David Hilbert
(1862–1943)

David Hilbert was born in Königsberg, East Prussia (now Kaliningrad, Russia), on January 23, 1862. He was the first of two children of Otto and Maria Therese Hilbert.

In 1872, Hilbert entered the Friedrichskolleg Gymnasium (senior secondary school) but transferred in the fall of 1879 to the more science oriented Wilhelm Gymnasium, where he graduated in 1880. Subsequently, he attended the University of Königsberg and received his doctorate there in 1885. In 1886, Hilbert qualified as an unpaid lecturer at the University of Königsberg and acted in this capacity until 1892, when he replaced Adolf Hurwitz as assistant professor. In 1895, he was appointed to chair at the University of Göttingen, where he remained until he retired in 1930.

Influenced by Ferdinand von Lindemann (who proved the transcendence of π), Hilbert's first work was on the theory of invariants. His activity moved from algebraic forms to algebraic number theory, foundations of geometry, analysis (including the calculus of variations and integral equations), theoretical physics, and, finally, to the foundations of mathematics. The invention of the space that bears Hilbert's name grew from his work in the field of integral equations.

The treatise *Der Zahlbericht* (literally, "report on numbers") was begun in 1893 in partnership with Hermann Minkowski (who subsequently abandoned the project). In this report, Hilbert collected, reorganized, and reshaped the information of algebraic number theory into a master work of mathematical literature—for 50 years, *Der Zahlbericht* was the sacred canon of algebraic number theory. Hilbert also wrote *Grundlagen der Geometrie* (Foundations of Geometry), a text first published in 1899 and reaching its ninth edition in 1962, which put geometry in a formal axiomatic setting.

In 1925, Hilbert contracted pernicious anemia, and although he recovered from this illness, he did not resume his full scientific activity. Hilbert died in Göttingen, Germany, on February 14, 1943.

13

Hilbert Spaces and Banach Spaces

The theory of normed spaces applies ideas from linear algebra, geometry, and topology to problems of analysis. In this chapter we will study in detail the most important examples of normed spaces, namely, Hilbert spaces and the classical Banach spaces. These spaces, which are natural generalizations of Euclidean n-space $\mathcal{R}^n$ and unitary n-space $\mathbb{C}^n$, are ubiquitous in analysis. The examples we study in this chapter also serve to motivate some general theorems that appear in Chapter 14.

Section 13.1 discusses preliminaries on normed spaces; Sections 13.2 and 13.3 consider Hilbert spaces and bases and duality of Hilbert spaces; Section 13.4 examines $\mathcal{L}^p$ spaces; and Sections 13.5 and 13.6 investigate nonnegative linear functionals on $C(\Omega)$ and the dual spaces of $C(\Omega)$ and $C_0(\Omega)$.

13.1 PRELIMINARIES ON NORMED SPACES

In this section, we study some elementary properties of normed spaces. Specifically, we examine the relationship between continuity and linearity for mappings of a normed space. We also present a criterion for a normed space to be complete.

In calculus, the following properties of derivative and integral are used so often that their fundamental importance is indisputable:

$$(\alpha f + \beta g)'(t) = \alpha f'(t) + \beta g'(t)$$

$$\int_a^b (\alpha f + \beta g)(t)\, dt = \alpha \int_a^b f(t)\, dt + \beta \int_a^b g(t)\, dt.$$

These two formulas show that differentiation and integration are *linear mappings* on appropriate spaces of functions.

A Course in Real Analysis. Second edition.

DEFINITION 13.1 Linear Mappings, Operators, and Functionals

Let Ω and Λ be linear spaces with the same scalar field. A function $L: \Omega \to \Lambda$ is said to be a **linear mapping** if for all $x, y \in \Omega$ and all scalars α the following two conditions are satisfied:

a) $L(x+y) = L(x) + L(y)$.
b) $L(\alpha x) = \alpha L(x)$.

Linear mappings are also referred to as **linear operators** or **linear transformations;** and in cases where Λ is the scalar field, linear mappings are usually called **linear functionals.**

It follows easily from Definition 13.1 that a linear mapping L takes the **linear combination** $\sum_{j=1}^{n} \alpha_j x_j$ to the linear combination $\sum_{j=1}^{n} \alpha_j L(x_j)$; that is, for each $n \in \mathcal{N}$,

$$L\left(\sum_{j=1}^{n} \alpha_j x_j\right) = \sum_{j=1}^{n} \alpha_j L(x_j)$$

for all $x_1, x_2, \ldots, x_n \in \Omega$ and scalars $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n$.

EXAMPLE 13.1 Illustrates Definition 13.1

a) Let $C_1([0,1])$ denote the collection of all complex-valued functions on $[0,1]$ that have everywhere defined and continuous derivatives. Then the function $D: C_1([0,1]) \to C([0,1])$ defined by $D(f) = f'$ is a linear mapping.
b) The function $J: C([0,1]) \to C([0,1])$ defined by $J(f)(x) = \int_0^x f(t)\,dt$ is a linear operator.
c) The function $\ell: C([0,1]) \to \mathbb{C}$ defined by $\ell(f) = \int_0^1 f(t)\,dt$ is a linear functional.
d) Let A be an $m \times n$ real matrix. Then the function $T: \mathcal{R}^m \to \mathcal{R}^n$ defined by $T(x) = xA$ is a linear mapping. Here xA denotes the product of x with A as matrices, where x is considered a $1 \times m$ matrix. Such mappings are the classical linear transformations studied in linear algebra. □

The next proposition, whose proof is left to the reader as Exercise 13.1, considers the relationship between continuity and linearity of mappings of normed spaces. In the statement of the proposition, as often elsewhere in the text, we use the symbol $\|\ \|$ as a generic norm, letting context determine its exact meaning.

□ □ □ PROPOSITION 13.1

Let Ω and Λ be normed spaces with the same scalar field and $L: \Omega \to \Lambda$ a linear mapping. Then the following properties are equivalent:

a) L is continuous.
b) L is continuous at some point of Ω.
c) L is continuous at 0.
d) $\sup\{\, \|L(x)\| : \|x\| \le 1 \,\} < \infty$.
e) There is a constant c such that $\|L(x)\| \le c\|x\|$ for all $x \in \Omega$.

Part (d) of Proposition 13.1 motivates the definition of a bounded linear mapping, as given in Definition 13.2.

DEFINITION 13.2 **Bounded Linear Mapping**

Suppose Ω and Λ are normed spaces with the same scalar field and that $L: \Omega \to \Lambda$ is a linear mapping. If

$$\|L\| = \sup\{\, \|L(x)\| : \|x\| \leq 1 \,\} < \infty,$$

then L is said to be a **bounded linear mapping.**

Proposition 13.1 shows that a linear mapping is bounded if and only if it is continuous. Note that if L is a bounded linear mapping on Ω, then we have $\|L(x)\| \leq \|L\|\|x\|$ for all $x \in \Omega$.

EXAMPLE 13.2 *Illustrates Definition 13.2*

a) Let Ω be a normed space and let $I: \Omega \to \Omega$ be the identity function, that is, $I(x) = x$ for all $x \in \Omega$. Then I is a bounded linear operator and we have $\|I\| = 1$; I is called the **identity operator** on Ω.
b) The linear operator J defined in Example 13.1(b) is bounded and, in fact, it is easy to show that $\|J\| = 1$.
c) The linear functional ℓ defined in Example 13.1(c) is also bounded and, again, it is easy to show that $\|\ell\| = 1$.
d) The linear mapping D defined in Example 13.1(a) is not bounded if $C_1([0,1])$ is given the norm $\| \; \|_{[0,1]}$. To establish this fact, consider the sequence of functions defined by $s_n(x) = \sin n\pi x$. Clearly, $\|s_n\|_{[0,1]} = 1$. However, because $\|D(s_n)\|_{[0,1]} = n\pi$, it follows that $\|D\| = \infty$.
e) The linear mappings discussed in Example 13.1(d) are bounded, as is implied by Exercise 13.11(b). □

When Ω and Λ are normed spaces with the same scalar field, the collection of all bounded linear operators from Ω to Λ is denoted by $B(\Omega, \Lambda)$. If we define addition and scalar multiplication in $B(\Omega, \Lambda)$ by

$$(L_1 + L_2)(x) = L_1(x) + L_2(x) \qquad \text{and} \qquad (\alpha L_1)(x) = \alpha L_1(x),$$

then $B(\Omega, \Lambda)$ becomes a linear space. Furthermore, $\| \; \|$, as given in Definition 13.2, defines a norm on $B(\Omega, \Lambda)$. See Exercise 13.3.

From now on, unless specified otherwise, we will abbreviate the normed space $(B(\Omega, \Lambda), \| \; \|)$ by $B(\Omega, \Lambda)$. When $\Omega = \Lambda$, we usually denote $B(\Omega, \Lambda)$ by $B(\Omega)$; and when Λ is the scalar field, $B(\Omega, \Lambda)$ is denoted by Ω^* and the norm $\| \; \|$ by $\| \; \|_*$. This latter space has a special name.

DEFINITION 13.3 **Dual Space**

Let Ω be a normed space. Then the space $(\Omega^*, \| \; \|_*)$ of bounded linear functionals on Ω is called the **dual space** of Ω.

The following proposition, whose proof is left to the reader as Exercise 13.6, provides a sufficient condition for the completeness of $B(\Omega, \Lambda)$.

□ □ □ PROPOSITION 13.2

Let Ω and Λ be normed spaces. If Λ is complete, then so is $B(\Omega, \Lambda)$. In particular, the dual space $(\Omega^*, \| \, \|_*)$ is complete.

We will discover that in many notable cases it is possible to find a concrete description of the dual of a normed space. For example, we will prove later that $\ell \in C([0,1])^*$ if and only if there is a unique complex Borel measure μ on $[0,1]$ such that $\ell(f) = \int f \, d\mu$ for all $f \in C([0,1])$.

Banach Spaces

For normed spaces, completeness is a property of such consequence that those possessing it are called *Banach spaces,* after the noted mathematician Stefan Banach. (See the biography at the beginning of Chapter 14 for more on Banach.)

DEFINITION 13.4 Banach Space

A complete normed space is called a **Banach space.**

EXAMPLE 13.3 Illustrates Definition 13.4

a) Exercises 10.59 and 10.60 on page 378 show $\mathcal{R}^n$ and $\mathbb{C}^n$ are Banach spaces.
b) By Proposition 13.2, $B(\Omega, \Lambda)$ is a Banach space whenever Λ is; in particular, Ω^* is always a Banach space.
c) If Ω is a compact topological space, then $C(\Omega)$ is a Banach space.
d) If Ω is locally compact but not compact, then Exercise 11.57(c) on page 422 shows that $C_0(\Omega)$ and $C_b(\Omega)$ are Banach spaces.
e) If $(\Omega, \mathcal{A}, \mu)$ is a measure space, then $\mathcal{L}^\infty(\mu)$ is a Banach space. □

Our next proposition characterizes completeness in normed spaces in terms of infinite series. First let us recall some concepts from Chapter 10. If $\{x_n\}_{n=1}^\infty$ is a sequence of elements in a normed space Ω, then the expression $\sum_{n=1}^\infty x_n$ is called an **infinite series.** The sequence $\{s_n\}_{n=1}^\infty$ of elements of Ω defined by $s_n = \sum_{k=1}^n x_k$ is called the associated **sequence of partial sums.** We say the infinite series **converges** if the sequence of partial sums converges, that is, if $\lim_{n\to\infty} s_n$ exists.

Closely related to the concept of convergence of series is the concept of absolute convergence of series. If $\{x_n\}_{n=1}^\infty$ is a sequence of elements in a normed space Ω, then the infinite series $\sum_{n=1}^\infty x_n$ is said to be **absolutely convergent** or to **converge absolutely** if $\sum_{n=1}^\infty \|x_n\| < \infty$. In the normed space $\mathcal{R}$, a series of nonnegative terms converges if and only if it converges absolutely. On the other hand, the series $\sum_{n=1}^\infty (-1)^n/n$ converges but does not converge absolutely.

From calculus, we know that every absolutely convergent series of real numbers converges. The following proposition shows that this property characterizes Banach spaces.

□ □ □ PROPOSITION 13.3

A normed space Ω is a Banach space if and only if every absolutely convergent series in Ω converges.

PROOF Suppose Ω is a Banach space. Let $\sum_{n=1}^{\infty} x_n$ be an absolutely convergent series. The sequence of partial sums $s_n = \sum_{k=1}^{n} x_n$ satisfies $\|s_n - s_m\| \leq \sum_{k=m+1}^{n} \|x_k\|$ for $m < n$. Therefore, it follows that $\{s_n\}_{n=1}^{\infty}$ is a Cauchy sequence. So, by completeness, $\lim_{n\to\infty} s_n$ exists.

Conversely, suppose that every absolutely convergent series in Ω converges. Let $\{y_n\}_{n=1}^{\infty}$ be a Cauchy sequence. In view of Exercise 10.79 on page 385, to prove that $\{y_n\}_{n=1}^{\infty}$ converges, we need only show that it has a convergent subsequence. By repeatedly applying the Cauchy property, we obtain a subsequence $\{y_{n_k}\}_{k=1}^{\infty}$ such that $\|y_{n_{k+1}} - y_{n_k}\| < 2^{-k}$. Now, let $x_1 = y_{n_1}$ and $x_k = y_{n_k} - y_{n_{k-1}}$ for $k \geq 2$. Then $\sum_{k=1}^{\infty} x_k$ converges absolutely. Because $y_{n_k} = \sum_{j=1}^{k} x_j$, we conclude that $\lim_{k\to\infty} y_{n_k}$ exists. ∎

Exercises for Section 13.1

13.1 Prove Proposition 13.1 on page 454.

13.2 Let $L \in B(\Omega, \Lambda)$, where Ω and Λ are normed spaces. Prove that

$$\|L\| = \sup\{\, \|L(x)\| : \|x\| < 1 \,\} = \sup\{\, \|L(x)\| : \|x\| = 1 \,\}.$$

13.3 Suppose that Ω and Λ are normed spaces. Prove that $\|\ \|$, as defined in Definition 13.2 on page 455, is a norm on the space $B(\Omega, \Lambda)$.

13.4 Let $g \in C([0,1])$, and consider the linear operator $L_g\colon C([0,1]) \to C([0,1])$ defined by $L_g(f) = gf$. Show that L_g is continuous and find $\|L_g\|$.

13.5 Show that each of the following functions is a continuous linear functional on $C([0,1])$ and find its norm:

a) $\ell(f) = f(0)$.

b) $\ell(f) = \int_0^1 f(t)\,dt$.

c) $\ell(f) = \int_0^1 f(t)h(t)\,dt$, where $h \in \mathcal{L}^1([0,1])$.

13.6 Prove Proposition 13.2.

13.7 Let $C_1([0,1])$ be defined as in Example 13.1(a) on page 454.

a) Show that $C_1([0,1])$ is not a closed subspace of $C([0,1])$.

b) Conclude that $C_1([0,1])$ equipped with the norm $\|\ \|_{[0,1]}$ is not a Banach space.

13.8 Show that the space $C_1([0,1])$ defined in Example 13.1(a) on page 454 becomes a Banach space if it is equipped with the norm $\|f\| = |f(0)| + \|f'\|_{[0,1]}$.

13.9 Refer to Example 10.6 on pages 366–367, and let Ω be a nonempty set. Show that the spaces $\ell^1(\Omega)$, $\ell^2(\Omega)$, and $\ell^\infty(\Omega)$ are all Banach spaces.

13.10 Prove that there exist discontinuous linear functionals on any infinite dimensional normed space.

13.11 This exercise shows that all linear mappings on Euclidean n-space or unitary n-space are continuous.

a) Show that all linear functionals on $\mathbb{C}^n$ or $\mathcal{R}^n$ are continuous.

b) Show that all linear mappings from $\mathbb{C}^n$ or $\mathcal{R}^n$ into a normed space are continuous.

13.12 Let S be a linear subspace of the normed space Ω. Prove that if $S^\circ \neq \emptyset$, then $S = \Omega$.

13.13 Let Γ and Λ be normed spaces. Define

$$\|(x,y)\|_1 = \|x\| + \|y\|,$$
$$\|(x,y)\|_2 = (\|x\|^2 + \|y\|^2)^{1/2},$$
$$\|(x,y)\|_\infty = \max\{\|x\|, \|y\|\}.$$

a) Prove that each of the three expressions defines a norm on the Cartesian product space $\Gamma \times \Lambda$.
b) Prove that all three norms are equivalent.

13.14 Let $\| \ \|_1$ be the norm on $C([0,1])$ defined by $\|f\|_1 = \int_0^1 |f(t)|\,dt$.
a) Show that $\|f\|_1 \le \|f\|_{[0,1]}$.
b) Are $\| \ \|_1$ and $\| \ \|_{[0,1]}$ equivalent?

13.2 HILBERT SPACES

Perhaps because they are such natural generalizations of the standard Euclidean space $(\mathcal{R}^n, \| \ \|_2)$, Hilbert spaces appear more frequently in mathematics than other Banach spaces. In addition to being intrinsically important, the theory of Hilbert spaces also merits an extensive discussion because it serves as a model for the general theory of Banach spaces. In this section, we begin our treatment of Hilbert space theory.

