net statement.

Let us consider two examples; in both of them, we will describe what it means for sequences to converge. This does *not* completely describe the topology, but it will give the reader an impressionistic view of the underlying topology.

Example 1 Let $\mathscr{H}$ be a Hilbert space. Let $\{\varphi_\alpha\}_{\alpha \in I}$ be an orthonormal basis for $\mathscr{H}$. Given a *sequence* $\psi_n \in \mathscr{H}$, let $\psi_n^{(\alpha)} = \langle \varphi_\alpha, \psi_n \rangle$ be the coordinates of ψ_n. We claim $\psi_n \to \psi$ in the weak topology if and only if (a) $\psi_n^{(\alpha)} \to \psi^{(\alpha)}$ for each α and (b) $\|\psi_n\|$ is bounded. For suppose $\psi_n \xrightarrow{w} \psi$; then (a) follows by definition and (b) comes from (ii) of the proposition. On the other hand, let (a) and (b) hold and let $F \subset \mathscr{H}$ be the subspace of finite linear combinations of the φ_α. By (a), $\langle \varphi, \psi_n \rangle \to \langle \varphi, \psi \rangle$ if $\varphi \in F$. Using the fact that X is dense, (b), and an $\varepsilon/3$ argument, the weak convergence follows.

Example 2 Let X be a compact Hausdorff space and consider the weak topology on $C(X)$. Let $\{f_n\}$ be a sequence in $C(X)$. We claim $f_n \to f$ in the weak topology if and only if (a) $f_n(x) \to f(x)$ for each $x \in X$, and (b) $\|f_n\|$ is bounded. For if $f_n \xrightarrow{w} f$, then (a) holds since $f \mapsto f(x)$ is an element of $C(X)^*$ and (b) comes from (ii) of the proposition. On the other hand, if (a) and (b) hold, then $|f_n(x)| \leq \sup_n \|f_n\|_\infty$ which is L^1 with respect to any Baire measure μ.

Thus, by the dominated convergence theorem, for any $\mu \in \mathcal{M}_+(X)$, $\int f_n \, d\mu \to \int f \, d\mu$. Since any $\ell \in \mathcal{M}(X)$ is a finite linear combination of measures in $\mathcal{M}_+$, we conclude that $f_n \to f$ weakly.

We have seen that the weak topology is weaker than the norm topology; actually, it is very weak indeed! To see this, we note that having few open sets is the same as having few closed sets and this is the same as big closures. In Problem 40, the reader will prove that the weak closure of the unit sphere, $\{x \in X \mid \|x\| = 1\}$, in X is the unit ball, $\{x \mid \|x\| \le 1\}$, in any infinite dimensional Banach space.

We will shortly study general "dual" topologies. As a special case of Theorem IV.20, we state;

Theorem IV.19 A linear functional ℓ on a Banach space is weakly continuous if and only if it is norm continuous.

While this theorem follows from Theorem IV.20, it has a simple direct proof (Problem 42).

Finally, we should like to discuss the weak-$*$ topology and prove a compactness theorem which will often be of use to us. Suppose $Y = X^*$ is the dual of some Banach space X. $Y^* = X^{**}$, of course, induces the weak topology on Y, but we may instead consider the topology induced by X acting on X^*; explicitly:

Definition Let X^* be the dual of a Banach space. The **weak-$*$ topology** is the weakest topology on X^* in which all the functions $\ell \mapsto \ell(x)$, $x \in X$, are continuous.

Notice that the weak-$*$ topology is even weaker than the weak topology. As one might expect, X is reflexive if and only if the weak and weak-$*$ topologies coincide, and many characterizations of reflexivity depend on relations involving the weak and weak-$*$ topologies.

To avoid confusion and to be able to state our next theorem in its natural setting, let us introduce a new notion:

Definition Let X be a vector space and let Y be a family of linear functionals on X which separates points of X. Then the **Y-weak topology on X**, written $\sigma(X, Y)$, is the weakest topology on X for which all the functionals in Y are continuous.

Because Y is assumed to separate points, $\sigma(X, Y)$ is a Hausdorff topology on X. For example, the weak topology on X is the $\sigma(X, X^*)$ topology while the $\sigma(X^*, X)$ topology is the weak-$*$ topology on X^*. The $\sigma(X, Y)$ topology depends only on the vector space generated by Y, so we henceforth suppose that Y is a vector space.

