Chapter 6

Hilbert Spaces

So far, in increasing order of specialization, we have studied topological spaces, metric spaces, normed linear spaces, and Banach spaces. Hilbert spaces are Banach spaces with a norm that is derived from an inner product, so they have an extra feature in comparison with arbitrary Banach spaces, which makes them still more special. We can use the inner product to introduce the notion of orthogonality in a Hilbert space, and the geometry of Hilbert spaces is in almost complete agreement with our intuition of linear spaces with an arbitrary (finite or infinite) number of orthogonal coordinate axes. By contrast, the geometry of infinite-dimensional Banach spaces can be surprisingly complicated and quite different from what naive extrapolations of the finite-dimensional situation would suggest.

6.1 Inner products

To be specific, we consider complex linear spaces throughout this chapter. We use a bar to denote the complex conjugate of a complex number. The corresponding results for real linear spaces are obtained by replacing $\mathbb{C}$ by $\mathbb{R}$ and omitting the complex conjugates.

Definition 6.1 An *inner product* on a complex linear space X is a map

$$(\cdot, \cdot) : X \times X \to \mathbb{C}$$

such that, for all $x, y, z \in X$ and $\lambda, \mu \in \mathbb{C}$:

(a) $(x, \lambda y + \mu z) = \lambda(x, y) + \mu(x, z)$ (linear in the second argument);
(b) $(y, x) = \overline{(x, y)}$ (Hermitian symmetric);
(c) $(x, x) \geq 0$ (nonnegative);
(d) $(x, x) = 0$ if and only if $x = 0$ (positive definite).

We call a linear space with an inner product an *inner product space* or a *pre-Hilbert space.*

From (a) and (b) it follows that $(\cdot, \cdot)$ is *antilinear*, or *conjugate linear*, in the first argument, meaning that

$$(\lambda x + \mu y, z) = \overline{\lambda}(x, z) + \overline{\mu}(y, z).$$

If X is real, then $(\cdot, \cdot)$ is bilinear, meaning that it is a linear function of each argument. If X is complex, then $(\cdot, \cdot)$ is said to be *sesquilinear*, a name that literally means "one-and-half" linear.

There are two conventions for the linearity of the inner product. In most of the mathematically oriented literature $(\cdot, \cdot)$ is linear in the first argument. We adopt the convention that the inner product is linear in the second argument, which is more common in applied mathematics and physics.

If X is a linear space with an inner product $(\cdot, \cdot)$, then we can define a norm on X by

$$\|x\| = \sqrt{(x, x)}. \tag{6.1}$$

Thus, any inner product space is a normed linear space. We will always use the norm defined in (6.1) on an inner product space.

Definition 6.2 A *Hilbert space* is a complete inner product space.

In particular, every Hilbert space is a Banach space with respect to the norm in (6.1).

Example 6.3 The standard inner product on $\mathbb{C}^n$ is given by

$$(x, y) = \sum_{j=1}^{n} \overline{x_j}\, y_j,$$

where $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$, with $x_j, y_j \in \mathbb{C}$. This space is complete, and therefore it is a finite-dimensional Hilbert space.

Example 6.4 Let $C([a, b])$ denote the space of all complex-valued continuous functions defined on the interval $[a, b]$. We define an inner product on $C([a, b])$ by

$$(f, g) = \int_a^b \overline{f(x)} g(x)\, dx,$$

where $f, g : [a, b] \to \mathbb{C}$ are continuous functions. This space is not complete, so it is not a Hilbert space. The completion of $C([a, b])$ with respect to the associated norm,

$$\|f\| = \left(\int_a^b |f(x)|^2 \, dx \right)^{1/2},$$

is denoted by $L^2([a, b])$. The spaces $L^p([a, b])$, defined in Example 5.6, are Banach spaces but they are not Hilbert spaces when $p \neq 2$.