DEFINITION 13.5 Inner Product, Inner Product Space

Let $\mathcal{X}$ be a linear space with scalar field F either $\mathcal{R}$ or $\mathbb{C}$. An **inner product** on $\mathcal{X}$ is a function $\langle \ , \ \rangle \colon \mathcal{X} \times \mathcal{X} \to F$ that satisfies the following conditions for all $x, y, z \in \mathcal{X}$ and $\alpha, \beta \in F$:
a) $\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$.
b) $\langle x, y\rangle = \langle y, x\rangle$ if $F = \mathcal{R}$ or $\langle x, y\rangle = \overline{\langle y, x\rangle}$ if $F = \mathbb{C}$.
c) $\langle x, x\rangle \ge 0$.
d) $\langle x, x\rangle = 0$ if and only if $x = 0$.

If $\langle \ , \ \rangle$ is an inner product on $\mathcal{X}$, then the pair $(\mathcal{X}, \langle \ , \ \rangle)$ is called an **inner product space.**

Note: When it is clear from context which inner product is being considered, the inner product space $(\mathcal{X}, \langle \ , \ \rangle)$ will be indicated simply by $\mathcal{X}$. And, although we usually denote an inner product by $\langle \ , \ \rangle$, it is sometimes convenient to have slight variations of this notation such as $\langle \ , \ \rangle_2$ or $[\ , \ ]$.

EXAMPLE 13.4 Illustrates Definition 13.5

a) $\mathbb{C}^n$ is an inner product space if we define

$$\langle z, w\rangle = \sum_{k=1}^{n} z_k \overline{w_k},$$

where $z = (z_1, \ldots, z_n)$ and $w = (w_1, \ldots, w_n)$.

b) $\mathcal{R}^n$ is an inner product space if we define

$$\langle x, y \rangle = \sum_{k=1}^{n} x_k y_k,$$

where $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$. This inner product is the classical "dot product" encountered in vector-calculus courses.

Note: When we consider $\mathbb{C}^n$ or $\mathcal{R}^n$ as an inner product space, we will assume that the inner product is as in this example unless we state otherwise. □

□ □ □ THEOREM 13.1

Let $\mathcal{X}$ be an inner product space. Then, for all $x, y \in \mathcal{X}$,
a) $\langle x+y, x+y \rangle = \langle x, x \rangle + 2\Re\langle x, y \rangle + \langle y, y \rangle$.
b) $|\langle x, y \rangle|^2 \le \langle x, x \rangle \langle y, y \rangle$. **(Cauchy's inequality)**

Moreover, if $y \neq 0$, then equality holds in (b) if and only if $x = \alpha y$ for some scalar α.

PROOF a) From Definition 13.5, we have

$$\begin{aligned}\langle x+y, x+y \rangle &= \langle x, x+y \rangle + \langle y, x+y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \overline{\langle x, y \rangle} + \langle y, y \rangle \\ &= \langle x, x \rangle + 2\Re\langle x, y \rangle + \langle y, y \rangle,\end{aligned} \tag{13.1}$$

as required.

b) If in (13.1) we replace y by $-ty$ where t is a real scalar, then we obtain the polynomial

$$p(t) = \langle x - ty, x - ty \rangle = \gamma + \beta t + \alpha t^2,$$

where $\alpha = \langle y, y \rangle$, $\beta = -2\Re\langle x, y \rangle$, $\gamma = \langle x, x \rangle$. By Definition 13.5(c), we have $p(t) \ge 0$. It follows that $p(t)$ has at most one real root. Thus, $\beta^2 - 4\alpha\gamma \le 0$, that is,

$$(\Re\langle x, y \rangle)^2 \le \langle x, x \rangle \langle y, y \rangle. \tag{13.2}$$

The proof of (b) is now complete in the case of real scalars. If the scalar field is $\mathbb{C}$, we choose $\theta \in [0, 2\pi)$ so that $e^{i\theta}\langle x, y \rangle = |\langle x, y \rangle|$ and use Definition 13.5 and (13.2) to obtain

$$\begin{aligned}|\langle x, y \rangle|^2 &= (\Re\langle e^{i\theta} x, y \rangle)^2 \le \langle e^{i\theta} x, e^{i\theta} x \rangle \langle y, y \rangle \\ &= e^{i\theta} e^{-i\theta} \langle x, x \rangle \langle y, y \rangle = \langle x, x \rangle \langle y, y \rangle.\end{aligned}$$

Therefore, (b) holds in any case.

Suppose now that the scalar field is $\mathcal{R}$, $y \neq 0$, and that equality holds in (b). Then the polynomial $p(t)$ has a root at $t = -\beta/(2\alpha)$. It follows from Definition 13.5(d) that $x = -(\beta/(2\alpha))y$. If the scalar field is $\mathbb{C}$, we choose θ as in the preceding paragraph. Then equality in (b) yields $e^{i\theta} x = -(\beta/(2\alpha))y$ by an argument similar to that used in the real case. ■

We have referred to the inequality in part (b) of Theorem 13.1 as *Cauchy's inequality*. But it is also known as the *Schwarz, Cauchy-Schwarz, Bunyakovski,* or *Cauchy-Bunyakovski-Schwarz (CBS)* inequality.

EXAMPLE 13.5 *Illustrates Definition 13.5 and Theorem 13.1*

Suppose $z_1, \ldots, z_n, w_1, \ldots, w_n \in \mathbb{C}$. Then it follows from Theorem 13.1 and Example 13.4 that

$$\left|\sum_{k=1}^{n} z_k \overline{w_k}\right|^2 \leq \left(\sum_{k=1}^{n} |z_k|^2\right)\left(\sum_{k=1}^{n} |w_k|^2\right).$$

This result is Cauchy's inequality for finite sequences of complex numbers. □

EXAMPLE 13.6 *Illustrates Definition 13.5 and Theorem 13.1*

Refer to Example 10.6(b) on page 367. Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. Recall that $\mathcal{L}^2(\mu)$ consists of all complex-valued $\mathcal{A}$-measurable functions that satisfy $\int_\Omega |f|^2 \, d\mu < \infty$. Also recall that we identify functions that are equal μ-ae. We will show that

$$\langle f, g \rangle = \int_\Omega f \overline{g} \, d\mu \tag{13.3}$$

defines an inner product on $\mathcal{L}^2(\mu)$.

Because of properties of Lebesgue integration that we established in Chapter 5, we need only prove that

$$f, g \in \mathcal{L}^2(\mu) \Rightarrow f\overline{g} \in \mathcal{L}^1(\mu). \tag{13.4}$$

But this follows immediately from the simple inequality $2|f\overline{g}| \leq |f|^2 + |g|^2$.

From now on, whenever we consider $\mathcal{L}^2(\mu)$ in the context of inner product spaces, we will always use the inner product defined by (13.3). □

EXAMPLE 13.7 *Illustrates Definition 13.5 and Theorem 13.1*

Let $(\Omega, \mathcal{A}, P)$ be a probability space. By Example 13.6, the function $\langle \, , \, \rangle$ defined by $\langle X, Y \rangle = \mathcal{E}(XY)$ is an inner product on the space of all random variables with finite variances where, again, we identify two random variables that are equal with probability one. Note that

$$\operatorname{Cov}(X, Y) = \mathcal{E}((X - \mathcal{E}(X))(Y - \mathcal{E}(Y))) = \langle (X - \mathcal{E}(X)), (Y - \mathcal{E}(Y)) \rangle$$

and, in particular, $\operatorname{Var}(X) = \langle (X - \mathcal{E}(X)), (X - \mathcal{E}(X)) \rangle$.

The **correlation coefficient** of two random variables X and Y with finite variances is defined by

$$\rho_{X,Y} = \operatorname{Cov}(X, Y) / \sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}.$$

This quantity is used extensively in probability, statistics, and stochastic processes. From Cauchy's inequality, we see that $-1 \leq \rho_{X,Y} \leq 1$. □

□ □ □ COROLLARY 13.1

Let $\mathcal{X}$ be an inner product space. Define $\| \; \| \colon \mathcal{X} \to \mathcal{R}$ by

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

Then the following properties hold.

a) *The function* $\| \ \|$ *is a norm on* $\mathcal{X}$.

b) *We have*

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

for all $x, y \in \mathcal{X}$.

c) *The inner product is continuous with respect to the product topology induced on* $\mathcal{X} \times \mathcal{X}$ *by the norm* $\| \ \|$.

PROOF a) Definition 10.9 on page 365 gives the three conditions for being a norm. It is easy to check that $\| \ \|$ satisfies the first two conditions. To verify the third condition, we use Theorem 13.1 to conclude that

$$\begin{aligned}\|x+y\|^2 &= \|x\|^2 + 2\Re\langle x, y\rangle + \|y\|^2 \\ &\le \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2.\end{aligned}$$

This display gives the required result.

b) Applying Theorem 13.1 again, we obtain that

$$\|x+y\|^2 = \|x\|^2 + 2\Re\langle x, y\rangle + \|y\|^2$$

and, replacing y by $-y$ in the previous equation, we get

$$\|x-y\|^2 = \|x\|^2 - 2\Re\langle x, y\rangle + \|y\|^2.$$

Adding corresponding sides of the two preceding equalities yields (b).

c) We leave the proof of part (c) to the reader as Exercise 13.15. ■

In the future, we will assume that every inner product space is also a normed space, equipped with the norm defined in Corollary 13.1. If an inner product space is complete, it is called a *Hilbert space* in honor of the mathematician David Hilbert. (See the biography at the beginning of this chapter for more about Hilbert.)

DEFINITION 13.6 Hilbert Space

An inner product space that is complete with respect to its norm is called a **Hilbert space.**

We already know that $\mathcal{R}^n$ and $\mathbb{C}^n$ are Hilbert spaces. Later in this chapter, we will prove that all spaces of the form $\mathcal{L}^2(\mu)$ are Hilbert spaces. But for now, we will content ourselves with knowing that $\mathcal{L}^2(\mu)$-type spaces are inner product spaces, as we showed in Example 13.6.

Nearest Points

The standard Euclidean plane $(\mathcal{R}^2, \| \|_2)$ serves to illustrate an essential property of Hilbert spaces that we will prove in Theorem 13.2. We know that the linear subspaces of $\mathcal{R}^2$ are $\{(0,0)\}$, $\mathcal{R}^2$, and lines passing through $(0,0)$. If L is a line through $(0,0)$ and if $x \in \mathcal{R}^2$, then the point y_0 of intersection of L and the line through x perpendicular to L is the unique point on L that is nearest to x. What is important for us is that y_0 is completely determined by the conditions

$$y_0 \in L \qquad \text{and} \qquad \langle x - y_0, y \rangle = 0 \quad \text{for all} \quad y \in L,$$

as seen in Fig. 13.1.

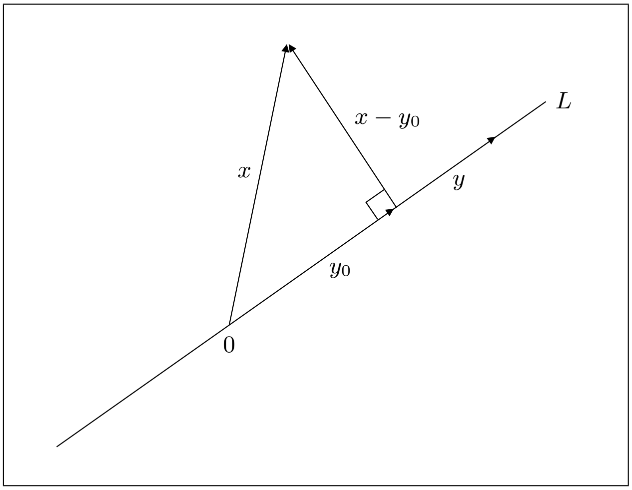

FIGURE 13.1

This property of the Euclidean plane serves to motivate the following important theorem about Hilbert spaces.

□ □ □ THEOREM 13.2

Let $\mathcal{H}$ be a Hilbert space and K a closed linear subspace of $\mathcal{H}$. For each $x \in \mathcal{H}$ there is a unique point $y_0 \in K$ such that

$$\|x - y_0\| = \rho(x, K),$$

where $\rho(x,K) = \inf\{\, \|x - y\| : y \in K \,\}$. Furthermore, the point y_0 is determined by the conditions

$$y_0 \in K \qquad \text{and} \qquad \langle x - y_0, y \rangle = 0 \quad \text{for all} \quad y \in K. \tag{13.5}$$

In other words, (13.5) determines the unique **nearest point** *of K to x.*

PROOF We establish the theorem when the scalar field is $\mathbb{C}$; the proof for real scalars is obtained by a slight modification. To begin, we select a sequence $\{y_n\}_{n=1}^{\infty} \subset K$ such that $\lim_{n\to\infty} \|x - y_n\| = \rho(x, K)$. We claim that $\{y_n\}_{n=1}^{\infty}$ is a Cauchy sequence. Setting $x = x - y_n$ and $y = x - y_m$ in Corollary 13.1, we obtain

$$4\|x - (y_n + y_m)/2\|^2 + \|y_n - y_m\|^2 = 2\|x - y_n\|^2 + 2\|x - y_m\|^2.$$

Since K is a linear subspace, $(y_n + y_m)/2 \in K$. It follows that

$$\|y_n - y_m\|^2 \le 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\rho(x, K)^2. \tag{13.6}$$

Because the right-hand side of (13.6) tends to 0 as $n, m \to \infty$, we conclude that $\{y_n\}_{n=1}^{\infty}$ is a Cauchy sequence.

By completeness, $y_0 = \lim_{n\to\infty} y_n$ exists and, because K is closed, $y_0 \in K$. Moreover,

$$\|x - y_0\| = \lim_{n\to\infty} \|x - y_n\| = \rho(x, K).$$

To verify (13.5), it suffices to consider the case where $y \in K \setminus \{0\}$. Suppose that y_0 is a point of K nearest to x. By Theorem 13.1(a), we have

$$\|x - y_0 - \alpha y\|^2 = \|x - y_0\|^2 - 2\Re\overline{\alpha}\langle x - y_0, y\rangle + |\alpha|^2\|y\|^2$$

for all scalars α. Choosing $\alpha = \langle x - y_0, y\rangle/\|y\|^2$, we obtain

$$\|x - y_0 - \alpha y\|^2 = \|x - y_0\|^2 - |\langle x - y_0, y\rangle|^2/\|y\|^2.$$

Because K is a linear subspace, it follows that $y_0 + \alpha y \in K$. Hence,

$$\|x - y_0\|^2 = \rho(x, K)^2 \le \|x - (y_0 + \alpha y)\|^2 = \|x - y_0\|^2 - |\langle x - y_0, y\rangle|^2/\|y\|^2$$

and, consequently, $\langle x - y_0, y\rangle = 0$.

Suppose, on the other hand, that y_0 is an element of K that satisfies (13.5). Then, for every $y \in K$,

$$\begin{aligned}\|x - y\|^2 &= \|x - y_0 + y_0 - y\|^2 \\ &= \|x - y_0\|^2 + 2\Re\langle x - y_0, y_0 - y\rangle + \|y_0 - y\|^2 \\ &= \|x - y_0\|^2 + \|y_0 - y\|^2 \ge \|x - y_0\|^2.\end{aligned} \tag{13.7}$$

Thus, y_0 is a point of K nearest to x.