Example The weak-$*$ topology on $\mathcal{M}(X)$, X a compact Hausdorff space, is often called the **vague** topology. To get an idea of how weak it is, let us show the linear combinations of point masses are weak-$*$ dense in $\mathcal{M}(X)$. In Problem 41, the reader is asked to show they are actually norm closed. Suppose that μ is a given measure. We must show that every weak neighborhood of μ contains a sum of point measures, or equivalently, given $f_1, \ldots, f_n$ and ε, that we can find $\alpha_1, \ldots, \alpha_m$ and $x_1, \ldots, x_m$ so that

$$|\mu(f_i) - \sum_{j=1}^{m} \alpha_j f_i(x_j)| < \varepsilon \qquad \text{for} \quad i = 1, \ldots, n$$

For then $\sum \alpha_j \delta_{x_j}$ will be in the vague neighborhood $N(f_1, \ldots, f_n, \varepsilon) + \mu$. Without loss, suppose that $f_1, \ldots, f_n$ are linearly independent. For each x, consider the vector $\mathbf{f}_x = \langle f_1(x), \ldots, f_n(x)\rangle \in \mathbb{R}^n$. If the $\{\mathbf{f}_x\}$ do not span $\mathbb{R}^n$ there is an $\mathbf{a} = \langle a_1, \ldots, a_n\rangle \neq 0 \in \mathbb{R}^n$ with $\mathbf{a} \cdot \mathbf{f}_x = 0$ for all x, that is, $\sum_{i=1}^n a_i f_i = 0$ contradicting linear independence. Thus, the $\mathbf{f}_x$ span $\mathbb{R}^n$. So, we can find $x_1, \ldots, x_n$ and $\alpha_1, \ldots, \alpha_n$ with

$$\langle \mu(f_1), \ldots, \mu(f_n)\rangle = \sum_{i=1}^{n} \alpha_i \mathbf{f}_{x_i}$$

So, $\mu(f_i) = \sum_{j=1}^n \alpha_j f_i(x_j)$, which proves our claim.

The $\sigma(X^*, X)$ topology is of course weaker than the norm topology on X^* so all the $\sigma(X^*, X)$-continuous linear functionals are in X^{**}. In general, however, not all of X^{**} is weak-$*$ continuous on X^*; in fact:

Theorem IV.20 The $\sigma(X, Y)$ continuous linear functionals on X are precisely Y; in particular the only weak-$*$ continuous functionals on X^* are the elements of X.

Proof Suppose that ℓ is a $\sigma(X, Y)$ continuous functional on X. Then $\{x \mid |\ell(x)| < 1\} \supset \{x \mid |y_i(x)| < \varepsilon;\ i = 1, \ldots, n\}$ for some ε and some $y_1, \ldots, y_n \in Y$. Now suppose that $y_i(x) = 0$ for $i = 1, \ldots, n$. Then $|\ell(\varepsilon^{-1}x)| < 1$ for all $\varepsilon > 0$, which implies that $\ell(x) = 0$. As a result, ℓ lifts to a functional $\tilde{\ell}$ on X/K where $K = \{x \mid y_i(x) = 0,\ i = 1, \ldots, n\}$. Elementary abstract algebra shows $\tilde{y}_1, \ldots, \tilde{y}_n$ span the dual space of X/K. Thus $\tilde{\ell} = \sum_{i=1}^n \alpha_i \tilde{y}_i$, so that $\ell = \sum_{i=1}^n \alpha_i y_i \in Y$. ∎

Finally, we conclude this section with its most important result, a result which is perhaps the most important consequence of Tychonoff's theorem:

Theorem IV. 21 (the Banach–Alaoglu theorem) Let X^* be the dual of some Banach space, X. Then the unit ball in X^* is compact in the weak-$*$ topology.

Proof For each $x \in X$, let $B_x = \{\lambda \in \mathbb{C} \mid |\lambda| \le \|x\|\}$. Each B_x is compact, so, by Tychonoff's theorem, $B = \bigtimes_{x \in X} B_x$ is compact in the product topology. Now what is B? An element of B is just an assignment of a number $b(x) \in B_x$ for each x in X, that is, b is a function from X to $\mathbb{C}$ with $|b(x)| \le \|x\|$. In particular, the unit ball $(X^*)_1$ is a subset of B, namely those $b \in B$ which are linear. What is the relative topology induced on $(X^*)_1$ by the product topology on B? It is precisely the weakest topology making $\ell \mapsto \ell(x)$ continuous for each x, that is, the weak-$*$ topology.