Similarly, if $\mathbb{T}$ is the circle, then $L^2(\mathbb{T})$ is the Hilbert space of square-integrable functions $f : \mathbb{T} \to \mathbb{C}$ with the inner product

$$(f, g) = \int_{\mathbb{T}} \overline{f(x)} g(x)\, dx.$$

Example 6.5 We define the Hilbert space $\ell^2(\mathbb{Z})$ of bi-infinite complex sequences by

$$\ell^2(\mathbb{Z}) = \left\{ (z_n)_{n=-\infty}^{\infty} \;\middle|\; \sum_{n=-\infty}^{\infty} |z_n|^2 < \infty \right\}.$$

The space $\ell^2(\mathbb{Z})$ is a complex linear space, with the obvious operations of addition and multiplication by a scalar. An inner product on it is given by

$$(x, y) = \sum_{n=-\infty}^{\infty} \overline{x_n} y_n.$$

The name "ℓ^2" is pronounced "little ell two" to distinguish it from L^2 or "ell two" in the previous example. The space $\ell^2(\mathbb{N})$ of square-summable sequences $(z_n)_{n=1}^{\infty}$ is defined in an analogous way.

Example 6.6 Let $\mathbb{C}^{m\times n}$ denote the space of all $m \times n$ matrices with complex entries. We define an inner product on $\mathbb{C}^{m\times n}$ by

$$(A, B) = \mathrm{tr}\,(A^* B),$$

where tr denotes the trace and * denotes the Hermitian conjugate of a matrix — that is, the complex-conjugate transpose. In components, if $A = (a_{ij})$ and $B = (b_{ij})$, then

$$(A, B) = \sum_{i=1}^{m} \sum_{j=1}^{n} \overline{a}_{ij} b_{ij}.$$

This inner product is equal to the one obtained by identification of a matrix in $\mathbb{C}^{m\times n}$ with a vector in $\mathbb{C}^{mn}$. The corresponding norm,

$$\|A\| = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2},$$

is called the *Hilbert-Schmidt norm.*

Example 6.7 Let $C^k([a, b])$ be the space of functions with k continuous derivatives on $[a, b]$. We define an inner product on $C^k([a, b])$ by

$$(f, g) = \sum_{j=0}^{k} \int_a^b \overline{f^{(j)}(x)} g^{(j)}(x)\, dx,$$

where $f^{(j)}$ denotes the jth derivative of f. The corresponding norm is

$$\|f\| = \left(\sum_{j=0}^{k} \int_a^b |f^{(j)}(x)|^2 \, dx \right)^{1/2}. \tag{6.2}$$

The space $C^k([a,b])$ is an inner product space, but it is not complete. The Hilbert space obtained by completion of $C^k([a,b])$ with respect to the norm $\|\cdot\|$ is a *Sobolev space*, denoted by $H^k((a,b))$. In the notation of Example 5.7, we have

$$H^k((a,b)) = W^{k,2}((a,b)).$$

The following fundamental inequality on an inner product space is called the *Cauchy-Schwarz inequality.*

Theorem 6.8 (Cauchy-Schwarz) If $x, y \in X$, where X is an inner product space, then

$$|(x,y)| \leq \|x\|\|y\|, \tag{6.3}$$

where the norm $\|\cdot\|$ is defined in (6.1).

Proof. By the nonnegativity of the inner product, we have

$$0 \leq (\lambda x - \mu y, \lambda x - \mu y)$$

for all $x, y \in X$ and $\lambda, \mu \in \mathbb{C}$. Expansion of the inner product, and use of (6.1), implies that

$$\overline{\lambda}\mu(x,y) + \lambda\overline{\mu}(y,x) \leq |\lambda|^2\|x\|^2 + |\mu|^2\|y\|^2.$$

If $(x,y) = re^{i\varphi}$, where $r = |(x,y)|$ and $\varphi = \arg(x,y)$, then we choose

$$\lambda = \|y\|e^{i\varphi}, \qquad \mu = \|x\|.$$

It follows that

$$2\|x\|\|y\||(x,y)| \leq 2\|x\|^2\|y\|^2,$$

which proves the result. □

An inner product space is a normed space with respect to the norm defined in (6.1). The converse question of when a norm is derived from an inner product in this way is answered by the following theorem.

Theorem 6.9 A normed linear space X is an inner product space with a norm derived from the inner product by (6.1) if and only if

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \qquad \text{for all } x, y \in X. \tag{6.4}$$

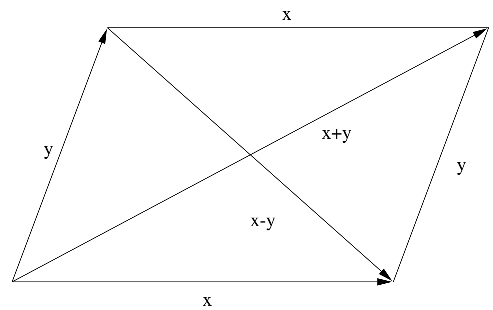

Fig. 6.1 The geometric interpretation of the parallelogram law (6.4).