It remains to prove that y_0 is unique. Let y_1 be a point of K nearest to x. Then, by (13.7),

$$\|x - y_0\|^2 = \|x - y_1\|^2 = \|x - y_0\|^2 + \|y_0 - y_1\|^2$$

and, therefore, $\|y_0 - y_1\|^2 = 0$. It follows that $y_0 = y_1$. ■

EXAMPLE 13.8 Illustrates Theorem 13.2

a) Let (x_1, y_1), (x_2, y_2), ..., (x_n, y_n) be n points in the plane. In statistics and other fields, it is important to find the straight line that best fits the n points in the sense of minimizing the sum of squared errors. That is, the problem is to find real numbers α and β that minimize

$$\sum_{j=1}^{n} \bigl(y_j - (\alpha + \beta x_j)\bigr)^2.$$

The resulting line is called the *least-squares line* or *regression line.*

We can apply Theorem 13.2 to determine the regression line by proceeding as follows. Let $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n)$, $w = (1, 1, \ldots, 1)$, and $K = \{\, aw + bx : a, b \in \mathcal{R} \,\}$. Finding the regression line is equivalent to obtaining the element y_0 of K nearest to y. Writing $y_0 = \alpha w + \beta x$, we apply (13.5) to get the equations

$$\langle \alpha w + \beta x, w\rangle = \langle y, w\rangle \qquad \text{and} \qquad \langle \alpha w + \beta x, x\rangle = \langle y, x\rangle$$

or, equivalently,

$$n\alpha + \beta \sum_{j=1}^{n} x_j = \sum_{j=1}^{n} y_j \qquad \text{and} \qquad \alpha \sum_{j=1}^{n} x_j + \beta \sum_{j=1}^{n} x_j^2 = \sum_{j=1}^{n} x_j y_j.$$

We thus have two linear equations in the two unknowns α and β. The solution, which we leave to the reader, gives the slope and y-intercept of the regression line.

b) Let μ be the measure on $[-1, 1]$ defined by $\mu(E) = \lambda(E)/2$. The quantity

$$\|f - g\|_2 = \left(\frac{1}{2}\int_{-1}^{1} |f(x) - g(x)|^2 \, dx\right)^{1/2}$$

can be thought of as the average distance between f and g. We will use Theorem 13.2 to find the function of the form $g(x) = \alpha x + \beta$ that minimizes the average distance to $f(x) = x^2$. The function g must satisfy

$$\int_{-1}^{1} (x^2 - \alpha x - \beta)(\gamma x + \delta) \, dx = 0$$

for all $\gamma, \delta \in \mathbb{C}$. A calculation shows that $2(\delta - \alpha\gamma)/3 - 2\beta\delta = 0$ for all γ and δ. It follows that $\alpha = 0$ and $\beta = 1/3$. Thus, the best approximation to x^2 of the form $\alpha x + \beta$ in the sense of the $\mathcal{L}^2(\mu)$-norm is the constant function $g(x) = 1/3$.

c) Refer to Example 13.7. Let $(\Omega, \mathcal{A}, P)$ be a probability space and X a random variable with finite variance. We will use Theorem 13.2 to determine the constant c that minimizes $\mathcal{E}((X - c)^2)$. Applying (13.5) to the subspace generated by the random variable 1, we obtain the equation $\mathcal{E}((X - c)1) = 0$. Thus, $c = \mathcal{E}(X)$ minimizes $\mathcal{E}((X - c)^2)$, and we see that the minimum value is $\mathrm{Var}(X)$. ◻

A close reading of the proof of Theorem 13.2 reveals that more than just that theorem has been established. We did not fully use the assumption that $\mathcal{H}$ is complete; rather, we only needed the completeness of the linear subspace K. The assumption that K is a linear subspace of $\mathcal{H}$ can also be relaxed.

Recall that a subset S of a linear space is said to be a **convex set** if for all $x, y \in S$ and $0 \le \alpha \le 1$, we have $\alpha x + (1-\alpha)y \in S$; in words, whenever S contains two points, it also contains the entire line segment that connects the two points. If C is a closed convex subset, but not necessarily a linear subspace, of a Hilbert space $\mathcal{H}$, then we can still obtain a unique nearest point. However, (13.5) is in general no longer valid. (See Exercise 13.22.)

Theorem 13.2 enables us to associate with each closed linear subspace K of a Hilbert space $\mathcal{H}$ the function $P_K\colon \mathcal{H} \to \mathcal{H}$, where $P_K(x)$ is the point of K nearest to x. The properties of the function P_K are explored in Exercise 13.26 where, in particular, it is shown that it is a bounded linear operator on $\mathcal{H}$ with range K. The operator P_K is often referred to as the **orthogonal projection** of $\mathcal{H}$ onto K.

Orthogonality

From calculus, the ordinary dot product on $\mathcal{R}^2$ satisfies $\langle x, y\rangle = \|x\|\|y\|\cos\theta$, where θ is the angle between x and y. Thus, two vectors in $\mathcal{R}^2$ are perpendicular if and only if their dot product is 0. Similarly, the condition $\langle x, y\rangle = 0$ captures the notion of perpendicularity of two elements of a general inner product space $\mathcal{X}$. The term used for "perpendicular" in the context of inner product spaces is *orthogonal.*

DEFINITION 13.7 Orthogonality

Let $\mathcal{X}$ be an inner product space. Two elements x and y of $\mathcal{X}$ are said to be **orthogonal** if $\langle x, y\rangle = 0$. For a subset S of $\mathcal{X}$, we define the **orthogonal complement** of S, denoted $\boldsymbol{S^\perp}$, to be the set of all elements of $\mathcal{X}$ that are orthogonal to every element of S, that is,

$$S^\perp = \{\, y \in \mathcal{X} : \langle x, y\rangle = 0 \text{ for all } x \in S \,\}.$$

EXAMPLE 13.9 *Illustrates Definition 13.7*

a) The elements $(1,0)$ and $(0,1)$ of $\mathcal{R}^2$ are orthogonal and the orthogonal complement of $\{(1,0)\}$ is $\{\,(0,y) : y \in \mathcal{R}\,\}$.
b) Recall that two random variables with finite variances are said to be *uncorrelated* if $\mathrm{Cov}(X,Y) = 0$. We see from Example 13.7 on page 460 that two random variables are uncorrelated if and only if $X - \mathcal{E}(X)$ and $Y - \mathcal{E}(Y)$ are orthogonal. □

It is left to the reader as Exercise 13.23 to prove that $S^\perp$ is always a closed linear subspace. Moreover, it can be shown that in Hilbert spaces, $(S^\perp)^\perp = \overline{\operatorname{span} S}$, as the reader is asked to verify in Exercise 13.25. Here we are using span S to represent the **span** of S, that is, the linear subspace of all finite linear combinations of elements of S.

Our next result is a version of Theorem 13.2 that emphasizes the role of the orthogonal complement. It also serves as the prototype for an important theorem in the general theory of normed spaces that appears in Chapter 14.

□ □ □ THEOREM 13.3

Let K be a proper closed linear subspace of the Hilbert space $\mathcal{H}$ and $x \in K^c$. Then there exists a unique $z_0 \in K^\perp$ such that $\|z_0\| = 1$ and

$$\begin{aligned}\rho(x, K) &= \inf\{\, \|x - y\| : y \in K \,\} \\ &= \sup\{\, |\langle x, z\rangle| : z \in K^\perp \text{ and } \|z\| \le 1 \,\} = \langle x, z_0\rangle.\end{aligned} \tag{13.8}$$

PROOF Let y_0 be the nearest point of K to x. If $z \in K^\perp$ is such that $\|z\| \le 1$, then, by the definition of $K^\perp$ and Theorem 13.1, we have

$$|\langle x, z\rangle| = |\langle x - y_0, z\rangle| \le \|x - y_0\|\|z\| \le \inf\{\, \|x - y\| : y \in K \,\}. \tag{13.9}$$

It follows that $\inf\{\, \|x - y\| : y \in K \,\} \ge \sup\{\, |\langle x, z\rangle| : z \in K^\perp \text{ and } \|z\| \le 1 \,\}$. Let $z_0 = (x - y_0)/\|x - y_0\|$. By (13.5), $z_0 \in K^\perp$ and, furthermore,

$$\begin{aligned}\inf\{\, \|x - y\| : y \in K \,\} &= \|x - y_0\| = \langle x - y_0, z_0\rangle = \langle x, z_0\rangle \\ &\le \sup\{\, |\langle x, z\rangle| : z \in K^\perp \text{ and } \|z\| \le 1 \,\}.\end{aligned} \tag{13.10}$$

The equations in (13.8) now follow from (13.9) and (13.10). The uniqueness of z_0 is left to the reader as Exercise 13.28. ■

As a visual aid to understanding Theorem 13.3, we have constructed a simple illustration of the theorem in Fig. 13.2.

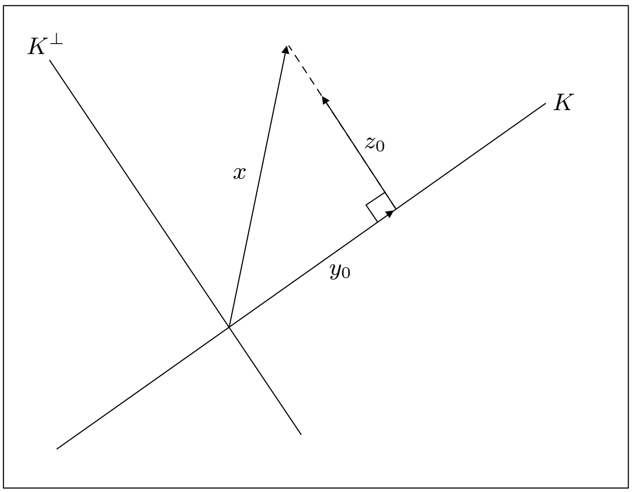

FIGURE 13.2

Exercises for Section 13.2

Note: A ★ denotes an exercise that will be subsequently referenced.

13.15 Prove part (c) of Corollary 13.1 on pages 460–461.

13.16 Let $(\mathcal{X}, \| \ \|)$ be a normed space with scalar field $\mathcal{R}$.
a) Suppose the norm satisfies the identity in Corollary 13.1(b) on page 461. Show that there is an inner product on $\mathcal{X}$ such that $\| \ \|$ is the induced norm.
b) Repeat part (a) in case the scalar field is $\mathbb{C}$.

13.17 A **semi inner product** on a linear space $\mathcal{X}$ is a function $\langle \ , \ \rangle: \mathcal{X} \times \mathcal{X} \to F$ that satisfies conditions (a), (b) and (c) of Definition 13.5 on page 458 and the following weakening of condition (d): $\langle x, x \rangle = 0$ if $x = 0$. Show that (a) and (b) of Theorem 13.1 remain valid for semi inner products.

13.18 Let $\mathcal{X}$ be a linear space with inner product $\langle \ , \ \rangle$ and $L: \mathcal{X} \to \mathcal{X}$ a linear operator. Show that $[x, y] = \langle L(x), L(y) \rangle$ defines a semi inner product on $\mathcal{X}$ in the sense of Exercise 13.17.

13.19 Let Ω be a nonempty set. Prove that $\ell^2(\Omega)$ is a Hilbert space with respect to the inner product given by $\langle f, g \rangle = \int_\Omega f\overline{g}\, d\mu$, where μ is counting measure on Ω.

13.20 Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. Show that if $f \in \mathcal{L}^2(\mu)$, then there is a sequence of simple functions $\{r_n\}_{n=1}^\infty \subset \mathcal{L}^2(\mu)$ such that, as $n \to \infty$, $\|f - r_n\|_2 \to 0$, $\|r_n\|_2 \to \|f\|_2$, and $r_n \to f$ μ-ae.

13.21 Let $\{\mathcal{H}_n\}_{n=1}^\infty$ be a sequence of Hilbert spaces and set

$$\mathcal{H} = \left\{ x \in \bigtimes_{n=1}^\infty \mathcal{H}_n : \sum_{n=1}^\infty \|x_n\|^2 < \infty \right\}.$$

Denote by $\langle \ , \ \rangle$ the inner product for each $\mathcal{H}_n$. Show that $\mathcal{H}$ is a Hilbert space with respect to the inner product defined by $[x, y] = \sum_{n=1}^\infty \langle x_n, y_n \rangle$.

13.22 Let C be a closed convex subset of a Hilbert space $\mathcal{H}$. Show that for each $x \in \mathcal{H}$, there is a unique point $y_0 \in C$ such that $\|x - y_0\| = \rho(x, C)$.

13.23 Let S be a subset of an inner product space $\mathcal{X}$. Show that $S^\perp$ is a closed linear subspace of $\mathcal{X}$.

13.24 Verify the following properties of orthogonal complements:
a) $A \subset B \Rightarrow B^\perp \subset A^\perp$.
b) $A^\perp = (\operatorname{span} A)^\perp$.
c) $D^\perp \cap E^\perp = (D \cup E)^\perp$.

13.25 Prove that in Hilbert spaces, $(A^\perp)^\perp = \overline{\operatorname{span} A}$.

★*13.26* Let K be a closed linear subspace of a Hilbert space $\mathcal{H}$ and P_K the associated orthogonal projection. Verify the following properties.
a) P_K is linear.
b) $\|P_K(x)\| \le \|x\|$, so that P_K is continuous.
c) $P_K \circ P_K = P_K$.
d) $P_K^{-1}(\{0\}) = K^\perp$.
e) The range of P_K is K.
f) $P_{K^\perp} = I - P_K$, where I is the identity operator on $\mathcal{H}$. (See Exercise 13.25.)
g) Deduce from part (f) that each $x \in \mathcal{H}$ can be written uniquely as $x = y + y^\perp$, where $y \in K$ and $y^\perp \in K^\perp$.

13.27 Let y_0 be a nonzero element of a Hilbert space $\mathcal{H}$ and set $K = \operatorname{span}\{y_0\}$. Find an explicit formula for P_K.

13.28 Verify the uniqueness of z_0 in Theorem 13.3.

13.3 BASES AND DUALITY IN HILBERT SPACES

As we know, the concepts of linear independence and basis play an essential role in the theory of finite dimensional linear spaces. In the infinite dimensional case, one can use Zorn's lemma to prove the existence of a **Hamel basis** — a maximal linearly independent set B — and then show that every element of the space can be written uniquely as a finite linear combination of members of B.

Hamel bases are of little use in analysis, however, because they generally cannot be obtained by a formula or constructive process. Fortunately, in Hilbert spaces, there is an analogue of Hamel basis that is much better suited to the needs of analysis. It is this notion of basis to which we now turn our attention.

DEFINITION 13.8 **Orthogonal Set; Orthonormal Set and Basis**

Let $(\mathcal{X}, \langle\ ,\ \rangle)$ be an inner product space. A subset $S \subset \mathcal{X}$ is said to be an **orthogonal set** if every two distinct elements of S are orthogonal, that is, if $\langle x, y\rangle = 0$ for all $x, y \in S$ with $x \neq y$. An orthogonal set S is said to be an **orthonormal set** if $\|x\| = 1$ for each $x \in S$. If S is a nonempty orthonormal set and is contained in no strictly larger orthonormal set, then S is called an **orthonormal basis,** or simply a **basis.**

EXAMPLE 13.10 *Illustrates Definition 13.8*

a) The set of elements $\{(1,0,0,\dots,0),(0,1,0,\dots,0),\dots,(0,0,0,\dots,1)\}$ is an orthonormal set in $\mathbb{C}^n$. Clearly, it is also a basis.
b) Let Ω be a nonempty set. For each $x \in \Omega$, let d_x denote the function that is 1 at x and 0 at all other points of Ω. Then $\{\, d_x : x \in \Omega \,\}$ is an orthonormal set in $\ell^2(\Omega)$. We will see later that it is also an orthonormal basis.
c) For each $n \in \mathcal{Z}$, define $e_n(x) = (2\pi)^{-1/2}e^{inx}$. It is easy to see that the collection of functions $\{\, e_n : n \in \mathcal{Z} \,\}$ is an orthonormal set in $\mathcal{L}^2([-\pi,\pi])$. Later we will show that it is an orthonormal basis as well. □

Our next theorem provides some fundamental properties of orthonormal sets.

□ □ □ THEOREM 13.4

Let $\mathcal{X}$ be an inner product space and $E = \{e_1, e_2, \dots, e_n\}$ a finite orthonormal subset of $\mathcal{X}$. Then the following properties hold.

a) E is linearly independent.
b) $\|\sum_{j=1}^n \alpha_j e_j\|^2 = \sum_{j=1}^n |\alpha_j|^2$ for any choice of scalars $\alpha_1, \alpha_2, \dots, \alpha_n$.
c) For each $x \in \mathcal{X}$, we have $\sum_{j=1}^n |\langle x, e_j\rangle|^2 \le \|x\|^2$.
d) $x = \sum_{j=1}^n \langle x, e_j\rangle e_j$ for each $x \in \operatorname{span} E$.
e) $\operatorname{span} E$ is a complete subspace of $\mathcal{X}$, in particular, a closed subset of $\mathcal{X}$.
f) For each $x \in \mathcal{X}$, $y_0 = \sum_{j=1}^n \langle x, e_j\rangle e_j$ is the unique nearest point of $\operatorname{span} E$ to x, that is, it is the unique member y of $\operatorname{span} E$ such that $\|x - y\| = \rho(x, \operatorname{span} E)$.

PROOF The proofs of (a), (b), and (d) are left to the reader as Exercise 13.30. To prove (c), let $x \in \mathcal{X}$ and $y = \sum_{j=1}^{n} \langle x, e_j \rangle e_j$. By (b), $\|y\|^2 = \sum_{j=1}^{n} |\langle x, e_j \rangle|^2$. Also,

$$\langle x, y \rangle = \langle x, \textstyle\sum_{j=1}^{n} \langle x, e_j \rangle e_j \rangle = \displaystyle\sum_{j=1}^{n} \overline{\langle x, e_j \rangle} \langle x, e_j \rangle = \sum_{j=1}^{n} |\langle x, e_j \rangle|^2.$$

Applying Theorem 13.1(a) on page 459, we now obtain that

$$0 \le \|x - y\|^2 = \|x\|^2 - 2\Re\langle x, y \rangle + \|y\|^2 = \|x\|^2 - \sum_{j=1}^{n} |\langle x, e_j \rangle|^2,$$

from which (c) follows immediately.