Thus, we must only show that $(X^*)_1$ is closed in the product topology. Suppose that ℓ_α is a net in $(X^*)_1$ with $\ell_\alpha \to \ell$. Since $|\ell(x)| \le \|x\|$, we need only show ℓ is linear. But this is easy; if $x,y \in X$ and $\lambda,\mu \in \mathbb{C}$, then

$$\ell(\lambda x + \mu y) = \lim_\alpha \ell_\alpha(\lambda x + \mu y) = \lim_\alpha \lambda\ell_\alpha(x) + \mu\ell_\alpha(y)$$

$$= \lambda\ell(x) + \mu\ell(y) \quad \blacksquare$$

Appendix to IV.5 Weak and strong measurability

In Section II.1, we briefly discussed vector-valued measurable functions with values in an infinite dimensional Hilbert space $\mathscr{H}$. f was called measurable (in Problem 12 of Chapter II) if $(y, f(\cdot))$ was a complex-valued measurable function for each $y \in \mathscr{H}$. This notion might be called **weak measurability**. Another natural candidate for measurability is the a priori stronger notion of measurability which requires that $f^{-1}[C]$ be measurable for each open set $C \subset \mathscr{H}$. *Throughout this book, by a vector-valued measurable function, we will mean a function measurable in the weak sense.* However, to satisfy the reader's natural curiosity, a brief comparison of the various notions of measurability of vector-valued functions seems in order.

Definition Let f be a function on a measure space $\langle M, \mu, \mathscr{R} \rangle$ taking values in a Banach space E.

(i) f is called **strongly measurable** if and only if there is a sequence of functions f_n so that $f_n(x) \to f(x)$ in norm for a.e. $x \in M$ and each f_n takes only finitely many values, each value being taken on a set in $\mathscr{R}$.
(ii) f is called **Borel measurable** if $f^{-1}[C] \in \mathscr{R}$ for each open set C in E (in the metric space topology on E).
(iii) f is called **weakly measurable** if and only if $\ell(f(x))$ is a complex-valued measurable function for each $\ell \in E^*$.

Proposition (a) A pointwise limit of a sequence of Borel measurable functions is a Borel measurable function.

(b) Let f be a function from M to E. If f is strongly measurable, then f is Borel measurable.

(c) Let f be a function from M to E. If f is Borel measurable, it is weakly measurable.

Proof (a) Let $f_n \to f$ pointwise in norm. Let C be an open set in E. Let $C_k = \{x \mid B^x_{k^{-1}} \subset C\}$ where B^x_ε is the ball of radius ε about x. Then,

$$f^{-1}[C] = \bigcup_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m>n} f^{-1}[C_k]$$

so f is Borel measurable.

(b) This is a direct consequence of (a) and the definitions.

(c) The composition of Borel functions is Borel. ∎

Theorem IV.22 Let $\mathscr{H}$ be a separable Hilbert space. Let f be a function from a measure space $\langle M, \mu, \mathscr{R}\rangle$ to $\mathscr{H}$. Then the following three statements are equivalent:

(a) f is strongly measurable.
(b) f is Borel measurable.
(c) f is weakly measurable.

Proof By the last proposition, we need only show that (c) implies (a). Let $\{\psi_n\}_{n=1}^{\infty}$ be an orthonormal basis for $\mathscr{H}$. Let $a_n = (\psi_n, f(x))$. Each a_n is a complex-valued measurable function. It is easy to construct $a_{n,m}(x)$ finite valued, $|a_{n,m}(x)| < |a_n(x)|$ for all x and $\lim_{m\to\infty} a_{n,m}(x) = a_n(x)$ for all $x \in X$. Define $f_N = \sum_{n=1}^{N} a_{n,N}(x)\psi_n$. f_N is finite valued and $f_N \to f$ in norm so f is strongly measurable. ∎

Example Let $\mathbb{C}_t$ be a copy of the complex numbers $\mathbb{C}$ and let $\mathscr{H} = \bigoplus_{t \in \mathbb{R}} \mathbb{C}_t$, that is, $\mathscr{H}$ consists of functions φ on $\mathbb{R}$, nonzero at only countably many t with $\sum_{t \in \mathbb{R}} |\varphi(t)|^2 < \infty$. Let φ_s be given by

$$\varphi_s(t) = \begin{cases} 1 & \text{if } t = s \\ 0 & \text{otherwise} \end{cases}$$

Then $\{\varphi_s\}_{s \in \mathbb{R}}$ is an orthonormal basis for $\mathscr{H}$. Let $f\colon \mathbb{R} \to \mathscr{H}$ be defined by $f(s) = \varphi_s$. For any $\psi \in \mathscr{H}$, $(\psi, f(s)) = 0$ except for a countable set so $(\psi, f(s))$ is measurable. Thus $f(s)$ is weakly measurable. But f is not strongly measurable; for if $f = \lim f_n$ pointwise in norm, then $\operatorname{Ran} f \in \overline{\bigcup \operatorname{Ran} f_n}$. If each f_n were finite valued, $\operatorname{Ran} f$ would be separable, which it is not.