Proof. Use of (6.1) to write norms in terms of inner products, and expansion of the result, implies that (6.4) holds for any norm that is derived from an inner product. Conversely, if a norm satisfies (6.4), then the equation

$$(x, y) = \frac{1}{4}\left\{\|x+y\|^2 - \|x-y\|^2 - i\|x+iy\|^2 + i\|x-iy\|^2\right\} \tag{6.5}$$

defines an inner product on X. We leave a detailed verification of this fact to the reader. □

The relation (6.4) is called the *parallelogram law*. Its geometrical interpretation is that the sum of the squares of the sides of a parallelogram is equal to the sum of the squares of the diagonals (see Figure 6.1). As the *polarization formula* (6.5) shows, an inner product is uniquely determined by its values on the diagonal, that is, by its values when the first and second arguments are equal.

Let $(X, (\cdot, \cdot)_X)$ and $(Y, (\cdot, \cdot)_Y)$ be two inner product spaces. Then there is a natural inner product on the Cartesian product space

$$X \times Y = \{(x, y) \mid x \in X, y \in Y\}$$

given by

$$((x_1, y_1), (x_2, y_2))_{X\times Y} = (x_1, x_2)_X + (y_1, y_2)_Y.$$

The associated norm on $X \times Y$ is

$$\|(x, y)\| = \sqrt{\|x\|^2 + \|y\|^2}$$

Unless stated otherwise, we will use this inner product and norm on the Cartesian product of two inner product spaces.

Theorem 6.10 Let X be an inner product space. The inner product is a continuous map from $X \times X \to \mathbb{C}$.

Proof. For all $x_1, x_2, y_1, y_2 \in X$, the Cauchy-Schwarz inequality implies that

$$\begin{aligned} |(x_1, y_1) - (x_2, y_2)| &= |(x_1 - x_2, y_1) + (x_2, y_1 - y_2)| \\ &\leq \|x_1 - x_2\| \, \|y_1\| + \|x_2\| \, \|y_1 - y_2\|. \end{aligned}$$

This estimate implies the continuity of the inner product. □

6.2 Orthogonality

Let $\mathcal{H}$ be a Hilbert space. We denote its inner product by $\langle \cdot, \cdot \rangle$, which is another common notation for inner products that is often reserved for Hilbert spaces. The inner product structure of a Hilbert space allows us to introduce the concept of orthogonality, which makes it possible to visualize vectors and linear subspaces of a Hilbert space in a geometric way.

Definition 6.11 If x, y are vectors in a Hilbert space $\mathcal{H}$, then we say that x and y are *orthogonal*, written $x \perp y$, if $\langle x, y \rangle = 0$. We say that subsets A and B are orthogonal, written $A \perp B$, if $x \perp y$ for every $x \in A$ and $y \in B$. The *orthogonal complement* $A^\perp$ of a subset A is the set of vectors orthogonal to A,

$$A^\perp = \{x \in \mathcal{H} \mid x \perp y \text{ for all } y \in A\}.$$

Theorem 6.12 The orthogonal complement of a subset of a Hilbert space is a closed linear subspace.

Proof. Let $\mathcal{H}$ be a Hilbert space and A a subset of $\mathcal{H}$. If $y, z \in A^\perp$ and $\lambda, \mu \in \mathbb{C}$, then the linearity of the inner product implies that

$$\langle x, \lambda y + \mu z \rangle = \lambda \langle x, y \rangle + \mu \langle x, z \rangle = 0 \qquad \text{for all } x \in A.$$

Therefore, $\lambda y + \mu z \in A^\perp$, so $A^\perp$ is a linear subspace.

To show that $A^\perp$ is closed, we show that if (y_n) is a convergent sequence in $A^\perp$, then the limit y also belongs to $A^\perp$. Let $x \in A$. From Theorem 6.10, the inner product is continuous and therefore

$$\langle x, y \rangle = \langle x, \lim_{n \to \infty} y_n \rangle = \lim_{n \to \infty} \langle x, y_n \rangle = 0,$$

since $\langle x, y_n \rangle = 0$ for every $x \in A$ and $y_n \in A^\perp$. Hence, $y \in A^\perp$. □

The following theorem expresses one of the fundamental geometrical properties of Hilbert spaces. While the result may appear obvious (see Figure 6.2), the proof is not trivial.