To prove (e), let $\{y_m\}_{m=1}^{\infty}$ be a Cauchy sequence in span E. From Cauchy's inequality, we have

$$|\langle y_m, e_k \rangle - \langle y_\ell, e_k \rangle| \le \|y_m - y_\ell\|.$$

Thus, $\{\langle y_m, e_k \rangle\}_{m=1}^{\infty}$ is a Cauchy sequence for $k = 1, 2, \ldots, n$. Applying part (d) and using the completeness of the scalars, we conclude that the limit

$$y = \lim_{m \to \infty} y_m = \sum_{k=1}^{n} \left(\lim_{m \to \infty} \langle y_m, e_k \rangle \right) e_k$$

exists. Clearly, $y \in \operatorname{span} E$. We have now shown that span E is complete. Since a complete subset of a metric space is closed, it follows that span E is closed in $\mathcal{X}$.

Next we establish (f). By Theorem 13.2 on page 462 and the defining properties of inner product, it is enough to show that $\langle x - y_0, e_k \rangle = 0$ for $k = 1, 2, \ldots, n$. Using the fact that E is an orthonormal set, we get

$$\langle x - y_0, e_k \rangle = \langle x, e_k \rangle - \sum_{j=1}^{n} \langle x, e_j \rangle \langle e_j, e_k \rangle = \langle x, e_k \rangle - \langle x, e_k \rangle = 0,$$

as required. ■

As an immediate consequence of Theorem 13.4(c), we get the following important result, known as *Bessel's inequality.* Refer to Exercise 2.37 on page 46 for the meaning of the summation that occurs in that inequality.

□ □ □ COROLLARY 13.2 Bessel's Inequality

Let E be an orthonormal subset of an inner product space $\mathcal{X}$. Then

$$\sum_{e \in E} |\langle x, e \rangle|^2 \le \|x\|^2$$

for all $x \in \mathcal{X}$.

EXAMPLE 13.11 Illustrates Theorem 13.4

In the space $\mathcal{L}^2([-\pi,\pi])$, consider the linear subspace

$$\mathcal{U}_n = \operatorname{span}\{\, e_k : -n \le k \le n \,\},$$

where $e_k(x) = (2\pi)^{-1/2} e^{ikx}$. As we noted in Example 13.10(c) on page 468, $\{\, e_n : n \in \mathcal{Z} \,\}$ is an orthonormal set. Therefore, $\{\, e_k : -n \le k \le n \,\}$ is a finite orthonormal subset of $\mathcal{L}^2([-\pi,\pi])$. It is clear that $\mathcal{U}_n$ is the space of complex trigonometric polynomials of degree at most n.

Let $f \in \mathcal{L}^2([-\pi,\pi])$. Then, from Theorem 13.4(f), the nearest member of $\mathcal{U}_n$ to f is given by

$$s_n = \sum_{|k|\le n} \langle f, e_k\rangle e_k.$$

The number

$$\hat{f}(k) = (2\pi)^{-1/2}\langle f, e_k\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}\,dx$$

is called the kth **Fourier coefficient** of f. Thus, the best approximation,

$$s_n(x) = \sum_{|k|\le n} \langle f, e_k\rangle e_k = \sum_{k=-n}^{n} \hat{f}(k)e^{ikx},$$

is the nth partial sum of the **Fourier series** $\sum_{k=-\infty}^{\infty} \hat{f}(k)e^{ikx}$ associated with the function f. □

More examples of orthonormal sets can be found by using the procedure described in the proof of the following theorem.

□ □ □ THEOREM 13.5

Let $\{x_m\}_{m=1}^{\infty}$ be a sequence of elements in an inner product space $\mathcal{X}$ and assume that $x_1 \ne 0$. Then there is a countable orthonormal set $\{y_1, y_2, \ldots\}$ and a nondecreasing sequence of integers $\{k(m)\}_{m=1}^{\infty}$ such that

$$\operatorname{span}\{x_1, x_2, \ldots, x_m\} = \operatorname{span}\{y_1, y_2, \ldots, y_{k(m)}\}$$

for each $m \in \mathcal{N}$.

PROOF We outline an argument by mathematical induction, but leave the details for Exercise 13.31.

Let $y_1 = x_1/\|x_1\|$. Proceeding inductively, suppose y_1, y_2, …, $y_{k(m)}$ have been chosen so that $\{y_1, y_2, \ldots, y_{k(m)}\}$ is an orthonormal set and

$$\operatorname{span}\{x_1, x_2, \ldots, x_m\} = \operatorname{span}\{y_1, y_2, \ldots, y_{k(m)}\}.$$

Define

$$v = x_{m+1} - \sum_{j=1}^{k(m)} \langle x_{m+1}, y_j\rangle y_j.$$

Then we find that v is orthogonal to y_j for $j = 1, 2, \ldots, k(m)$.

If $v = 0$, then $x_{m+1} \in \operatorname{span}\{y_1, y_2, \ldots, y_{k(m)}\}$, and we let $k(m+1) = k(m)$. If $v \neq 0$, we let $k(m+1) = k(m) + 1$, and we define $y_{k(m+1)} = v/\|v\|$; then $\{y_1, y_2, \ldots, y_{k(m)}, y_{k(m+1)}\}$ is an orthonormal set. In either case, we have that

$$\operatorname{span}\{x_1, x_2, \ldots, x_m, x_{m+1}\} = \operatorname{span}\{y_1, y_2, \ldots, y_{k(m+1)}\},$$

as required. ■

The following theorem provides several equivalent conditions for an orthonormal set in a Hilbert space to be a basis. It also makes clear why bases in the sense of Definition 13.8 are appropriate analogues of Hamel bases.

Before stating the theorem, we need to discuss **generalized sums** in normed spaces. Let $\{x_\iota\}_{\iota \in I}$ be an indexed collection of elements of a normed space. Then we say that the sum $\sum_{\iota \in I} x_\iota$ converges if there are only countably many nonzero terms and if for every enumeration of these terms, the resulting series converges to the same element.

□ □ □ THEOREM 13.6

Let $\mathcal{H}$ be a Hilbert space and E an orthonormal subset of $\mathcal{H}$. Then the following properties are equivalent:

a) E is a basis.

b) $\overline{\operatorname{span} E} = \mathcal{H}$.

c) $\langle x, e \rangle = 0$ for each $e \in E$ implies $x = 0$.

d) For each $x \in \mathcal{H}$, we have $x = \sum_{e \in E} \langle x, e \rangle e$.

e) $\|x\|^2 = \sum_{e \in E} |\langle x, e \rangle|^2$ for each $x \in \mathcal{H}$.

PROOF (a) ⇒ (b): If $\overline{\operatorname{span} E} \neq \mathcal{H}$, then by Theorem 13.2 on page 462, we can find a nonzero element $z \in (\overline{\operatorname{span} E})^\perp$. Let $e_0 = z/\|z\|$. We note that $E \cup \{e_0\}$ is orthonormal and properly contains E. Thus, E is not a basis.

(b) ⇒ (c): Suppose that $\langle x, e \rangle = 0$ for each $e \in E$. It follows from the properties of an inner product that $\langle x, y \rangle = 0$ for each $y \in \operatorname{span} E$. Using the continuity of the inner product, we conclude that x is orthogonal to every element of $\overline{\operatorname{span} E}$, which by assumption equals H. Therefore, $\langle x, x \rangle = 0$ and, so, $x = 0$.

(c) ⇒ (d): It follows from Bessel's inequality that $\sum_{e \in E} |\langle x, e \rangle|^2 < \infty$. Using that fact and Exercise 2.37(c) on page 46, we see that $E_0 = \{\, e \in E : \langle x, e \rangle \neq 0 \,\}$ is either countably infinite or finite. We will deal with the former case; the latter one is handled in a similar manner. Let $\{e_n\}_{n=1}^\infty$ be an enumeration of E_0 and define $x_n = \sum_{j=1}^n \langle x, e_j \rangle e_j$. If $n < m$, then Theorem 13.4(b) implies that $\|x_n - x_m\|^2 = \sum_{j=n+1}^m |\langle x, e_j \rangle|^2$. It now follows that $\{x_n\}_{n=1}^\infty$ is Cauchy and, therefore, converges to some $y \in \mathcal{H}$. We claim that $y = x$. For each $e \in E$, we have

$$\langle x - y, e \rangle = \langle x, e \rangle - \sum_{j=1}^\infty \langle x, e_j \rangle \langle e_j, e \rangle. \tag{13.11}$$

If e is not in E_0, then $\langle x, e \rangle = 0$ and $\langle e_j, e \rangle = 0$ for each j. If $e = e_k$ for some k, then the right-hand side of (13.11) reduces to $\langle x, e_k \rangle - \langle x, e_k \rangle$. Thus, $x - y$ is orthogonal to each element of E. It follows from (c) that $y = x$.

(d) ⇒ (e): It follows from (d) and the continuity of the inner product that

$$\|x\|^2 = \langle x, x\rangle = \sum_{e\in E} \langle x, e\rangle\langle e, x\rangle = \sum_{e\in E} |\langle x, e\rangle|^2,$$

as required.

(e) ⇒ (a): If E is not a basis, we can find an element $e_0 \in \mathcal{H}$ such that $\|e_0\| = 1$ and $\langle e_0, e\rangle = 0$ for each $e \in E$. Thus,

$$\|e_0\|^2 = 1 \neq 0 = \sum_{e\in E} |\langle e_0, e\rangle|^2.$$

The proof of the theorem is now complete. ■

EXAMPLE 13.12 *Illustrates Theorem 13.6*

Assume as known that $\mathcal{L}^2([-\pi,\pi])$ is complete, a fact that will be proved in the next section. We will show that the orthonormal set $\{ e_n : n \in \mathcal{Z} \}$, introduced in Example 13.10(c), is a basis for $\mathcal{L}^2([-\pi,\pi])$. By Theorem 13.6, it suffices to show that if $f \in \mathcal{L}^2([-\pi,\pi])$ is such that

$$\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx = 0, \qquad n \in Z, \tag{13.12}$$

then $f = 0$ ae.

From (13.12), it follows immediately that $\int_{-\pi}^{\pi} f(x)p(x)\,dx = 0$ for all trigonometric polynomials p. As the reader is asked to show in Exercise 13.34, there is a sequence $\{p_n\}_{n=1}^{\infty}$ of trigonometric polynomials such that $\lim_{n\to\infty} \|\overline{f} - p_n\|_2 = 0$. Using the continuity of the inner product, we conclude that

$$\int_{-\pi}^{\pi} |f(x)|^2\,dx = \int_{-\pi}^{\pi} f(x)\overline{f(x)}\,dx = \lim_{n\to\infty} \int_{-\pi}^{\pi} f(x)p_n(x)\,dx = 0.$$

Hence, f vanishes ae.

Because $\{ e_n : n \in Z \}$ is a basis for $\mathcal{L}^2([-\pi,\pi])$, Theorem 13.6(d) implies that each function $f \in \mathcal{L}^2([-\pi,\pi])$ has the **Fourier series expansion**

$$f(x) = \sum_{n=-\infty}^{\infty} \hat{f}(n)e^{inx},$$

where the convergence is in $\mathcal{L}^2([-\pi,\pi])$. Furthermore, Theorem 13.6(e) yields

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2\,dx = \sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2,$$

which is called **Parseval's identity.** □

Unless we know that a Hilbert space possesses a basis, Theorem 13.6 is of little consequence. That every Hilbert space does in fact have a basis is part of our next theorem.

□ □ □ THEOREM 13.7

Let $\mathcal{H}$ be a Hilbert space, not $\{0\}$. Then the following properties hold.
a) $\mathcal{H}$ has a basis.
b) If E is a basis for a closed linear subspace K of $\mathcal{H}$, then there exists a basis for $\mathcal{H}$ that contains E as a subset.
c) $\mathcal{H}$ has a countable basis if and only if $\mathcal{H}$ is separable.

PROOF We prove (a) and leave (b) and (c) to the reader as Exercises 13.35 and 13.36. Let $\mathcal{O}$ denote the collection of orthonormal subsets of H, ordered by $\subset$. Suppose that $\mathcal{C}$ is a chain of $\mathcal{O}$. Then $\bigcup_{O\in\mathcal{C}} O \in \mathcal{O}$ is an upper bound for $\mathcal{C}$. Thus, we may apply Zorn's lemma (page 15) to obtain a maximal element of $\mathcal{O}$. ■

The Dual of a Hilbert Space

Let y be an element of the Hilbert space $\mathcal{H}$. The mapping defined by

$$\ell(x) = \langle x, y\rangle, \qquad x \in \mathcal{H}, \tag{13.13}$$

is a linear functional and satisfies $|\ell(x)| \le \|x\|\|y\|$. Thus, ℓ belongs to the dual space $\mathcal{H}^*$. It is an important property of Hilbert spaces that all continuous linear functionals are of the form (13.13).

□ □ □ THEOREM 13.8

Let $\mathcal{H}$ be a Hilbert space. Then $\ell \in \mathcal{H}^*$ if and only if there is a $y \in \mathcal{H}$ such that $\ell(x) = \langle x, y\rangle$ for each $x \in \mathcal{H}$. Furthermore, $\|\ell\|_* = \|y\|$.

PROOF We have already observed that functionals of the form (13.13) belong to $\mathcal{H}^*$. Conversely, suppose that $\ell \in \mathcal{H}^*$. If ℓ is identically 0, then (13.13) holds with $y = 0$. Otherwise, $K = \ell^{-1}(\{0\})$ is a proper closed linear subspace of $\mathcal{H}$ and, consequently, $K^\perp$ contains at least one nonzero element z. For each $x \in \mathcal{H}$, we have $\ell(\ell(z)x - \ell(x)z) = 0$. Thus,

$$0 = \langle \ell(z)x - \ell(x)z, z\rangle = \ell(z)\langle x, z\rangle - \ell(x)\langle z, z\rangle.$$

It follows that $\ell(x) = \langle x, y\rangle$, where $y = (\overline{\ell(z)}/\langle z, z\rangle)z$.

To find the norm of the linear functional ℓ, we first apply Cauchy's inequality to get

$$\|\ell\|_* = \sup\{\, |\langle x, y\rangle| : \|x\| \le 1 \,\} \le \|y\|.$$

Thus, if $y = 0$, then, trivially, $\|\ell\|_* = \|y\|$. If $y \ne 0$, we choose $w = y/\|y\|$ in order to obtain $\|y\| = \langle w, y\rangle \le \|\ell\|_*$. ■

Remark: If E is a basis for a Hilbert space $\mathcal{H}$, then we can write a formula for the element y given in Theorem 13.8 in terms of the basis elements. Indeed, noting that $\ell(e) = \langle e, y\rangle$, we have by Theorem 13.6 that

$$y = \sum_{e\in E} \langle y, e\rangle e = \sum_{e\in E} \overline{\ell(e)} e.$$

Theorem 13.8 is a prototype for results appearing in subsequent sections where we find explicit formulas for bounded linear functionals on various Banach spaces.

Exercises for Section 13.3

13.29 Verify the assertions of parts (b) and (c) of Example 13.10 on page 468.

13.30 Prove (a), (b), and (d) of Theorem 13.4 on page 468.

13.31 Provide the details for the proof of Theorem 13.5 on page 470.

13.32 In this exercise, E denotes an orthonormal set and $\mathcal{H}$ a Hilbert space.
a) Show that if e and e' are distinct members of E, then $\|e - e'\|^2 = 2$.
b) Show that if the closed unit ball $\overline{B}_1(0)$ of $\mathcal{H}$ is compact, then $\mathcal{H}$ is finite dimensional.

13.33 Let $[a, b]$ be a closed bounded interval.
a) Prove that the continuous functions are dense in $\mathcal{L}^2([a, b])$.
b) Formulate and prove a similar result for unbounded intervals.

13.34 Prove that the trigonometric polynomials are dense in $\mathcal{L}^2([-\pi, \pi])$. *Hint:* Refer to Exercise 13.33.

13.35 Prove part (b) of Theorem 13.7 on page 473.

13.36 Prove part (c) of Theorem 13.7 on page 473.

13.37 Let E be an orthonormal set of a Hilbert space $\mathcal{H}$. Establish the following facts.
a) $P_{\overline{\text{span } E}}(x) = \sum_{e \in E} \langle x, e\rangle e$ for all $x \in \mathcal{H}$.
b) $\rho(x, \overline{\text{span } E})^2 = \|x\|^2 - \sum_{e \in E} |\langle x, e\rangle|^2$ for all $x \in \mathcal{H}$.
c) If α is a scalar-valued function on E such that $\sum_{e \in E} |\alpha(e)|^2 < \infty$, then the sum $\sum_{e \in E} \alpha(e)e$ converges.

13.38 Refer to Theorem 13.5 on page 470.
a) Apply the technique used in the proof of that theorem to the subset of $\mathcal{L}^2([-1, 1])$ that consists of $1, x, x^2, \ldots$ to obtain an orthonormal set of polynomials $L_0, L_1, \ldots$. Show that

$$L_n(x) = (n + 1/2)^{1/2}(2^n n!)^{-1} d^n(x^2 - 1)^n/dx^n.$$

The polynomials $(2^n n!)^{-1} d^n(x^2 - 1)^n/dx^n$ are called **Legendre polynomials.**
b) Show that $\{L_0, L_1, \ldots\}$ is a basis for $\mathcal{L}^2([-1, 1])$.