NOTES

Section IV.1 For the reader who wishes to delve further into the realm of general point set topology, we recommend J. Kelley's *General Topology*, Van Nostrand-Reinhold, Princeton, New Jersey, 1955, most enthusiastically. The best way to read the book is to do all the problems; it is time consuming but well worth the effort if the reader can afford the time. Other good references on elementary (and sophisticated) topological notions include: K. Kuratowski, *Topology*, Vol. I, Academic Press, New York, 1966, W. Pervin, *Foundations of General Topology*, Academic Press, New York, 1964, and W. Thron, *Topological Structures*, Holt, New York, 1966.

The notion of topological spaces grew out of work of Fréchet and Hausdorff. The T_1–T_4 classification is due to P. Alexandroff and H. Hopf, in *Topologie I*, Berlin, 1935.

The concept of "Cauchy sequence" does not extend to an arbitrary topological space. However, one can add a "uniform structure" to the topological structure and thereby have spaces in which Cauchy sequence and completeness make sense. One thinks of neighborhoods of x as describing closeness to x. To have a notion of "closeness to x" uniform in x, we need a family $\mathscr{U}$ of subsets of $X \times X$ each containing the set $\Delta = \{\langle x, x\rangle \mid x \in X\}$. We need enough conditions on $\mathscr{U}$ so that $\mathscr{U}_x = \{U_x \mid U \in \mathscr{U}\}$ with $U_x = \{y \mid \langle x, y\rangle \in U\}$ is a neighborhood system for a topology. The canonical example is to let $\mathscr{U}$ be the family of all sets in $X \times X$ containing a set of the form $\{\langle x, y\rangle \mid \rho(x, y) < \varepsilon\}$ with ρ a metric. If G is a topological group (in particular, if G is a topological vector space), there is also a natural uniform structure given by $\mathscr{U} = \{U_N \mid N \in \eta\}$ where η the family of neighborhoods of the identity and $U_N = \{\langle x, y\rangle \mid xy^{-1} \in N\}$. Given a uniform structure $\mathscr{U}$, a net $\{x_\alpha \mid \alpha \in D\}$ is called a Cauchy net if and only if for each $U \in \mathscr{U}$, there is an $\alpha_0 \in D$ so that $\alpha, \beta > \alpha_0$ implies $\langle x_\alpha, x_\beta\rangle \in U$.

The notion of uniform space was first formalized in A. Weil, "Sur les espaces à structure uniforme et sur la topologie générale," *Actualités Sci. Ind.* **551**, Paris (1937). For a modern treatment of uniform spaces, see Kelley, Chapter 6, or G. Choquet, *Lectures on Analysis*, Benjamin, New York, §5.

Section IV.2 Nets were first introduced in E. H. Moore and H. L. Smith, "A General Theory of Limits," *Amer. J. Math.* **44** (1922), 102, and the theory is sometimes called Moore–Smith convergence in the older literature. See Kelley, Chapter 2, for additional discussion.

There is an alternate approach to convergence in topological spaces popularized by Bourbaki. For a discussion of this *theory of filters* see Choquet, §4, or Bourbaki, *Topologie générale*, Chapter 1. We find the filter theory of convergence very unintuitive and prefer the use of nets in all cases.

Section IV.3 It was Tychonoff who realized the utility of the product topology (and proved the Tychonoff theorem) in two fundamental papers: "Über die topologische Erweiterung von Räumen," *Math. Ann.* **102** (1929), 544–556, and "Über einen Funktionenraum," *Math. Ann.* **111** (1935), 762–766. The usual proof of Tychonoff's theorem (c.f. Kelley), depends on the f.i.p. criterion and is a little complicated. The machinery of filters, especially ultrafilters is ideal for a simpler looking proof of the theorem (cf. Choquet). This filter theoretic proof has a net theory translation which we should like to sketch. (1) A net $\{x_\alpha\}$ in a space X is called *universal* if for any $A \subset X$, $x_\alpha \in A$ eventually or $x_\alpha \in X\backslash A$ eventually. Note: A is arbitrary and the definition of universal net makes no mention of topology. (2) If x is a cluster point of a universal net, one has $x_\alpha \to x$, for it cannot happen that $x_\alpha \in A$ frequently without $x_\alpha \in A$ eventually. (3) Any net has a universal subnet. This is the technical heart of the proof and requires the axiom of choice. (4) X is compact if and only if every universal net converges. Given (3), this is just the Bolzano–Weierstrass theorem. (5) To prove Tychonoff's theorem, let $\{x_\alpha\}_{\alpha \in D}$ be a universal net in $\mathsf{X}_{l \in I} A_l$ with each A_l compact. Write $x_\alpha = \{x_\alpha^{(l)}\}_{l \in I}$ with $x_\alpha^{(l)} \in A_l$. Since $\{x_\alpha\}$ is universal, $\{x_\alpha^{(l)}\}$ is universal for each l. Since A_l is compact, $x_\alpha^{(l)} \to x^{(l)}$ for some $x^{(l)} \in A_l$. Let x be the element $\{x^{(l)}\}_{l \in I}$ in $\mathsf{X}_{l \in I} A_l$. Then $x_\alpha \to x$, so every universal net converges. We first learned this proof from O. E. Lanford, III, Les Houches lectures, 1970.