Theorem 6.13 (Projection) Let $\mathcal{M}$ be a closed linear subspace of a Hilbert space $\mathcal{H}$.

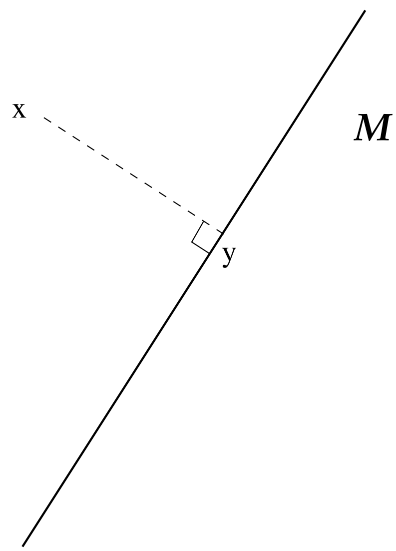

Fig. 6.2 y is the point in $\mathcal{M}$ closest to x.

(a) For each $x \in \mathcal{H}$ there is a unique closest point $y \in \mathcal{M}$ such that

$$\|x - y\| = \min_{z \in \mathcal{M}} \|x - z\|. \tag{6.6}$$

(b) The point $y \in \mathcal{M}$ closest to $x \in \mathcal{H}$ is the unique element of $\mathcal{M}$ with the property that $(x - y) \perp \mathcal{M}$.

Proof. Let d be the distance of x from $\mathcal{M}$,

$$d = \inf \{\|x - z\| \mid z \in \mathcal{M}\}. \tag{6.7}$$

First, we prove that there is a closest point $y \in \mathcal{M}$ at which this infimum is attained, meaning that $\|x - y\| = d$. From the definition of d, there is a sequence of elements $y_n \in \mathcal{M}$ such that

$$\lim_{n \to \infty} \|x - y_n\| = d.$$

Thus, for all $\epsilon > 0$, there is an N such that

$$\|x - y_n\| \leq d + \epsilon \qquad \text{when } n \geq N.$$

We show that the sequence (y_n) is Cauchy. From the parallelogram law, we have

$$\|y_m - y_n\|^2 + \|2x - y_m - y_n\|^2 = 2\|x - y_m\|^2 + 2\|x - y_n\|^2.$$

Since $(y_m + y_n)/2 \in \mathcal{M}$, equation (6.7) implies that

$$\|x - (y_m + y_n)/2\| \geq d.$$

Combining these equations, we find that for all $m, n \geq N$,

$$\begin{aligned}\|y_m - y_n\|^2 &= 2\|x - y_m\|^2 + 2\|x - y_n\|^2 - \|2x - y_m - y_n\|^2 \\ &\leq 4(d+\epsilon)^2 - 4d^2 \\ &\leq 4\epsilon(2d+\epsilon).\end{aligned}$$

Therefore, (y_n) is Cauchy. Since a Hilbert space is complete, there is a y such that $y_n \to y$, and, since $\mathcal{M}$ is closed, we have $y \in \mathcal{M}$. The norm is continuous, so $\|x - y\| = \lim_{n\to\infty} \|x - y_n\| = d$.

Second, we prove the uniqueness of a vector $y \in \mathcal{M}$ that minimizes $\|x - y\|$. Suppose y and y' both minimize the distance to x, meaning that

$$\|x - y\| = d, \qquad \|x - y'\| = d.$$

Then the parallelogram law implies that

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

Hence, since $(y + y')/2 \in \mathcal{M}$,

$$\|y - y'\|^2 = 4d^2 - 4\|x - (y + y')/2\|^2 \leq 0.$$

Therefore, $\|y - y'\| = 0$ so that $y = y'$.