13.39 The **Haar functions** are functions on $[0, 1]$ defined as follows. $H_0(t) = 1$,

$$H_1(t) = \begin{cases} 1, & t \in [0, 1/2]; \\ -1, & t \in (1/2, 1], \end{cases}$$

and

$$H_j(t) = \begin{cases} 2^{n/2} H_1(2^n t - j + 2^n), & t \in [-1 + j/2^n, -1 + (j + 1)/2^n]; \\ 0, & \text{otherwise}, \end{cases}$$

for $2^n \le j < 2^{n+1}$. Show that the Haar functions form a basis for $\mathcal{L}^2([0, 1])$.

13.40 Let $n \in \mathcal{N}$. Define a linear functional S on $\mathcal{L}^2([-\pi, \pi])$ by

$$S(f) = \sum_{k=-n}^{n} \hat{f}(k).$$

Find a function $g \in \mathcal{L}^2([-\pi, \pi])$ such that $S(f) = \int_{-\pi}^{\pi} f(x)g(x)\, dx$.

In Exercises 13.41–13.44, we will need the concepts of an isometric function and an isomorphism of normed spaces. Let Ω and Λ be normed spaces and $L\colon \Omega \to \Lambda$. Then L is said to be **isometric** *(or to be an* **isometry***) if $\|L(x)\| = \|x\|$ for each $x \in \Omega$. It is said to be an* **isomorphism** *if it is linear, one-to-one, onto, and continuous and L^{-1} is also continuous.*

13.41 Let $\mathcal{H}$ be a separable Hilbert space. Show that there is an isometric isomorphism from $\mathcal{H}$ onto $\ell^2(\mathcal{N})$.

13.42 Let $\mathcal{H}$ be a Hilbert space. Prove there is an isometric isomorphism from $\mathcal{H}$ onto $\ell^2(S)$ for some set S.

13.43 Prove that the function $g \to \langle \cdot, \overline{g} \rangle$ defines an isometric linear mapping from $\mathcal{L}^2(\mu)$ onto $\mathcal{L}^2(\mu)^*$.

13.44 Show that there is no isometric isomorphism from $\mathcal{L}^2(\mathcal{R})$ onto $\mathcal{L}^1(\mathcal{R})$.

13.4 $\mathcal{L}^p$-SPACES

In Example 10.6 on pages 366–367, we introduced three normed spaces of measurable functions: $\mathcal{L}^1(\mu)$, $\mathcal{L}^2(\mu)$, and $\mathcal{L}^\infty(\mu)$. Now we will generalize to $\mathcal{L}^p(\mu)$, where p is any positive extended real number. These spaces are called *$\mathcal{L}^p$-spaces.*

We will show that for $p \geq 1$, $\mathcal{L}^p(\mu)$ is a Banach space and will describe its dual space in the spirit of Theorem 13.8 (page 473). The $\mathcal{L}^p$-spaces, along with spaces of the form $C(\Omega)$ where Ω is a compact Hausdorff space, are sometimes referred to in the literature as the *classical Banach spaces.*

DEFINITION 13.9 $\mathcal{L}^p$-Spaces

Let $(\Omega, \mathcal{A}, \mu)$ be a measure space, f a complex-valued $\mathcal{A}$-measurable function on Ω, and $0 < p \leq \infty$.

- For $0 < p < \infty$, we define

$$\sigma_p(f) = \int_\Omega |f|^p \, d\mu$$

  and

$$\|f\|_p = \left(\int_\Omega |f|^p \, d\mu \right)^{1/p}.$$

- For $p = \infty$, we define

$$\|f\|_\infty = \inf\{\, M : |f| \leq M \ \mu\text{-ae} \,\}.$$

The collection of complex-valued $\mathcal{A}$-measurable functions f such that $\|f\|_p < \infty$ is denoted $\boldsymbol{\mathcal{L}^p(\Omega, \mathcal{A}, \mu)}$ or, when no confusion can arise, simply $\boldsymbol{\mathcal{L}^p(\mu)}$. The spaces $\mathcal{L}^p(\Omega, \mathcal{A}, \mu)$, $0 < p \leq \infty$, are called **$\mathcal{L}^p$-spaces.**

Note: Under certain conditions, special notation is used for $\mathcal{L}^p$-spaces:

- When μ is Lebesgue measure restricted to some Lebesgue measurable subset Ω of $\mathcal{R}^n$, we write $\mathcal{L}^p(\Omega)$ for $\mathcal{L}^p(\mu)$.
- When μ is counting measure on some set Ω, we write $\ell^p(\Omega)$ for $\mathcal{L}^p(\mu)$ and, in the special case, $\Omega = \mathcal{N}$, we sometimes write simply ℓ^p.

As mentioned earlier, we identify functions that are equal μ-ae. Keeping that in mind, we will see later that $\| \, \|_p$ is a norm on the linear space $\mathcal{L}^p(\mu)$ when $1 \le p \le \infty$. When $0 < p < 1$, the space $\mathcal{L}^p(\mu)$ is still a linear space, but $\| \, \|_p$ is no longer a norm. Rather, in this case, $\mathcal{L}^p(\mu)$ is a metric space with metric given by $\rho_p(f,g) = \sigma_p(f-g)$. See Exercises 13.53–13.55.

EXAMPLE 13.13 Illustrates Definition 13.9

a) Let $[a,b]$ be a closed bounded interval of $\mathcal{R}$ and $0 < p < \infty$. A complex-valued Lebesgue measurable function f on $[a,b]$ is in $\mathcal{L}^p([a,b])$ if and only if $\int_a^b |f(x)|^p \, dx < \infty$.
b) Let μ be counting measure on $\{1,2\}$. Then the space of real-valued functions in $\ell^p(\{1,2\})$ can be identified with $\mathcal{R}^2$. We have

$$\|(x_1,x_2)\|_p = \begin{cases} (|x_1|^p + |x_2|^p)^{1/p}, & 0 < p < \infty; \\ \max\{|x_1|,|x_2|\}, & p = \infty. \end{cases}$$

Figure 13.3 shows the unit "circles" centered at $(0,0)$ in the metric space $(\mathcal{R}^2, \rho_{0.5})$ and in the normed space $(\mathcal{R}^2, \| \, \|_p)$ for $p = 1$, 2, 3, and ∞.

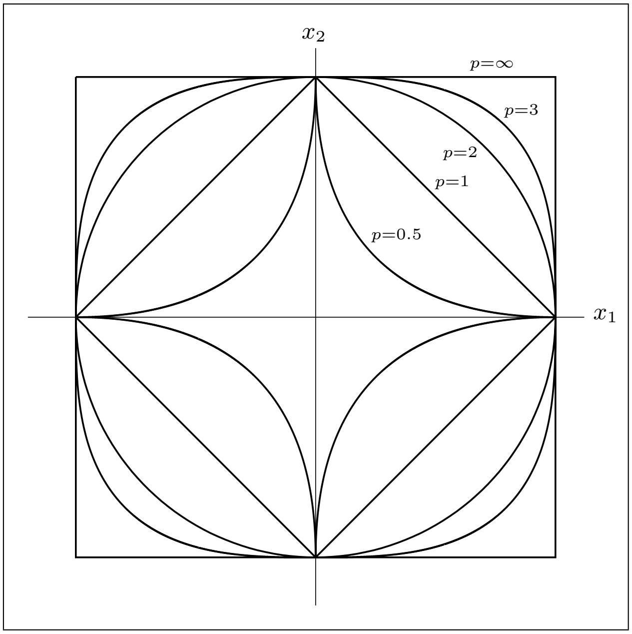

FIGURE 13.3 Selected unit circles

c) Refer to Example 7.10(c) on page 251. Let $(\Omega, \mathcal{A}, P)$ be a probability space. The random variables with finite nth moments are precisely those in $\mathcal{L}^n(P)$.

d) Let μ be counting measure on $\mathcal{N}$ and $0 < p < \infty$. A sequence $\{a_n\}_{n=1}^{\infty}$ of complex numbers is in ℓ^p if and only if $\sum_{n=1}^{\infty} |a_n|^p < \infty$. □

Our next proposition, whose proof is left to the reader as Exercise 13.45, provides some basic properties of $\mathcal{L}^p$-spaces.

□ □ □ PROPOSITION 13.4

Let p be a positive extended real number. Then the following properties hold.

a) $\|\alpha f\|_p = |\alpha| \|f\|_p$ for all $f \in \mathcal{L}^p(\mu)$ and scalars α.

b) $\mathcal{L}^p(\mu)$ is a linear space.

c) For each $f \in \mathcal{L}^p(\mu)$, there exists a sequence of simple functions $\{s_n\}_{n=1}^{\infty}$ in $\mathcal{L}^p(\mu)$ such that $s_n \to f$ μ-ae, $\|f - s_n\|_p \to 0$, and $\int_\Omega |s_n|^p \, d\mu \to \int_\Omega |f|^p \, d\mu$, as $n \to \infty$.

In Section 13.2, we used Cauchy's inequality to prove that an inner product $\langle\ ,\ \rangle$ induces a norm via $\|x\| = \sqrt{\langle x, x\rangle}$. Similarly, we will use Hölder's inequality, a generalization of Cauchy's inequality, to show that $\|\ \|_p$ is a norm when $p \geq 1$.

□ □ □ THEOREM 13.9 Hölder's Inequality

Let $1 \leq p \leq \infty$ and q be such that $1/p + 1/q = 1$. Then for any two $\mathcal{A}$-measurable functions f and g, we have

$$\int_\Omega |fg| \, d\mu \leq \|f\|_p \|g\|_q. \tag{13.14}$$

Furthermore, if $1 < p < \infty$, then equality holds in (13.14) if and only if there are constants α and β not both zero such that $\alpha|f|^p = \beta|g|^q$.

PROOF Without loss of generality we can assume that $\|f\|_p$ and $\|g\|_q$ are finite and nonzero. Suppose that $1 < p < \infty$. By the concavity of the natural log function we have

$$\ln|fg| = (1/p)\ln|f|^p + (1/q)\ln|g|^q \leq \ln((1/p)|f|^p + (1/q)|g|^q).$$

Thus,

$$|fg| \leq (1/p)|f|^p + (1/q)|g|^q. \tag{13.15}$$

If $\|f\|_p = \|g\|_q = 1$, it follows from (13.15) that

$$\int_\Omega |fg| \, d\mu \leq (1/p) \int_\Omega |f|^p \, d\mu + (1/q) \int_\Omega |g|^q \, d\mu = 1/p + 1/q = 1 \tag{13.16}$$

and, hence, (13.14) holds in that case. In general, we can replace f and g by $f/\|f\|_p$ and $g/\|g\|_q$, respectively, and use Proposition 13.4(a) and (13.16) to obtain $(\|f\|_p\|g\|_q)^{-1} \int_\Omega |fg| \, d\mu \leq 1$.

We leave the cases $p = 1$ and $p = \infty$ and the "Furthermore, ..." part to the reader as Exercises 13.46–13.47. ■

□ □ □ THEOREM 13.10 Minkowski's Inequality

Let $1 \le p \le \infty$. *Then*

$$\|f+g\|_p \le \|f\|_p + \|g\|_p$$

for all $f, g \in \mathcal{L}^p(\mu)$.

PROOF The case $p = 1$ follows immediately from the inequality $|f+g| \le |f| + |g|$, and the case $p = \infty$ follows from the fact that if $|f| \le M_1$ μ-ae and $|g| \le M_2$ μ-ae, then $|f+g| \le M_1 + M_2$ μ-ae.

Suppose that $p \in (1, \infty)$ and let q be defined via $1/p + 1/q = 1$. From

$$|f+g|^p \le |f||f+g|^{p-1} + |g||f+g|^{p-1}$$

we get

$$\|f+g\|_p^p \le \int_\Omega |f||f+g|^{p-1}\, d\mu + \int_\Omega |g||f+g|^{p-1}\, d\mu. \tag{13.17}$$

Noting that

$$\int_\Omega (|f+g|^{p-1})^q\, d\mu = \int_\Omega |f+g|^{qp-q}\, d\mu = \|f+g\|_p^p,$$

it follows from (13.17) and Hölder's inequality that

$$\|f+g\|_p^p \le \|f\|_p \|f+g\|_p^{p/q} + \|g\|_p \|f+g\|_p^{p/q}.$$

Hence,

$$\|f+g\|_p^{p-p/q} \le \|f\|_p + \|g\|_p.$$

Whereas $p - p/q = 1$, the proof is complete. ■

It follows from Proposition 13.4 and Theorem 13.10 that $\mathcal{L}^p(\mu)$ is a normed space when $p \in [1, \infty]$. The next theorem shows that it is in fact a Banach space.

□ □ □ THEOREM 13.11 Riesz's Theorem

For $1 \le p \le \infty$, *the normed space* $(\mathcal{L}^p(\mu), \|\ \|_p)$ *is a Banach space, that is, a complete metric space in the metric induced by the norm* $\|\ \|_p$.

PROOF We leave the case $p = \infty$ to the reader as Exercise 13.51. By Proposition 13.3 on page 457, it suffices to show that the series $\sum_{n=1}^\infty f_n$ converges with respect to the norm $\|\ \|_p$ whenever $\sum_{n=1}^\infty \|f_n\|_p < \infty$.

Consider the nondecreasing sequence of functions defined by $g_n = \sum_{k=1}^n |f_k|$, and set $g = \lim_{n\to\infty} g_n$. It follows immediately from Minkowski's inequality that $\int_\Omega g_n^p\, d\mu \le (\sum_{k=1}^n \|f_k\|_p)^p$. Applying the monotone convergence theorem, we obtain

$$\int_\Omega g^p\, d\mu \le \left(\sum_{n=1}^\infty \|f_n\|_p\right)^p < \infty.$$

Hence, g must be finite μ-ae.

It is easy to see that, whenever $g(x) < \infty$, the sequence of partial sums $s_n(x) = \sum_{k=1}^{n} f_k(x)$ is Cauchy and, hence, convergent. Let

$$s(x) = \begin{cases} \lim_{n\to\infty} s_n(x), & \text{if } g(x) < \infty; \\ 0, & \text{if } g(x) = \infty. \end{cases}$$

Then $s \in \mathcal{L}^p(\mu)$ because $\int_\Omega |s|^p\,d\mu \le \int_\Omega |g|^p\,d\mu < \infty$. Also, by using the fact that $|s - s_n|^p \le g^p$ and applying the dominated convergence theorem, we get

$$\lim_{n\to\infty} \|s - s_n\|_p^p = \lim_{n\to\infty} \int_\Omega |s - s_n|^p\,d\mu = 0.$$

We have now shown that the series $\sum_{n=1}^{\infty} f_n$ converges with respect to the norm $\|\ \|_p$. ■

The Dual Space of $\mathcal{L}^p(\mu)$

We will now consider the problem of describing the bounded linear functionals on $\mathcal{L}^p(\mu)$. At this point, we restrict ourselves to the case where $1 < p < \infty$. To begin, we observe that for $g \in \mathcal{L}^q(\mu)$, where $1/p + 1/q = 1$, the linear functional defined by

$$\ell(f) = \int_\Omega fg\,d\mu \tag{13.18}$$

is continuous on $\mathcal{L}^p(\mu)$. Indeed, by Hölder's inequality, $|\ell(f)| \le \|f\|_p\|g\|_q$ and, therefore,

$$\|\ell\|_* \le \|g\|_q. \tag{13.19}$$

We claim that equality holds in (13.19). If $g = 0$, there is nothing to prove. So assume $\|g\|_q \ne 0$ and set

$$s(x) = \begin{cases} \overline{g(x)}/|g(x)|, & \text{if } g(x) \ne 0; \\ 0, & \text{if } g(x) = 0. \end{cases}$$

Then the function $f_0 = s|g|^{q-1}/\|g\|_q^{q-1}$ satisfies

$$\int_\Omega |f_0|^p\,d\mu = \int_\Omega |s|^p|g|^{pq-p}/\|g\|_q^{pq-p}\,d\mu = \int_\Omega |g|^q/\|g\|_q^q\,d\mu = 1.$$

Hence, $f_0 \in \mathcal{L}^p(\mu)$ and $\|f_0\|_p = 1$. Furthermore,

$$\ell(f_0) = \frac{1}{\|g\|_q^{q-1}} \int_\Omega s|g|^{q-1}g\,d\mu = \frac{1}{\|g\|_q^{q-1}} \int_\Omega |g|^q\,d\mu = \|g\|_q.$$

It follows from this last equality and (13.19) that $\|\ell\|_* = \|g\|_q$.

We have shown that functions in $\mathcal{L}^q(\mu)$ induce bounded linear functionals on $\mathcal{L}^p(\mu)$ via the formula (13.18). Now the question is whether these exhaust all bounded linear functionals on $\mathcal{L}^p(\mu)$. The following theorem shows that the answer is yes!

□ □ □ **THEOREM 13.12 Riesz Representation Theorem**

Let $1 < p < \infty$ and $1/p + 1/q = 1$. Then $\ell \in \mathcal{L}^p(\mu)^*$ if and only if there exists a unique $g \in \mathcal{L}^q(\mu)$ such that

$$\ell(f) = \int_\Omega fg\,d\mu, \qquad f \in \mathcal{L}^p(\mu).$$

Furthermore, g satisfies $\|\ell\|_* = \|g\|_q$.