When does a topological space have a topology given by a metric? In general, there is not a simple answer, but for compact Hausdorff spaces, X is metrizable (has a topology given by a metric) if and only if it is second countable. In Section V.2, we see that a similar result holds for topological vector spaces. Both the compact and the vector space results are best understood in the context of uniform spaces; see Kelley, Chapter 6.

K. Weierstrass' original proof of the polynomial approximation theorem can be found on page 5 of Vol. 3 of his *Mathematische Werke*, Mayer and Müller, Berlin, 1903. Stone's generalization first appeared in M. H. Stone, "Applications of the Theory of Boolean Rings to General Topology," *Trans. Amer. Math. Soc.* **41** (1937), 325–481, and a simplified proof was given in his classic article "The Generalized Weierstrass Approximation Theorem," *Math. Mag.* **21** (1947/48), 167–184; 237–254.

Section IV.4 For a brief readable discussion of measure theory on compact spaces we especially recommend the first chapter of L. Nachbin, *The Haar Integral*, Van Nostrand–Reinhold, Princeton, New Jersey, 1965. For a more comprehensive discussion see N. Bourbaki, *Integration*, Chapters 1–4.

Much of our discussion on positive linear functionals goes through for vector spaces with an order allowing finite inf's and sup's, that is for vector lattices. For the deep relations between order notions and topology, see L. Nachbin, *Topology and Order*, Van Nostrand–Reinhold, Princeton, New Jersey, 1965.

For additional discussion of measure theory on locally compact spaces, see the quoted references of Nachbin and Bourbaki, or, for a discussion more similar to our approach, Choquet's book (see notes to Section IV.1).

Section IV.5 We will eventually prove a stronger result than our claim that the linear combinations of Dirac measures are vaguely dense in $\mathcal{M}(X)$. We will actually show that the linear combinations in $\mathcal{M}_{+,1}(X)$ are vaguely dense in $\mathcal{M}_{+,1}(X)$. Thus any positive measure μ with $\mu(X)=1$ can be vaguely approximated by measures $\sum_{n=1}^{N} t_n \delta_{x_n}$ with $0 \le t_n \le 1$, $\sum t_n = 1$. This will follow from the Krein–Milman theorem which we discuss in Section XIV.1. The Banach–Alaoglu theorem was proven in L. Alaoglu: "Weak Topologies of Normed Linear Spaces," *Ann. Math.* **41** (1940), 252–267.

Theorem IV.22 can be extended to an arbitrary separable Banach space. More generally, one has Pettis' theorem: A vector-valued function is strongly measurable if and only if it is weakly measurable and almost separably valued (in the sense that after changing f on a set of measure zero, Ran f is separable). This theorem was first proven in B. J. Pettis, "On Integration in Vector Spaces," *Trans. Amer. Math. Soc.* **44** (1938), 277–304.

One can define the integral of a strongly measurable function by methods analogous to the methods used for real-valued functions. This *Bochner integral* is discussed in K. Yosida, *Functional Analysis*, Springer, New York, 1965 and in many other texts. It was invented by S. Bochner in "Integration von Funktionen, deren Werte die Elemente eines Vektoraumes sind," *Fund. Math.* **20** (1933), 262–276. The Bochner integral obeys a norm dominated convergence theorem. Throughout this book, we use the weak integral defined by $\ell(\int f(x)\,d\mu) = \int \ell(f(x))\,d\mu$. The Bochner integral has nicer properties than this weak integral but we will not need these extra properties so we settle for the simpler weak integral.

PROBLEMS

1. Prove that the family of all topologies on a space is a complete lattice, that is, that any family of topologies has a least upper bound and a greatest lower bound.