Third, we show that the unique $y \in \mathcal{M}$ found above satisfies the condition that the "error" vector $x - y$ is orthogonal to $\mathcal{M}$. Since y minimizes the distance to x, we have for every $\lambda \in \mathbb{C}$ and $z \in \mathcal{M}$ that

$$\|x - y\|^2 \leq \|x - y + \lambda z\|^2.$$

Expanding the right-hand side of this equation, we obtain that

$$2\mathrm{Re}\,\lambda\langle x - y, z\rangle \leq |\lambda|^2\|z\|^2.$$

Suppose that $\langle x - y, z\rangle = |\langle x - y, z\rangle|\, e^{i\varphi}$. Choosing $\lambda = \epsilon e^{-i\varphi}$, where $\epsilon > 0$, and dividing by ϵ, we get

$$2\,|\langle x - y, z\rangle| \leq \epsilon\|z\|^2.$$

Taking the limit as $\epsilon \to 0^+$, we find that $\langle x - y, z\rangle = 0$, so $(x - y) \perp \mathcal{M}$.

Finally, we show that y is the only element in $\mathcal{M}$ such that $x - y \perp \mathcal{M}$. Suppose that y' is another such element in $\mathcal{M}$. Then $y - y' \in \mathcal{M}$, and, for any $z \in \mathcal{M}$, we have

$$\langle z, y - y'\rangle = \langle z, x - y'\rangle - \langle z, x - y\rangle = 0.$$

In particular, we may take $z = y - y'$, and therefore we must have $y = y'$. □

The proof of part (a) applies if $\mathcal{M}$ is any closed convex subset of $\mathcal{H}$ (see Exercise 6.1). Theorem 6.13 can also be stated in terms of the decomposition of $\mathcal{H}$ into an *orthogonal direct sum* of closed subspaces.

Definition 6.14 If $\mathcal{M}$ and $\mathcal{N}$ are orthogonal closed linear subspaces of a Hilbert space, then we define the *orthogonal direct sum*, or simply the *direct sum*, $\mathcal{M} \oplus \mathcal{N}$ of $\mathcal{M}$ and $\mathcal{N}$ by

$$\mathcal{M} \oplus \mathcal{N} = \{y + z \mid y \in \mathcal{M} \text{ and } z \in \mathcal{N}\}.$$

We may also define the orthogonal direct sum of two Hilbert spaces that are not subspaces of the same space (see Exercise 6.4).

Theorem 6.13 states that if $\mathcal{M}$ is a closed subspace, then any $x \in \mathcal{H}$ may be uniquely represented as $x = y + z$, where $y \in \mathcal{M}$ is the best approximation to x, and $z \perp \mathcal{M}$. We therefore have the following corollary

Corollary 6.15 If $\mathcal{M}$ is a closed subspace of a Hilbert space $\mathcal{H}$, then $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$.

Thus, every closed subspace $\mathcal{M}$ of a Hilbert space has a closed complementary subspace $\mathcal{M}^\perp$. If $\mathcal{M}$ is not closed, then we may still decompose $\mathcal{H}$ as $\mathcal{H} = \overline{\mathcal{M}} \oplus \mathcal{M}^\perp$. In a general Banach space, there may be no element of a closed subspace that is closest to a given element of the Banach space (see Exercise 6.2), and a closed linear subspace of a Banach space may have no closed complementary subspace. These facts are one indication of the much murkier geometrical properties of infinite-dimensional Banach spaces in comparison with Hilbert spaces.

6.3 Orthonormal bases

A subset U of nonzero vectors in a Hilbert space $\mathcal{H}$ is *orthogonal* if any two distinct elements in U are orthogonal. A set of vectors U is *orthonormal* if it is orthogonal and $\|u\| = 1$ for all $u \in U$, in which case the vectors u are said to be *normalized.* An *orthonormal basis* of a Hilbert space is an orthonormal set such that every vector in the space can be expanded in terms of the basis, in a way that we make precise below. In this section, we show that every Hilbert space has an orthonormal basis, which may be finite, countably infinite, or uncountable. Two Hilbert spaces whose orthonormal bases have the same cardinality are isomorphic — any linear map that identifies basis elements is an isomorphism — but many different concrete realizations of a given abstract Hilbert space arise in applications. The most important case in practice is that of a *separable* Hilbert space, which has a finite or countably infinite orthonormal basis. As shown in Exercise 6.10, this condition is equivalent to the separability of the Hilbert space as a metric space, meaning that it contains a countable dense subset.

Before studying orthonormal bases in general Hilbert spaces, we give some examples.