PROOF In view of our discussion directly before this theorem, we need only prove necessity. So, assume that $\ell \in \mathcal{L}^p(\mu)^*$. We will work under the assumption that $(\Omega, \mathcal{A}, \mu)$ is a finite measure space and leave the general case to the reader as Exercises 13.62–13.65. We also leave the proof of the uniqueness of g for Exercise 13.59.

Define the complex measure ν on $\mathcal{A}$ by $\nu(E) = \ell(\chi_E)$. If $\mu(E) = 0$, then we have $\chi_E = 0$ μ-ae and, so, $\nu(E) = \ell(\chi_E) = 0$. Thus, ν is absolutely continuous with respect to μ. Applying the complex version of the Radon-Nikodym theorem (page 326), we conclude that there exists a function $g \in \mathcal{L}^1(\mu)$ such that

$$\ell(\chi_E) = \int_E g\,d\mu, \qquad E \in \mathcal{A}.$$

By linearity, it follows that $\ell(\phi) = \int_\Omega \phi g\,d\mu$ for all ($\mathcal{A}$-measurable) simple functions ϕ. Thus, $|\int_\Omega \phi g\,d\mu| \le \|\ell\|_*\|\phi\|_p$ for all simple functions. Let

$$s(x) = \begin{cases} \overline{g(x)}/|g(x)|, & \text{if } g(x) \ne 0; \\ 0, & \text{if } g(x) = 0. \end{cases}$$

As the reader is asked to show in Exercise 13.60, we can find a sequence of simple functions $\{\psi_n\}_{n=1}^\infty$ such that $|\psi_n| \le 1$ μ-ae and $\psi_n \to s$ μ-ae. We have

$$\left|\int_\Omega \psi_n \phi g\,d\mu\right| \le \|\ell\|_*\|\psi_n\phi\|_p \le \|\ell\|_*\|\phi\|_p$$

and, applying the dominated convergence theorem, we obtain

$$\left|\int_\Omega \phi|g|\,d\mu\right| \le \|\ell\|_*\|\phi\|_p. \tag{13.20}$$

We will use (13.20) to show that g belongs to the space $\mathcal{L}^q(\mu)$.

Let $n \in \mathcal{N}$ and $E_n = \{ x : |g(x)| \le n \}$. The function $f_0 = \chi_{E_n} |g|^{q-1}$ belongs to $\mathcal{L}^p(\mu)$. Hence, by Proposition 13.4 on page 477, there is a sequence $\{\phi_k\}_{k=1}^{\infty}$ of simple functions such that, as $k \to \infty$, $\phi_k \to f_0$ μ-ae and $\|\phi_k\|_p \to \|f_0\|_p$. Replacing ϕ_k by $\chi_{E_n}|\phi_k|$ if necessary, we may assume without loss of generality that the ϕ_ks are nonnegative and vanish outside of E_n. Using Fatou's lemma and (13.20), we obtain

$$\int_{E_n} |g|^{q-1}|g| \, d\mu \le \liminf_{k\to\infty} \int_\Omega \phi_k |g| \, d\mu \le \|\ell\|_* \liminf_{k\to\infty} \|\phi_k\|_p = \|\ell\|_* \|f_0\|_p$$
$$= \|\ell\|_* \left(\int_{E_n} |g|^{p(q-1)} \, d\mu \right)^{1/p} = \|\ell\|_* \left(\int_{E_n} |g|^q \, d\mu \right)^{1/p}$$

and, hence, that

$$\left(\int_{E_n} |g|^q \, d\mu \right)^{1/q} = \left(\int_{E_n} |g|^q \, d\mu \right)^{1-1/p} \le \|\ell\|_*.$$

Letting $n \to \infty$ and applying the MCT, we get that $\|g\|_q \le \|\ell\|_*$. Thus, g belongs to $\mathcal{L}^q(\mu)$.

Because $g \in \mathcal{L}^q(\mu)$, the function ℓ_g defined by $\ell_g(f) = \int_\Omega fg \, d\mu$ is in $\mathcal{L}^p(\mu)^*$. As ℓ and ℓ_g agree on simple functions, Proposition 13.4 implies that they are identical. ■

Remark: If $p = 1$, Theorem 13.12 remains valid under the additional assumption that $(\Omega, \mathcal{A}, \mu)$ is σ-finite, as the reader is asked to prove in Exercise 13.61. An example given in Chapter 14 shows that Theorem 13.12 fails when $p = \infty$.

In view of Theorem 13.12, we can write $\mathcal{L}^p(\mu)^* = \mathcal{L}^q(\mu)$, for $1 < p < \infty$, and, in the σ-finite case, for $p = 1$. However, for $p = \infty$, we can assert only that

$$\mathcal{L}^\infty(\mu)^* \supset \mathcal{L}^1(\mu). \tag{13.21}$$

See Exercise 13.58.

EXAMPLE 13.14 *Illustrates Theorem 13.12*

Refer to Example 13.11 on page 470. Let $x \in [-\pi, \pi]$ and $1 < p < \infty$. Define the linear functional ℓ_x on $\mathcal{L}^p([-\pi, \pi])$ by

$$\ell_x(f) = \sum_{k=-n}^{n} \hat{f}(k) e^{ikx}.$$

Of course, ℓ_x just gives the value at x of the nth partial sum of the Fourier series of f.

First we show that ℓ_x is bounded and then we find the function $g \in \mathcal{L}^q([-\pi, \pi])$ guaranteed by Theorem 13.12. From Hölder's inequality,

$$|\hat{f}(k)| = \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) e^{-iky}\, dy \right| \le \frac{\|f\|_p}{2\pi} \left(\int_{-\pi}^{\pi} |e^{-iky}|^q\, dy \right)^{1/q} = \|f\|_p (2\pi)^{-1/p}.$$

It follows at once that $|\ell_x(f)| \le (2n+1)(2\pi)^{-1/p} \|f\|_p$. Thus, ℓ_x is bounded. Finally, we write

$$\ell_x(f) = \sum_{k=-n}^{n} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) e^{ik(x-y)}\, dy = \int_{-\pi}^{\pi} f(y) D_n(x-y)\, dy,$$

where

$$D_n(t) = \frac{1}{2\pi} \sum_{k=-n}^{n} e^{ikt} = \begin{cases} \dfrac{\sin((n+1/2)t)}{2\pi \sin(t/2)}, & t \ne 0; \\ \dfrac{2n+1}{2\pi}, & t = 0. \end{cases}$$

Thus, the function g guaranteed by Theorem 13.12 is $g(y) = D_n(x-y)$. □

Exercises for Section 13.4

13.45 Prove Proposition 13.4 on page 477.

13.46 Prove the "Furthermore, ..." part of Hölder's inequality (Theorem 13.9 on page 477).

13.47 Verify Hölder's inequality (Theorem 13.9 on page 477) for $p = 1$ and $p = \infty$.

13.48 Discuss the case of equality in (13.14) on page 477 when $p = 1$ or $p = \infty$.

13.49 Suppose that $p, q \in (0, \infty]$.
a) Let r be such that $1/r = 1/p + 1/q$. Show that if $f \in \mathcal{L}^p(\mu)$ and $g \in \mathcal{L}^q(\mu)$, then $fg \in \mathcal{L}^r(\mu)$ and $\|fg\|_r \le \|f\|_p \|g\|_q$.
b) Suppose that $(\Omega, \mathcal{A}, \mu)$ is a finite measure space. Show that if $0 < s < r \le \infty$, then $\mathcal{L}^r(\mu) \subset \mathcal{L}^s(\mu)$.

13.50 Let $(\Omega, \mathcal{A}, \mu)$ be a finite measure space. Show that for each $f \in \mathcal{L}^\infty(\mu)$, $\|f\|_p \to \|f\|_\infty$ as $p \to \infty$.

13.51 Prove that the normed space $(\mathcal{L}^\infty(\mu), \|\ \|_\infty)$ is a Banach space.

13.52 Show that $(\mathcal{L}^p([0,1]), \|\ \|_p)$ is not an inner product space unless $p = 2$.

13.53 Show that $\|\ \|_p$ does not define a norm on $\mathcal{L}^p([0,1])$ when $0 < p < 1$.

★*13.54* Refer to Definition 13.9 on page 475.
a) Show that if $0 < p < 1$, then $\sigma_p(f+g) \le \sigma_p(f) + \sigma_p(g)$.
b) Deduce that $\rho_p(f,g) = \sigma_p(f-g)$ defines a metric on $\mathcal{L}^p(\mu)$ for $0 < p < 1$.

13.55 Refer to Exercise 13.54. Show that if $0 < p < 1$, then $(\mathcal{L}^p(\mu), \rho_p)$ is a complete metric space.

★*13.56* Let J be a nonempty interval in $\mathcal{R}$ and $0 < p < \infty$.
a) Show that if J is closed and bounded, then $C(J)$ is dense in $\mathcal{L}^p(J)$.
b) Refer to Example 11.9 on page 421. Show that $C_c(J)$ is dense in $\mathcal{L}^p(J)$.
c) Show that $C_c(J)$ is not dense in $\mathcal{L}^\infty(J)$.

13.57 Let $0 < p < \infty$. Prove that the trigonometric polynomials are dense in $\mathcal{L}^p([-\pi, \pi])$.

13.58 The result of this exercise gives meaning to the relation (13.21) on page 481. Prove that if $g \in \mathcal{L}^1(\mu)$, then $\ell(f) = \int_\Omega fg\,d\mu$ defines a bounded linear functional on $\mathcal{L}^\infty(\mu)$ and that $\|\ell\|_* = \|g\|_1$.

13.59 Prove the uniqueness of the function g in Theorem 13.12 (page 480).

★*13.60* Suppose $f \in \mathcal{L}^\infty(\mu)$. Show that there exists a sequence of simple functions $\{\phi_n\}_{n=1}^\infty$ such that $|\phi_n| \le \|f\|_\infty$ μ-ae and $\lim_{n\to\infty} \phi_n = f$ μ-ae.

★*13.61* Prove Theorem 13.12 (page 480) when $p = 1$ under the assumption that $(\Omega, \mathcal{A}, \mu)$ is a σ-finite measure space.

In Exercises 13.62–13.65, we complete the proof of Theorem 13.12 (page 480) by eliminating the restriction $\mu(\Omega) < \infty$.

13.62 Suppose that $(\Omega, \mathcal{A}, \mu)$ is a measure space. For $E \in \mathcal{A}$, define the measure μ_E on $\mathcal{A}$ by $\mu_E(A) = \mu(E \cap A)$.

a) Show that $f \in \mathcal{L}^p(\mu_E)$ if and only if $\chi_E f \in \mathcal{L}^p(\mu)$.
b) Show that if $\ell \in \mathcal{L}^p(\mu)^*$, then $\ell_E(f) = \ell(\chi_E f)$ defines a continuous linear functional on $\mathcal{L}^p(\mu_E)$ and $\|\ell_E\|_* \le \|\ell\|_*$.
c) If $\mu(E) < \infty$, show there is a unique function $g_E \in \mathcal{L}^q(\mu)$ such that g_E vanishes outside of E, $\ell_E(f) = \int_\Omega f g_E\,d\mu$ for each $f \in \mathcal{L}^p(\mu_E)$, and $\|\ell_E\|_*^q = \int_\Omega |g_E|^q\,d\mu_E$.

13.63 Use Exercise 13.62 to prove Theorem 13.12 in case $(\Omega, \mathcal{A}, \mu)$ is σ-finite.

13.64 Let $(\Omega, \mathcal{A}, \mu)$ be an arbitrary measure space and $1 < p < \infty$. Show that if $\ell \in \mathcal{L}^p(\mu)^*$, then there exists a sequence $\{\Omega_n\}_{n=1}^\infty$ of $\mathcal{A}$-measurable sets such that $\mu(\Omega_n) < \infty$ for each $n \in \mathcal{N}$ and $\ell(\chi_A) = 0$ for each $A \in \mathcal{A}$ such that $\mu(A) < \infty$ and $A \subset (\bigcup_{n=1}^\infty \Omega_n)^c$.

13.65 Employ Exercises 13.62–13.64 to establish Theorem 13.12 for an arbitrary measure space $(\Omega, \mathcal{A}, \mu)$.

13.5 NONNEGATIVE LINEAR FUNCTIONALS ON $C(\Omega)$

We have now characterized the dual spaces of Hilbert spaces (Theorem 13.8 on page 473) and $\mathcal{L}^p$-spaces (Theorem 13.12 on page 480). Our next task, which we will begin in this section and complete in the following one, is to characterize the dual spaces of $C(\Omega)$ and $C_0(\Omega)$.

We will see that the linear functional defined on $C([0,1])$ by

$$\ell_\lambda(f) = \int_0^1 f(x)\,dx = \int_{[0,1]} f\,d\lambda$$

is typical in the sense that all bounded linear functionals on $C(\Omega)$ arise from integration with respect to some complex measure. Here we lay the foundation for the general result by characterizing those that arise from integration with respect to a (nonnegative) measure.

Borel Sets and Regular Borel Measures

In Chapter 3 we defined the collection $\mathcal{B}$ of Borel sets of $\mathcal{R}$. We showed in Theorem 3.4 that $\mathcal{B}$ is the smallest σ-algebra of subsets of $\mathcal{R}$ that contains the open sets of $\mathcal{R}$. This characterization allows us to extend the concept of Borel sets to any topological space.

DEFINITION 13.10 **Borel Set, Measure, and Measurable Function**

Let Ω be a topological space. The smallest σ-algebra of subsets of Ω that contains all the open sets is denoted $\boldsymbol{\mathcal{B}(\Omega)}$. We use the following terminology:

- **Borel set:** a member of $\mathcal{B}(\Omega)$.
- **Borel measurable function:** a function measurable with respect to $\mathcal{B}(\Omega)$.
- **Borel measure:** a signed or complex measure on $\mathcal{B}(\Omega)$.

EXAMPLE 13.15 Illustrates Definition 13.10

a) $\mathcal{B}(\mathcal{R}) = \mathcal{B}$, as defined in Chapter 3.
b) $\mathcal{B}(\mathcal{R}^2) = \mathcal{B}_2 = \mathcal{B} \times \mathcal{B}$, as discussed in Exercise 6.53 on page 211. More generally, we have that $\mathcal{B}(\mathcal{R}^n) = \mathcal{B}_n = \mathcal{B} \times \mathcal{B} \times \cdots \times \mathcal{B}$, as discussed in Exercise 6.77 on page 223.
c) Let Ω be any set and $\mathcal{T} = \{\Omega, \emptyset\}$. Then $\mathcal{B}(\Omega) = \mathcal{T}$.
d) Let Ω be any set and $\mathcal{T}$ be the discrete topology on Ω. Then we have that $\mathcal{B}(\Omega) = \mathcal{T} = \mathcal{P}(\Omega)$.
e) Let $(\Omega, \mathcal{T})$ be a topological space. Then all functions in $C(\Omega)$ are Borel measurable. □

To characterize the bounded linear functionals on $C(\Omega)$, we need the concept of a *regular Borel measure.* We recommend that the reader review the discussion of the total variation of a complex measure presented in Section 9.3 starting on page 321.

DEFINITION 13.11 **Regular Borel Measure**

Let Ω be a locally compact Hausdorff space. A complex Borel measure μ is said to be a **regular Borel measure** if for each $B \in \mathcal{B}(\Omega)$ and $\epsilon > 0$, there is a compact set K and an open set O such that $K \subset B \subset O$ and $|\mu|(O \setminus K) < \epsilon$.

The collection of all regular Borel measures on Ω is denoted by $\boldsymbol{M(\Omega)}$; the real-valued and nonnegative regular Borel measures are denoted, respectively, by $\boldsymbol{M_r(\Omega)}$ and $\boldsymbol{M_+(\Omega)}$.

Remark: Definition 13.11 requires that a regular Borel measure be finite valued. Other definitions of regular Borel measure exist and some permit certain extended real-valued measures, such as Lebesgue measure, to be regular.

EXAMPLE 13.16 Illustrates Definition 13.11

a) Lebesgue measure on $[0, 1]$ is a regular Borel measure. In fact, Lebesgue measure on any Borel set of finite Lebesgue measure is a regular Borel measure.
b) The Lebesgue-Stieltjes measure corresponding to a distribution function on $\mathcal{R}$ is a regular Borel measure, as the reader is asked to establish in Exercise 13.68.

c) Let Ω be a locally compact Hausdorff space. For $x \in \Omega$, the Dirac measure concentrated at x, restricted to the Borel sets of Ω, is a regular Borel measure. See Exercise 13.71. □

Suppose that Ω is a locally compact Hausdorff space. The spaces $M(\Omega)$ and $M_r(\Omega)$ are, respectively, complex and real linear spaces, where the operations of addition and scalar multiplication are defined by

$$(\mu + \nu)(B) = \mu(B) + \nu(B) \qquad \text{and} \qquad (\alpha\mu)(B) = \alpha\mu(B).$$

Referring to Exercise 9.48 on page 329, we see that the linear spaces $M(\Omega)$ and $M_r(\Omega)$ are also normed spaces, where the norm is given by the total variation, that is, $\|\mu\| = |\mu|(\Omega)$. Moreover, as the reader is asked to prove in Exercise 13.66, $M(\Omega)$ and $M_r(\Omega)$ are Banach spaces with respect to the norm $\| \ \|$.