2. (Kuratowski closure axioms) Show that the operation $A \mapsto \bar{A}$ in a topological space has the properties:

(i) $\overline{(\bar{A})} = \bar{A}$
(ii) $\overline{A \cup B} = \bar{A} \cup \bar{B}$
(iii) $A \subset \bar{A}$
(iv) $\bar{\emptyset} = \emptyset$

Conversely, suppose that $^{-} : 2^X \to 2^X$ is given ($2^X \equiv$ all subsets of X) obeying (i)–(iv). Show the family of sets B with $\overline{X\backslash B} = X\backslash B$ forms a topology for which the closure operation is $^{-}$.
Reference: Kelley, pp. 42–43.

3. (a) Let 2 be the topological space $\{0, 1\}$ with the discrete topology. Prove that a topological space X is connected if and only if any continuous function $f: X \to 2$ is constant.
(b) Prove that any product of connected spaces is connected.
(c) Let S be a topological space. Suppose that $A, B \subset S$ are connected in the relative topology and $A \cap B \neq \emptyset$; $A \cup B = S$. Show that S is connected.
(d) Let S be a topological space. Suppose that $S = \bar{D}$ and D is connected. Prove that S is connected.
(e) Prove that a continuous image of a connected space is connected.

(f) Prove the intermediate value theorem of freshman calculus, that is, if f is a continuous function on $[a, b]$, then for any $f(a) < x < f(b)$, there is a $c \in [a, b]$ with $f(c) = x$.
Hint: Use (a) to prove (b)–(e).

4. (a) A topological space X is called **Lindelöf** if every open cover has a countable subcover. Prove that any second countable space is Lindelöf.
(b) Prove that a second countable, regular (that is T_3) space is normal (that is T_4).
Reference: Kelley, pp. 49, 113.

5. (a) Prove that $\mathbb{R}$ and $\mathbb{R}^n$ are not homeomorphic for any $n > 1$.
(b) Prove that $\mathbb{R} \neq X \times X$ for any topological space X.
Hint: What happens to $\mathbb{R}$ if a single point is removed?

6. A topological space X is called **arcwise connected** if given $x, y \in X$, there is a continuous function (an arc!) $f\colon [0, 1] \to X$ with $f(0) = x, f(1) = y$.
(a) Show that if X is arcwise connected, it is connected.
(b) Let X_0 be the graph of the function $y = \sin 1/x$ on $\mathbb{R} - \{0\}$, given the relative topology as a subset of the plane. Let $X = X_0 \cup \{\langle x, y\rangle \mid x = 0\}$. Show that Y is connected but not arcwise connected.

7. Let $X = \mathbb{R}$ with the topology $\mathscr{T}$ generated by all sets of the form $\{[a, b) \mid a, b \in \mathbb{R}\}$ which is actually a base for $\mathscr{T}$. Prove that
(a) $\langle X, \mathscr{T}\rangle$ is separable.
(b) $\langle X, \mathscr{T}\rangle$ is first countable.
(c) $\langle X, \mathscr{T}\rangle$ is *not* second countable.

8. Prove that a subspace of a separable metric space is separable.

9. Let Y be $\mathbb{R}^2$ with the product topology given by taking the topology $\mathscr{T}$ of Problem 7 on each factor. Prove that:
(a) Y is separable.
(b) The line $x + y = 1$ is not separable in the relative topology.

10. Let X be any uncountable set and let $\mathscr{T}$ be the topology consisting of $\emptyset$ and complements of finite sets. Prove that
(a) X is separable.
(b) X is compact.
(c) X is T_1 but *not* T_2.
(d) X is neither first nor second countable.

†*11.* Prove Theorem IV.2.

12. Let X be the Banach space l_∞ and consider the sequence $\delta_1, \delta_2, \ldots$ in X^* given by

$$\delta_n(\{c_k\}_{k=1}^{\infty}) = c_n$$

Prove that $\{\delta_n\}, \ldots$ has no weak-$*$ convergent subsequence but that it has a weak-$*$ convergent subnet.

13. Give an example to show that a pointwise limit of a *net* of Borel functions on $\mathbb{R}$ may not be Borel.

14. Show that the space of the example in Section IV.2 is not compact but is Lindelöf (see Problem 4).

15. Let $\mathscr{A}$ be the family of continuous functions on $[0, 2\pi]$ with the property $\int_0^{2\pi} e^{ikx} f(x) = 0$ if k is a negative integer. Prove $\mathscr{A}$ is an algebra which is closed and separates points with $1 \in \mathscr{A}$ but for which $\mathscr{A} \neq C[0, 2\pi]$.

†*16.* Prove the conclusion of the Stone–Weierstrass theorem in the case where we do not suppose $1 \in \mathscr{B}$.