Example 6.16 A set of vectors $\{e_1, \ldots, e_n\}$ is an orthonormal basis of the finite-dimensional Hilbert spaces $\mathbb{C}^n$ if:

(a) $\langle e_j, e_k \rangle = \delta_{jk}$ for $1 \leq j, k \leq n$;
(b) for all $x \in \mathbb{C}^n$ there are unique coordinates $x_k \in \mathbb{C}$ such that

$$x = \sum_{k=1}^{n} x_k e_k, \tag{6.8}$$

where δ_{jk} is the Kronecker delta defined in (5.25). The orthonormality of the basis implies that $x_k = \langle e_k, x \rangle$. For example, the standard orthonormal basis of $\mathbb{C}^n$ consists of the vectors

$$e_1 = (1, 0, \ldots, 0), \quad e_2 = (0, 1, \ldots, 0), \ldots, \quad e_n = (0, 0, \ldots, 1).$$

Example 6.17 Consider the Hilbert space $\ell^2(\mathbb{Z})$ defined in Example 6.5. An orthonormal basis of $\ell^2(\mathbb{Z})$ is the set of coordinate basis vectors $\{e_n \mid n \in \mathbb{Z}\}$ given by

$$e_n = (\delta_{kn})_{k=-\infty}^{\infty}.$$

For example,

$$e_{-1} = (\ldots, 0, 1, 0, 0, 0, \ldots), \ e_0 = (\ldots, 0, 0, 1, 0, 0, \ldots), \ e_1 = (\ldots, 0, 0, 0, 1, 0, \ldots).$$

Example 6.18 The set of functions $\{e_n(x) \mid n \in \mathbb{Z}\}$, given by

$$e_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx},$$

is an orthonormal basis of the space $L^2(\mathbb{T})$ of 2π-periodic functions, called the *Fourier basis*. We will study it in detail in the next chapter. As we will see, the inverse Fourier transform $\mathcal{F}^{-1} : \ell^2(\mathbb{Z}) \to L^2(\mathbb{T})$, defined by

$$\mathcal{F}^{-1}(c_k) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} c_k e^{ikx},$$

is a Hilbert space isomorphism between $\ell^2(\mathbb{Z})$ and $L^2(\mathbb{T})$. Both Hilbert spaces are separable with a countably infinite basis.

Example 6.19 A function that is a sum of finitely many periodic functions is said to be *quasiperiodic*. If the ratios of the periods of the terms in the sum are rational, then the sum is itself periodic, but if at least one of the ratios is irrational, then the sum is not periodic. For example,

$$f(t) = e^{it} + e^{i\pi t}$$

is quasiperiodic but not periodic. Let X be the space of quasiperiodic functions $f : \mathbb{R} \to \mathbb{C}$ of the form

$$f(t) = \sum_{k=1}^{n} a_k e^{i\omega_k t},$$

where $n \in \mathbb{N}$, $a_k \in \mathbb{C}$, and $\omega_k \in \mathbb{R}$ are arbitrary constants. We may think of t as a time variable, in which case f is a sum of time-harmonic functions with amplitudes $|a_k|$, phases $\arg a_k$, and frequencies ω_k. When some of the frequencies are incommensurable, the function f "almost" repeats itself, but it is not periodic with any period, although it is bounded.

We define an inner product on X by means of the time average,

$$\langle f, g\rangle = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \overline{f(t)} g(t)\, dt.$$

If $f(t) = \sum_{k=1}^{n} a_k e^{i\omega_k t}$ and $g(t) = \sum_{k=1}^{n} b_k e^{i\omega_k t}$, where $\omega_j \neq \omega_k$ for $j \neq k$, then

$$\langle f, g\rangle = \sum_{k=1}^{n} \overline{a_k} b_k.$$

The inner product may also be written as

$$\langle f, g\rangle = \lim_{T\to\infty} \frac{1}{T}\int_{t_0}^{t_0+T} \overline{f(t)} g(t)\, dt,$$

where t_0 is any fixed time independent of T. The set of functions

$$\left\{e^{i\omega t} \mid \omega \in \mathbb{R}\right\} \tag{6.9}$$

is an orthonormal set in X. The space X is an inner product space, but it is not complete. We call the completion of X the space of L^2*-almost periodic functions.* This space consists of equivalence classes of functions of the form

$$f(t) = \sum_{k=1}^{\infty} a_k e^{i\omega_k t}, \tag{6.10}$$

where $\sum_{k=1}^{\infty} |a_k|^2 < \infty$. The sum converges in norm, meaning that for any $t_0 \in \mathbb{R}$,

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \left| f(t) - \sum_{k=1}^{n} a_k e^{i\omega_k t}\right|^2 dt \to 0 \quad \text{as } n \to \infty.$$

The set in (6.9) is an uncountable orthonormal basis of this Hilbert space, so the space is not separable.