If F is a closed subset of Ω, then any $\nu \in M(F)$ can be extended to a regular Borel measure ν' on Ω by defining

$$\nu'(B) = \nu(B \cap F), \qquad B \in \mathcal{B}(\Omega).$$

It is convenient to view ν as a measure on Ω by identifying it with ν'. In this way we can identify $M(F)$ with the linear subspace

$$\{\mu \in M(\Omega) : \mu(B) = 0 \text{ for all } B \in \mathcal{B}(\Omega) \text{ with } B \subset F^c\}.$$

Nonnegative Linear Functionals

From here on in this section, unless explicitly stated otherwise, we assume that Ω is a *compact Hausdorff space.* If $\mu \in M(\Omega)$, then μ induces a linear functional on the space $C(\Omega)$ via

$$\ell_\mu(f) = \int_\Omega f \, d\mu, \qquad f \in C(\Omega).$$

That ℓ_μ is a bounded linear functional follows from

$$|\ell_\mu(f)| \le \|f\|_\Omega |\mu|(\Omega) = \|f\|_\Omega \|\mu\|,$$

where we have applied Exercise 9.53(b) on page 330.

In this section, we will show that any linear functional on $C(\Omega)$ that satisfies a certain nonnegativity condition must be of the form ℓ_μ for some $\mu \in M_+(\Omega)$. In the next section, we will extend this result to all bounded linear functionals on $C(\Omega)$ if Ω is a compact Hausdorff space and to $C_0(\Omega)$ if Ω is a locally compact Hausdorff space.

DEFINITION 13.12 Nonnegative Linear Functional

A linear functional ℓ on $C(\Omega)$ is said to be **nonnegative** if $\ell(f) \ge 0$ whenever $f \ge 0$.

As the reader is asked to show in Exercise 13.75, the linear functional ℓ_μ on $C(\Omega)$ induced by a regular Borel measure μ is nonnegative if and only if μ is nonnegative.

The next theorem, whose proof is left to the reader in Exercises 13.76–13.81, presents some basic properties of nonnegative linear functionals.

□ □ □ THEOREM 13.13

Let Ω be a compact Hausdorff space.
a) If ℓ is a nonnegative linear functional on $C(\Omega)$, then $\ell \in C(\Omega)^*$ and, moreover, $\|\ell\|_* = \ell(1)$.
b) If $\ell \in C(\Omega)^*$ and $\ell(C(\Omega, \mathcal{R})) \subset \mathcal{R}$, then there exist nonnegative linear functionals ℓ_+ and ℓ_- such that $\|\ell\|_* = \ell_+(1) + \ell_-(1)$ and $\ell = \ell_+ - \ell_-$.

We have noted that a nonnegative regular Borel measure on Ω induces a nonnegative linear functional on $C(\Omega)$. Our next theorem shows that all nonnegative linear functionals on $C(\Omega)$ are of that type. There are two main ideas in the proof of this result. One is the use of Urysohn's lemma to obtain suitable approximations to characteristic functions of closed sets. The other is to mimic the construction of Lebesgue measure from Lebesgue outer measure.

With regard to the latter, recall that the collection $\mathcal{M}$ of Lebesgue measurable sets is defined by using the Carathéodory criterion and Lebesgue outer measure: $E \in \mathcal{M}$ if and only if

$$\lambda^*(W) = \lambda^*(W \cap E) + \lambda^*(W \cap E^c)$$

for all $W \subset \mathcal{R}$. Theorem 3.11 on page 103 shows that $\mathcal{M}$ is a σ-algebra. A careful look at the proof reveals that it uses only the properties of Lebesgue outer measure given in (a), (b), (c), and (e) of Proposition 3.1 on page 92. In other words, we have already proved the following proposition.

□ □ □ PROPOSITION 13.5

Let Ω be a set and ν^* an extended real-valued function on $\mathcal{P}(\Omega)$ that satisfies the following conditions:
a) $\nu^*(A) \geq 0$ for each $A \subset \Omega$.
b) $\nu^*(\emptyset) = 0$.
c) $A \subset B \Rightarrow \nu^*(A) \leq \nu^*(B)$.
d) $\{A_n\}_n \subset \mathcal{P}(\Omega) \Rightarrow \nu^*\left(\bigcup_n A_n\right) \leq \sum_n \nu^*(A_n)$.
Then the collection of subsets E of Ω that satisfy

$$\nu^*(W) = \nu^*(W \cap E) + \nu^*(W \cap E^c)$$

for all $W \subset \Omega$ is a σ-algebra whose members are called $\boldsymbol{\nu^*}$**-measurable sets.**†

We now state and prove the main result of this section, known as the Riesz-Markov theorem.

---

† Proposition 6.2 on page 183 shows that the outer measure ν^* induced by an appropriate set function on a semialgebra of subsets of a set Ω satisfies (a)–(d) of Proposition 13.5. Thus, the concept of ν^*-measurability given here is the same as that in Definition 6.3 on page 184.

□ □ □ THEOREM 13.14 Riesz-Markov Theorem

Let Ω be a compact Hausdorff space and let ℓ be a nonnegative linear functional on $C(\Omega)$. Then there exists a unique $\mu \in M_+(\Omega)$ such that

$$\ell(f) = \int_\Omega f\,d\mu, \qquad f \in C(\Omega).$$

PROOF We start by assigning a nonnegative number $\overline{\mu}(O)$ to each open set O. If $O = \emptyset$, let $\overline{\mu}(O) = 0$; otherwise, let

$$\overline{\mu}(O) = \sup\{\,\ell(f) : 0 \le f \le 1 \text{ and } \operatorname{supp} f \subset O\,\}.$$

We note that $\overline{\mu}(O) \le \overline{\mu}(\Omega) = \ell(1)$ for all O. Next, for each $A \subset \Omega$, we define

$$\mu^*(A) = \inf\{\,\overline{\mu}(O) : O \text{ open and } O \supset A\,\}.$$

Observe that $\mu^*(O) = \overline{\mu}(O)$ whenever O is open.

We will show that μ^* satisfies the hypotheses of Proposition 13.5. Conditions (a)–(c) follow easily from the definition of μ^*. To verify condition (d), we first show that if $\{O_n\}_{n=1}^\infty$ is a sequence of open subsets of Ω, then

$$\overline{\mu}\left(\bigcup_{n=1}^\infty O_n\right) \le \sum_{n=1}^\infty \overline{\mu}(O_n). \tag{13.22}$$

Let f be a continuous function that satisfies $0 \le f \le 1$ and $\operatorname{supp} f \subset \bigcup_{n=1}^\infty O_n$. Applying Theorem 11.10 on page 412 with $K = \operatorname{supp} f$, we obtain continuous functions $f_1, f_2, \ldots, f_m$ that satisfy

- $0 \le f_j \le 1$, for each j,
- $\sum_{j=1}^m f_j(x) = 1$ for $x \in \operatorname{supp} f$,
- $\sum_{j=1}^m f_j \le 1$, and
- for each j, there is an m_j such that $\operatorname{supp} f_j \subset O_{m_j}$.

By replacing f_j by $\sum_{m_k = m_j} f_k$ if necessary, we can assume that the m_js are distinct. It is clear that $f = \sum_{j=1}^m f f_j$ and, so, $\ell(f) = \sum_{j=1}^m \ell(f f_j)$. Because $\operatorname{supp} f f_j \subset O_{m_j}$, it follows that

$$\ell(f) \le \sum_{n=1}^\infty \overline{\mu}(O_n). \tag{13.23}$$

Taking the supremum on the left-hand side of (13.23), we obtain (13.22). It is now easy to check that μ^* satisfies condition (d) of Proposition 13.5, as we ask the reader to verify in Exercise 13.82.

We complete the proof of the theorem by showing successively that

- all open sets are μ^*-measurable,
- $\mu = \mu^*_{|\mathcal{B}(\Omega)}$ is a regular Borel measure, and
- $\ell(f) = \int_\Omega f\,d\mu$ for all $f \in C(\Omega)$.

To show that an arbitrary open set O is μ^*-measurable, it suffices to prove that

$$\mu^*(A) \ge \mu^*(A \cap O) + \mu^*(A \cap O^c) \tag{13.24}$$

for all $A \subset \Omega$. Let U be an open set containing A, f a continuous function that satisfies $0 \le f \le 1$ and $\operatorname{supp} f \subset U \cap O$, and $V = U \cap (\operatorname{supp} f)^c$. If g is a continuous function that satisfies $0 \le g \le 1$ and $\operatorname{supp} g \subset V$, then

$$\operatorname{supp}(f+g) \subset \operatorname{supp} f \cup \operatorname{supp} g \subset U.$$

It follows that

$$\overline{\mu}(U) \ge \ell(f) + \ell(g). \tag{13.25}$$

From (13.25) we deduce that

$$\overline{\mu}(U) \ge \ell(f) + \overline{\mu}(V) \ge \ell(f) + \mu^*(A \cap O^c)$$

and, therefore, that

$$\overline{\mu}(U) \ge \overline{\mu}(U \cap O) + \mu^*(A \cap O^c) \ge \mu^*(A \cap O) + \mu^*(A \cap O^c).$$

As the open set U containing A was chosen arbitrarily, (13.24) holds.

Having shown that all open sets are μ^*-measurable, we can invoke Proposition 13.5 on page 486 and Proposition 6.5 on page 186 to conclude that all Borel sets are μ^*-measurable and that $\mu = \mu^*_{|\mathcal{B}(\Omega)}$ is a Borel measure. To show that μ is regular, we first observe that, by the definition of μ^*,

$$\mu(B) = \inf\{\, \mu(O) : O \text{ open and } O \supset B \,\}, \qquad B \in \mathcal{B}(\Omega). \tag{13.26}$$

Because $\mu(\Omega) = \ell(1) < \infty$, we have for each $B \in \mathcal{B}(\Omega)$ that

$$\begin{aligned} \mu(B) &= \mu(\Omega) - \mu(B^c) = \mu(\Omega) - \inf\{\, \mu(W) : W \text{ open}, W \supset B^c \,\} \\ &= \sup\{\, \mu(W^c) : W \text{ open}, W \supset B^c \,\} \\ &= \sup\{\, \mu(F) : F \text{ closed}, F \subset B \,\}. \end{aligned} \tag{13.27}$$

It follows at once from (13.26) and (13.27) that μ is regular.

Finally, we must show that

$$\ell(f) = \int_\Omega f\,d\mu, \qquad f \in C(\Omega). \tag{13.28}$$

Every function in $C(\Omega)$ is a linear combination of functions with values in the interval $[0,1)$. Therefore, by the linearity of ℓ, it suffices to establish (13.28) in case $0 \le f < 1$.

Let $n \in \mathcal{N}$. For each integer k, $0 \le k \le n$, the sets $F_k = f^{-1}([k/n, \infty))$ and $U_k = f^{-1}(((k-1)/n, \infty))$ are closed and open, respectively. Moreover, we have

$$F_{k+1} \subset U_{k+1} \subset F_k \subset U_k \qquad \text{and} \qquad \Omega = \bigcup_{k=0}^{n-1} (F_k \setminus F_{k+1}).$$

If $F_k = \emptyset$, we set $g_k = 0$. Otherwise, we first invoke the regularity of μ to choose an open set V_k such that $F_k \subset V_k \subset U_k$ and $\sum_{k=0}^{n-1} \mu(V_k \setminus F_k) < 1$ and then apply Proposition 10.14 on page 387 and Urysohn's lemma on page 388 to obtain a continuous function g_k such that $0 \le g_k \le 1$, $g_k(F_k) = \{1\}$, and $\operatorname{supp} g_k \subset V_k$.

Let $h = (1/n) \sum_{j=0}^{n-1} g_j$. We claim $f \le h$. For each $x \in \Omega$, choose the unique k such that $x \in f^{-1}([k/n, (k+1)/n)) = F_k \setminus F_{k+1}$. If $0 \le j \le k$, then $g_j(x) = 1$ since $F_k \subset F_j$; if $j > k+1$, then $g_j(x) = 0$ since $x \in F_{k+1}^c \subset U_{k+2}^c \subset U_j^c \subset V_j^c$. It follows that

$$h(x) = (k+1)/n + g_{k+1}(x)/n \ge f(x),$$

as required. Using the fact that $f \le h$ and the nonnegativity of ℓ, we obtain

$$\begin{aligned} \ell(f) \le \ell(h) &= (1/n) \sum_{j=0}^{n-1} \ell(g_j) \le (1/n) \sum_{j=0}^{n-1} \mu(V_j) \\ &= (1/n) \sum_{j=0}^{n-1} \bigl(\mu(V_j \setminus F_j) + \mu(F_j)\bigr) \le 1/n + (1/n) \sum_{j=0}^{n-1} \mu(F_j). \end{aligned} \tag{13.29}$$

For $j = 0, 1, \ldots, n-1$, we can write $F_j = \bigcup_{k=j}^{n-1} (F_k \setminus F_{k+1})$ and, therefore,

$$\mu(F_j) = \sum_{k=j}^{n-1} \mu(F_k \setminus F_{k+1}).$$

Applying (13.29), we get

$$\begin{aligned} \ell(f) &\le 1/n + (1/n) \sum_{j=0}^{n-1} \sum_{k=j}^{n-1} \mu(F_k \setminus F_{k+1}) \\ &= 1/n + (1/n) \sum_{k=0}^{n-1} (k+1) \mu(F_k \setminus F_{k+1}) \\ &= 1/n + \mu(\Omega)/n + \sum_{k=0}^{n-1} (k/n) \mu(F_k \setminus F_{k+1}) \\ &= 1/n + \ell(1)/n + \int_\Omega \sum_{k=0}^{n-1} (k/n) \chi_{(F_k \setminus F_{k+1})} \, d\mu \\ &\le (1 + \ell(1))/n + \int_\Omega f \, d\mu. \end{aligned}$$

Because n was chosen arbitrarily, it follows that

$$\ell(f) \le \int_\Omega f \, d\mu. \tag{13.30}$$

We can replace f by $(1-f)/2$ in (13.30) to get

$$\ell(1) - \ell(f) \le \mu(\Omega) - \int_\Omega f\,d\mu = \ell(1) - \int_\Omega f\,d\mu.$$

Thus (13.28) holds.

It remains only to prove the uniqueness of μ, which we leave to the reader as Exercise 13.83. ■

Exercises for Section 13.5

13.66 Let Ω be a locally compact Hausdorff space. Show that $(M(\Omega), \|\ \|)$ and $(M_r(\Omega), \|\ \|)$ are Banach spaces, where $\|\mu\| = |\mu|(\Omega)$.

13.67 Let Ω be a locally compact Hausdorff space. Show that if $\mu \in M(\Omega)$, then $|\mu| \in M(\Omega)$.

★*13.68* In this exercise, you are asked, among other things, to verify the statement of Example 13.16(b).

a) Prove that if a locally compact metric space Ω is the countable union of compact subsets, then every complex Borel measure on Ω is regular.
b) Show that the Lebesgue-Stieltjes measure associated with a distribution function on $\mathcal{R}$ is a regular Borel measure.

13.69 Suppose Ω is locally compact and $\mu \in M_+(\Omega)$. Prove that $C_0(\Omega)$ is dense in $\mathcal{L}^p(\mu)$ for $1 \le p < \infty$.

13.70 Let $\mu \in M([0,1])$ satisfy

$$\int_{[0,1]} x^n\,d\mu(x) = 0$$

for $n = 0, 1, 2, \ldots$. Show that $\mu = 0$, that is, μ vanishes identically.

13.71 Suppose that Ω is a locally compact Hausdorff space. Let $x \in \Omega$ and δ_x be defined on $\mathcal{B}(\Omega)$ by

$$\delta_x(B) = \begin{cases} 1, & \text{if } x \in B; \\ 0, & \text{if } x \notin B. \end{cases}$$

a) Show that δ_x is a regular Borel measure.
b) Determine $\int_\Omega f\,d\delta_x$ when $f \in C(\Omega)$.

13.72 Let δ_x be as in Exercise 13.71.

a) Show that $\|\delta_x - \delta_y\| = 1$ when $x \neq y$.
b) Deduce that $M(\Omega)$ is not separable if Ω is uncountable.

★*13.73* Show how to identify $M(\Omega)$ and $\ell^1(\Omega)$ when Ω is countable.

13.74 Let Ω be a locally compact Hausdorff space, $\mu \in M(\Omega)$, and $B \in \mathcal{B}(\Omega)$. Prove that there are sets F and G such that G is a countable intersection of open sets, F is a countable union of closed sets, $F \subset B \subset G$, and $|\mu|(G \setminus F) = 0$.