**17.* Let $\mathscr{B}$ be an ideal of $C_{\mathbb{R}}(X)$ which is closed. Let $Y = \{x \in X \mid f(x) = 0 \text{ for all } f \in \mathscr{B}\}$. Prove that Y is closed and that $\mathscr{B} = \{f \in C_{\mathbb{R}}(X) \mid f = 0 \text{ on } Y\}$.

18. Prove the Tietze theorem in the case when X is merely assumed normal. (See the hints given in Kelley, Chapter 7, Problem O.)

19. Let f be a continuous function on $[-\frac{1}{2}, \frac{1}{2}]$ with $f(\frac{1}{2}) = f(-\frac{1}{2}) = 0$. Let $s_k(x)$ be a sequence of functions with $\int_{-1}^{1} s_k(x)\, dx = 1$, each $s_k \geq 0$ so that for any $\delta > 0$,

$$\lim_{k \to \infty} \int_{1 \geq |x| \geq \delta} s_k(x) = 0$$

Prove that

$$\lim_{k \to \infty} \int_{-1/2}^{1/2} s_k(x - y) f(y)\, dy = f(x)$$

for any $x \in [-\frac{1}{2}, \frac{1}{2}]$ and that the convergence is uniform.

20. Let $s_k(x) = (I_k)^{-1}(1 - x^2)^k$ where $I_k = \int_{-1}^{1} (1 - x^2)^k\, dx$. Using Problem 19, prove that any continuous function on $[-\frac{1}{2}, \frac{1}{2}]$ is a limit of polynomials uniformly on $[-\frac{1}{2}, \frac{1}{2}]$.

21. Use the Stone–Weierstrass theorem to prove that:
(a) $\{e^{ikx}\}_{k=-\infty}^{\infty}$ are a complete *orthogonal* set for $L^2[0, 2\pi]$.
(b) The Legendre polynomials are a complete *orthogonal* set for $L^2[-1, 1]$.
*(c) The spherical harmonics are a complete orthonormal set for L^2 of the sphere. (Hint: Use your knowledge of Clebsch–Gordon coefficients!)

22. Prove *Dini's theorem*: Let X be a compact Hausdorff space. Suppose f_n is a monotone decreasing family of functions; let $f_n(x) \to f(x)$ pointwise. Then f_n converges uniformly if and only if f is continuous.

23. Let X be a locally compact Hausdorff space. Consider $\hat{X} = X \cup \{\infty\}$ where ∞ is a "point" not in X. Call $O \subset \hat{X}$ open if either $\infty \notin O$ and O is open in X or $\infty \in O$ and $\hat{X} \backslash O$ is compact. Prove that $\hat{X}$ is a compact Hausdorff space; it is called the **one-point compactification** of X.

24. Prove the Stone–Weierstrass theorem for a locally compact space X: If $\mathscr{A}$ is a closed subalgebra of $C_\infty(X)$, the continuous real-valued functions vanishing at ∞, and if $\mathscr{A}$ separates points and for each $x \in X$, there is $f \in \mathscr{A}$ with $f(x) \neq 0$, then $\mathscr{A} = C_\infty(X)$.

25. Let X be a locally compact Hausdorff space. Prove that for $C, D \subset X$, D closed, C compact, there is a continuous function f, $0 \le f \le 1$, on X with $f[C] = 0$, $f[D] = 1$.
Remark. Use the space $\hat{X}$ of Problem 23 to solve 24 and 25.

26. (a) Prove that any locally compact Hausdorff space is T_3.
(b) Prove that any second countable, locally compact Hausdorff space is normal.
(c) Prove that any σ-compact, locally compact Hausdorff space is normal.
Remark. There exist locally compact spaces which are Hausdorff but not normal, see Kelley, Chapter 4, Problem E.

27. A group G with a topology is called a topological group if the map $\langle x, y\rangle \mapsto xy^{-1}$ of $G \times G \to G$ is jointly continuous. A function f on a topological group G is called uniformly continuous if, for any ε, we can find a neighborhood N_ε of $e \in G$ (the identity) with $|f(x) - f(y)| < \varepsilon$ if $xy^{-1} \in N_\varepsilon$. Prove that any continuous function on a compact topological group is uniformly continuous.