13.75 Let Ω be a compact Hausdorff space and $\mu \in M(\Omega)$. Show that $\ell_\mu(f) = \int_\Omega f\,d\mu$ defines a nonnegative linear functional on $C(\Omega)$ if and only if $\mu \in M_+(\Omega)$.

Exercises 13.76–13.81 provide the proof of Theorem 13.13 on page 486.

13.76 Show that if ℓ is a nonnegative linear functional on $C(\Omega)$, then $\ell \in C(\Omega)^*$ and, moreover, $\|\ell\|_* = \ell(1)$.

13.77 Suppose that ℓ satisfies the hypotheses of part (b) of Theorem 13.13 on page 486. For each nonnegative continuous function f on Ω, let

$$\ell_+(f) = \sup\{\, \ell(g) : 0 \le g \le f,\ g \text{ continuous} \,\}.$$

a) Show that if f_1 and f_2 are nonnegative and continuous, then

$$\ell_+(f_1 + f_2) = \ell_+(f_1) + \ell_+(f_2).$$

b) Show that $0 \le f \le g$ implies $\ell_+(f) \le \ell_+(g)$.
c) Show that $\ell_+(\alpha f) = \alpha \ell_+(f)$ whenever $f \ge 0$ and α is a nonnegative real number.

13.78 Extend the function ℓ_+ defined in Exercise 13.77 to all of $C(\Omega, \mathcal{R})$ by the formula

$$\ell_+(f) = \ell_+(\|f\| + f) - \ell_+(\|f\|),$$

where $\|f\| = \|f\|_\Omega$.
a) Prove that this new definition of $\ell_+(f)$ agrees with the old one when f is nonnegative.
b) Show that this extended ℓ_+ is linear on the space $C(\Omega, \mathcal{R})$.

13.79 Extend the function ℓ_+ defined in Exercise 13.78 to all of $C(\Omega)$ by the formula

$$\ell_+(f) = \ell_+(\Re f) + i\ell_+(\Im f).$$

a) Prove that this new definition of $\ell_+(f)$ agrees with the old one when f is real valued.
b) Show that this extended function is linear and nonnegative.

13.80 Suppose that ℓ satisfies the hypotheses of part (b) of Theorem 13.13 on page 486. Let $\ell_- = \ell_+ - \ell$, where ℓ_+ is defined as in Exercise 13.79. Show that ℓ_- is nonnegative.

13.81 Suppose that ℓ satisfies the hypotheses of part (b) of Theorem 13.13 on page 486. Let ℓ_+ and ℓ_- be defined as in Exercise 13.80. Show that $\|\ell\|_* = \ell_+(1) + \ell_-(1)$. *Hint:* If $0 \le g \le 1$, then $\|2g - 1\| \le 1$ and, so, $\|\ell\|_* \ge 2\ell(g) - 1$.

13.82 Show that the set function μ^* defined in the proof of Theorem 13.14 on page 487 satisfies condition (d) of Proposition 13.5 on page 486.

13.83 Prove the uniqueness part of Theorem 13.14 on page 487.

13.6 THE DUAL SPACES OF $C(\Omega)$ AND $C_0(\Omega)$

In this section, we extend the Riesz-Markov theorem (page 487) to arbitrary bounded linear functionals on $C(\Omega)$. We will also characterize the bounded linear functionals on $C_0(\Omega)$ when Ω is a locally compact Hausdorff space. These results show that we are justified in writing $C(\Omega)^* = M(\Omega)$ and $C_0(\Omega)^* = M(\Omega)$ in the compact and locally compact cases, respectively.

□ □ □ LEMMA 13.1

Suppose that Ω is a compact Hausdorff space and $\mu \in M(\Omega)$. Further suppose that ϕ is a complex-valued Borel measurable function such that $|\phi| \le 1$ $|\mu|$-ae. Then there is a sequence $\{f_n\}_{n=1}^\infty$ of continuous functions such that $\|f_n\|_\Omega \le 1$ for each n and $\int_\Omega |f_n - \phi|\, d|\mu| \to 0$ as $n \to \infty$.

PROOF By applying Exercise 13.60 on page 483, we can choose a sequence $\{\phi_n\}_{n=1}^{\infty}$ of Borel measurable simple functions such that $|\phi_n| \le 1$ $|\mu|$-ae for all $n \in \mathcal{N}$ and $\lim_{n\to\infty} \phi_n = \phi$ $|\mu|$-ae. Applying the dominated convergence theorem, we get that

$$\lim_{n\to\infty} \int_\Omega |\phi_n - \phi| \, d|\mu| = 0. \tag{13.31}$$

Let $n \in \mathcal{N}$. We can write $\phi_n = \sum_{k=1}^{m} \alpha_k \chi_{E_k}$, where $|\alpha_k| \le 1$ for each k and the E_ks are pairwise disjoint Borel sets whose union is Ω. Using the regularity of μ, we can find compact sets $F_k \subset E_k$ such that $|\mu|(E_k \setminus F_k) < 1/nm$ for $k = 1, 2, \ldots, m$.

For each k, we can write $\alpha_k = |\alpha_k| e^{i\theta_k}$, where $\theta_k \in [0, 2\pi)$. If $x \in F_k$, define $u_0(x) = |\alpha_k|$ and $v_0(x) = \theta_k$. Since the F_ks are pairwise disjoint and closed, the functions u_0 and v_0 are well-defined and continuous on $\bigcup_{k=1}^{m} F_k$ and, furthermore, $|u_0| \le 1$.

By Tietze's extension theorem (page 389), we can extend u_0 and v_0 to continuous real-valued functions u and v on all of Ω such that $|u| \le 1$. Let $f_n = ue^{iv}$. Then $f_n = \phi_n$ on $\bigcup_{k=1}^{m} F_k$ and $\|f_n\|_\Omega \le 1$. Moreover,

$$\begin{aligned}
\int_\Omega |\phi_n - f_n| \, d|\mu| &= \sum_{k=1}^{m} \int_{E_k} |\alpha_k - f_n| \, d|\mu| \\
&\le \sum_{k=1}^{m} \int_{E_k \setminus F_k} (|\alpha_k| + |f_n|) \, d|\mu| \\
&\le \sum_{k=1}^{m} 2|\mu|(E_k \setminus F_k) \le \sum_{k=1}^{m} 2/mn = 2/n.
\end{aligned} \tag{13.32}$$

It follows from (13.31) and (13.32) that $\lim_{n\to\infty} \int_\Omega |f_n - \phi| \, d|\mu| = 0$. ■

□ □ □ THEOREM 13.15 Riesz Representation Theorem

Let Ω be a compact Hausdorff space. Then $\ell \in C(\Omega)^*$ if and only if there exists a $\mu \in M(\Omega)$ such that

$$\ell(f) = \int_\Omega f \, d\mu, \qquad f \in C(\Omega). \tag{13.33}$$

Furthermore, the measure μ is unique and satisfies

$$\|\ell\|_* = |\mu|(\Omega). \tag{13.34}$$

PROOF In the penultimate paragraph before Definition 13.12 (see page 485), we showed that each $\mu \in M(\Omega)$ induces a bounded linear functional on $C(\Omega)$ via the relation (13.33).

Conversely, suppose that $\ell \in C(\Omega)^*$. Define

$$\ell_{\text{re}}(f) = \frac{1}{2}(\ell(f) + \overline{\ell(\overline{f})}) \qquad \text{and} \qquad \ell_{\text{im}}(f) = \frac{1}{2i}(\ell(f) - \overline{\ell(\overline{f})}).$$

Then ℓ_{re} and ℓ_{im} satisfy $\ell = \ell_{\text{re}} + i\ell_{\text{im}}$ and the hypotheses of Theorem 13.13(b) on page 486. Therefore, by the Riesz-Markov theorem, there are measures $\mu_1, \mu_2, \mu_3, \mu_4 \in M_+(\Omega)$ such that

$$\ell_{\text{re}}(f) = \int_\Omega f\,d\mu_1 - \int_\Omega f\,d\mu_2 \qquad \text{and} \qquad \ell_{\text{im}}(f) = \int_\Omega f\,d\mu_3 - \int_\Omega f\,d\mu_4$$

for all $f \in C(\Omega)$. Thus, the measure $\mu = \mu_1 - \mu_2 + i(\mu_3 - \mu_4)$ belongs to $M(\Omega)$ and satisfies (13.33).

To verify (13.34), we note first that

$$\|\ell\|_* = \sup\left\{ \left| \int_\Omega f\,d\mu \right| : \|f\|_\Omega \le 1 \right\} \le \sup\{ \|f\|_\Omega |\mu|(\Omega) : \|f\|_\Omega \le 1 \} = |\mu|(\Omega).$$

To prove the reverse inequality, we first apply Exercise 9.53 on page 330 to obtain a Borel measurable complex-valued function ϕ such that $|\phi| = 1$ $|\mu|$-ae and $\int_\Omega v\,d\mu = \int_\Omega v\phi\,d|\mu|$ for all $v \in \mathcal{L}^1(|\mu|)$. Now applying Lemma 13.1 to $\overline{\phi}$, we choose a sequence $\{f_n\}_{n=1}^\infty$ of continuous functions such that $\|f_n\|_\Omega \le 1$ and $\int_\Omega |f_n - \overline{\phi}|\,d|\mu| \to 0$. We have

$$\left| \int_\Omega f_n\,d\mu - |\mu|(\Omega) \right| = \left| \int_\Omega \phi(f_n - \overline{\phi})\,d|\mu| \right| \le \int_\Omega |f_n - \overline{\phi}|\,d|\mu|.$$

It follows that $|\mu|(\Omega) \le \|\ell\|_*$ and, hence, (13.34) holds. The proof of uniqueness is left to the reader as Exercise 13.84. ■

The Case Ω Locally Compact

Next we extend Theorem 13.15 to locally compact, noncompact Hausdorff spaces. In this case, we work with $C_0(\Omega)$ rather than $C(\Omega)$ because $\|\ \|_\Omega$ is no longer a norm on $C(\Omega)$.

□ □ □ THEOREM 13.16 Riesz Representation Theorem

Let Ω be a locally compact, noncompact Hausdorff space. Then $\ell \in C_0(\Omega)^*$ if and only if there exists a $\mu \in M(\Omega)$ such that

$$\ell(f) = \int_\Omega f\,d\mu, \qquad f \in C_0(\Omega). \tag{13.35}$$

Furthermore, the measure μ is unique and satisfies $\|\ell\|_* = |\mu|(\Omega)$.

PROOF Let $\ell \in C_0(\Omega)^*$. We prove the existence of the measure μ that satisfies (13.35), but leave the proofs of the remainder of the assertions to the reader as Exercise 13.86.

Let $\Omega^* = \Omega \cup \{\omega\}$ be the one-point compactification of Ω, as described in Theorem 11.12 on page 414. Define the function L on $C(\Omega^*)$ by $L(g) = \ell(g_{|\Omega} - g(\omega))$. Clearly L is linear. That it is also bounded, follows from

$$|L(g)| = |\ell(g_{|\Omega} - g(\omega))| \leq \|\ell\|_* \|g_{|\Omega} - g(\omega)\|_\Omega \leq 2\|\ell\|_* \|g\|_{\Omega^*}.$$

Hence, by Theorem 13.15, there is a measure $\mu^* \in M(\Omega^*)$ such that

$$L(g) = \int_{\Omega^*} g\, d\mu^*, \qquad g \in C(\Omega^*).$$

Letting $\mu = \mu^*_{|\mathcal{B}(\Omega)}$, we obtain

$$L(g) = \int_\Omega g\, d\mu + g(\omega)\mu^*(\{\omega\}), \qquad g \in C(\Omega^*). \tag{13.36}$$

Now let $f \in C_0(\Omega)$. By defining $f^*(x) = f(x)$ for $x \in \Omega$ and $f^*(\omega) = 0$, we can extend f to a function $f^* \in C(\Omega^*)$ with the same norm; indeed, $C_0(\Omega)$ is the collection of restrictions to Ω of functions in $C(\Omega^*)$ that vanish at ω. We have by (13.36) that $\ell(f) = L(f^*) = \int_\Omega f\, d\mu$. The regularity of μ follows from Exercise 13.85. ■

Two simple but instructive illustrations of Theorem 13.16 are provided in Example 13.17. In the next chapter, we will see more elaborate applications of the results of this section.

EXAMPLE 13.17 *Illustrates Theorem 13.16*

a) When it is given the discrete topology, the set of positive integers $\mathcal{N}$ becomes a locally compact space. $C_0(\mathcal{N})$ is simply the collection of all sequences $\{a_n\}_{n=1}^\infty$ of complex numbers such that $\lim_{n\to\infty} a_n = 0$. Applying Exercise 13.73 on page 490, we can identify $M(\mathcal{N})$ with $\ell^1(\mathcal{N})$ and, consequently, we can write $C_0(\mathcal{N})^* = \ell^1(\mathcal{N})$. It follows from Theorem 13.16 that each bounded linear functional ℓ on $C_0(\mathcal{N})$ is of the form $\ell(a) = \sum_{n=1}^\infty a_n b_n$ for some $b \in \ell^1(\mathcal{N})$ and, furthermore, that $\|\ell\|_* = \sum_{n=1}^\infty |b_n|$.
b) Let Ω be a locally compact Hausdorff space and let $x_0 \in \Omega$. Define the function ℓ on $C_0(\Omega)$ by $\ell(f) = f(x_0)$. Clearly, $\ell \in C_0(\Omega)^*$ and $\|\ell\|_* \leq 1$. Since $f(x_0) = \int_\Omega f\, d\delta_{x_0}$, it follows from the uniqueness part of Theorem 13.16 that $\mu = \delta_{x_0}$. Moreover, $\|\ell\|_* = |\delta_{x_0}|(\Omega) = \delta_{x_0}(\Omega) = 1$. □

Exercises for Section 13.6

13.84 Verify the uniqueness assertion in Theorem 13.15 on page 492.

13.85 Let Ω be a locally compact, noncompact Hausdorff space and $\Omega^* = \Omega \cup \{\infty\}$ its one point compactification.
a) Show that $\mathcal{B}(\Omega) \subset \mathcal{B}(\Omega^*)$.

b) Show that $\mu \in M(\Omega)$ if and only if there exists $\mu^* \in M(\Omega^*)$ such that $\mu^*(B) = \mu(B)$ for all $B \in \mathcal{B}(\Omega)$.

13.86 Verify the assertions in Theorem 13.16 (page 493) that we did not prove.

13.87 Refer to Exercises 11.33 and 11.34 on page 410. Let Ω be a compact Hausdorff space, g a lower-semicontinuous function on Ω, and $\mu \in M_+(\Omega)$. Prove that

$$\int_\Omega g\,d\mu = \sup\left\{\int_\Omega f\,d\mu : f \in C(\Omega) \text{ and } f \le g\right\}.$$

13.88 Let Ω and Λ be compact Hausdorff spaces, $\mu \in M(\Omega)$, and $G\colon \Omega \to \Lambda$ be continuous.
a) Show that there is a $\nu \in M(\Lambda)$ such that $\int_\Lambda f\,d\nu = \int_\Omega f \circ G\,d\mu$ for all $f \in C(\Lambda)$.
b) Verify that $\nu = \mu \circ G^{-1}$, the measure induced by μ and G.

13.89 Define the linear functional ℓ on $C([0,1] \times [0,1])$ by $\ell(f) = \int_0^1 f(x,x)\,dx$. Describe explicitly the measure μ that satisfies $\ell(f) = \int_{[0,1]\times[0,1]} f\,d\mu$. *Hint:* Refer to Exercise 13.88.

13.90 In Exercise 6.64 on page 221, we defined the convolution product of two nonnegative σ-finite Borel measures on $\mathcal{R}$. An alternative definition that holds for any two (complex) Borel measures on $\mathcal{R}$ is given as follows. For $\mu, \nu \in M(\mathcal{R})$, define the convolution product of μ and ν to be the unique measure $\mu * \nu \in M(\mathcal{R})$ that satisfies

$$\int_\mathcal{R} f\,d\mu * \nu = \int_\mathcal{R}\int_\mathcal{R} f(x+y)\,d\mu(x)\,d\nu(y), \qquad f \in C_0(\mathcal{R}).$$

Show that for $\mu, \nu \in M_+(\mathcal{R})$, this definition agrees with the one presented in Exercise 6.64(d) on page 221.

13.91 Refer to Exercise 13.90. For $\mu \in M(\mathcal{R})$, find $\mu * \delta_0$.

13.92 Let Ω be a locally compact Hausdorff space and $\nu \in M_+(\Omega)$. Denote by $\mathcal{AC}(\nu)$ the collection of measures in $M(\Omega)$ that are absolutely continuous with respect to ν. Prove that $\mathcal{AC}(\nu)$ is a closed subspace of $M(\Omega)$.

13.93 Refer to Exercise 13.92. Show that $\mathcal{L}^1(\nu)$ is isometrically isomorphic to $\mathcal{AC}(\nu)$ via the correspondence $f \to \nu_f$, where $\nu_f(B) = \int_B f\,d\nu$.