**28.* (a) Let $\mathscr{A}$ be an algebra of real-valued bounded continuous functions on $\mathbb{R}$ which separates points and is closed in $\|\cdot\|_\infty$. Form $X_{\mathscr{A}} \equiv \times_{f \in \mathscr{A}} \{x \in \mathbb{R} \mid |x| \le \|f\|_\infty\}$ with the product topology. Map $\mathbb{R} \to X_{\mathscr{A}}$ by letting x go into the point whose coordinates are $\{f(x)\}_{f \in \mathscr{A}}$. Prove that the image of $\mathbb{R}$ in $X_{\mathscr{A}}$ is homeomorphic to $\mathbb{R}$ if and only if $\mathscr{A}$ contains the functions of compact support.
(b) A topological space X with a map $f: \mathbb{R} \to X$ is called a **compactification** of $\mathbb{R}$ if f is a homeomorphism of $\mathbb{R}$ and its image, if the image is dense in X and if X is a compact Hausdorff space. Two compactifications $f: \mathbb{R} \to X$ and $g: \mathbb{R} \to Y$ are considered identical if there is a homeomorphism $h: X \to Y$ with $h \circ f = g$. Prove that there is a one-one correspondence between compactifications of $\mathbb{R}$ and algebras $\mathscr{A} \subset C_{\mathbb{R}}$ obeying the conditions of (a).
(c) If we take $\mathscr{A} = C(\mathbb{R})$, the compactification we obtain via the construction in (a) is called the **Stone-Čech compactification**, $\check{\mathbb{R}}$. Prove that $\check{\mathbb{R}}$ is a universal compactification of $\mathbb{R}$ in the following sense: Given any compactification $f: \mathbb{R} \to X$ and given the Stone-Čech compactification $g: \mathbb{R} \to \check{\mathbb{R}}$ we can find $h: \check{\mathbb{R}} \to X$ continuous and surjective with $h \circ g = f$.

29. Let $\langle X, d\rangle$ be a metric space with no isolated points. Suppose that every continuous function on X is uniformly continuous. Show that X is compact.

30. (a) Prove that every metric space is normal.
(b) Prove that every closed set in a metric space is a G_δ.

†*31.* Prove the uniqueness statement of Theorem IV.16.

32. Let $\{a_n\}$ be a sequence of numbers with the following property: If $\sum_{n=0}^N \alpha_n x^n \ge 0$ for all $x \in [0, 1]$ then $\sum_{n=0}^N \alpha_n a_n \ge 0$. Prove that there is a unique, (positive) measure μ on $[0, 1]$ with $a_n = \int_0^1 x^n \, d\mu$.

33. Let X be a vector space with Y a family of functionals separating points. Prove that if the $\sigma(X, Y)$ topology comes from a metric, then Y has a countable *algebraic* dimension. An **algebraic basis** for Y is a subset whose *finite* linear combinations span Y. The **algebraic dimension** is the number of elements in a minimal algebraic basis.

34. Let X be a real Banach space and let C be the unit ball of X^* with the weak-$*$ topology. Prove that a continuous function on C can be uniformly approximated by polynomials in the elements of X acting as linear functionals on X^*.

35. Let X be a Banach space, X^* its dual. Let L_n, $n \geq 1$ be elements of X^* with $L_n \to L \in X^*$ in the weak-$*$ sense. Let $x_n \to x$ in norm. Is it necessarily true that $L_n(x_n) \to L(x)$?

36. Prove that X is dense in X^{**} in the $\sigma(X^{**}, X^*)$ topology.

37. Let T: $C(X) \to C(Y)$ be linear. We say T is positivity preserving (or positive) if $Tf > 0$ whenever $f \geq 0$. If T is positive, we write $T \geq 0$. If $S - T \geq 0$ we write $T \leq S$.
(a) Prove that any $T \geq 0$ is automatically continuous and that $\|T\| = \|T1\|_\infty$.
(b) Let S_n be an increasing family of maps. Prove that S_n converges in operator norm if and only if $S_n 1$ converges in function norm.

†*38.* Prove the first proposition in Section IV.2.

†*39.* Find a Banach space and a weakly convergent net which is not norm bounded.

†*40.* Let X be an infinite-dimensional Banach space with the weak topology. Prove that the closure of the unit sphere is the unit ball.

†*41.* Let X be a compact Hausdorff space. Prove that the set of convergent infinite linear combinations of point measures is norm closed in $\mathscr{M}(X)$.

†*42.* Prove Theorem IV.19 directly.

•*43.* (a) Let X be a compact set with a countable basis. Let μ be a Baire measure on X. Prove that $L^p(X, d\mu)$ is separable for all $p < \infty$. (Hint: Let A_n be a countable basis of sets. For all n, m with $\bar{A}_n \cap \bar{A}_m = \emptyset$, find $f_{n,m} \in C(X)$ with $f = 0$ on A_n, $f = 1$ on A_m. Use the $f_{n,m}$ to construct a countable dense set in $C(X)$. Then use the fact that $C(X)$ is dense in $L^p(X, d\mu)$).
(b) Extend the result of (a) to the case where X is only locally compact (Hint: Prove that X is σ-compact).

•*44.* Do any fifty problems in Kelley's book.