Although in the future we will mainly consider separable Hilbert spaces, it is worth postponing this restriction for a little while. First, we say what we mean by a sum with a possibly uncountable number of terms. This definition also clarifies the sense in which our infinite sums converge, which is stronger than the sense in which infinite series converge.

Definition 6.20 Let $\{x_\alpha \in X \mid \alpha \in I\}$ be an indexed set in a Banach space X, where the index set I may be countable or uncountable. For each finite subset J of I, we define the partial sum S_J by

$$S_J = \sum_{\alpha \in J} x_\alpha.$$

The *unordered sum* of the indexed set $\{x_\alpha \mid \alpha \in I\}$ converges to $x \in X$, written

$$x = \sum_{\alpha \in I} x_\alpha, \tag{6.11}$$

if for every $\epsilon > 0$ there is a finite subset J^ϵ of I such that $\|S_J - x\| < \epsilon$ for all finite subsets J of I that contain J^ϵ. An unordered sum is said to *converge unconditionally.*

All the sums in this chapter are to be interpreted as unordered sums. The convergence of finite partial sums S_J, indexed by finite subsets J of I, is a special case of the convergence of nets [12]. It is easy to see that an unordered sum converges if and only if any permutation of its terms converges, and the sum is independent of the ordering of its terms.

A sum $\sum_{\alpha \in I} x_\alpha$ is said to *converge absolutely* if the sum $\sum_{\alpha \in I} \|x_\alpha\|$ of non-negative numbers converges unconditionally. The unordered sum of a sequence of real or complex numbers exists if and only if the corresponding series is absolutely convergent. An absolutely convergent sum in an infinite-dimensional Banach space converges unconditionally, but an unconditionally convergent sum need not converge absolutely (see Exercise 6.8 for an example).

If an unordered sum $\sum_{\alpha \in I} x_\alpha$ converges to x, then for each $n \in \mathbb{N}$, there is a finite $J_n \subset I$ such that for all J containing J_n, one has $\|S_J - x\| \leq 1/n$. It follows that $x_\alpha = 0$ if $\alpha \notin \bigcup_{n \in \mathbb{N}} J_n$, so a convergent unordered sum has only countably many nonzero terms. Moreover, there is a sequence (S_{J_n}) of finite partial sums that converges to x as $n \to \infty$. The continuity of the inner product implies that

$$\left\langle \sum_{\alpha \in I} x_\alpha, \sum_{\beta \in J} y_\beta \right\rangle = \sum_{(\alpha, \beta) \in I \times J} \langle x_\alpha, y_\beta \rangle.$$

There is a generalization of the Cauchy criterion for the convergence of series to unordered sums.

Definition 6.21 An unordered sum $\sum_{\alpha \in I} x_\alpha$ is *Cauchy* if for every $\epsilon > 0$ there is a finite set $J^\epsilon \subset I$ such that $\|S_K\| < \epsilon$ for every finite set $K \subset I \setminus J^\epsilon$.

Proposition 6.22 An unordered sum in a Banach space converges if and only if it is Cauchy.

Proof. First, suppose that the unordered sum $\sum_{\alpha\in I} x_\alpha$ converges to x. Let $\epsilon > 0$. By the definition of convergence, there is a finite set J^ϵ such that $\|S_J - x\| < \epsilon/2$ for all finite sets J that contain J^ϵ. If K is any finite subset of $I \setminus J^\epsilon$, then we let $J = J^\epsilon \cup K$. Since J contains J^ϵ, we have

$$\|S_K\| = \|S_J - S_{J^\epsilon}\| \leq \|S_J - x\| + \|x - S_{J^\epsilon}\| < \epsilon.$$

Hence, the sequence is Cauchy.

Conversely, suppose that an unordered sum is Cauchy. Let J_n be finite subsets of I such that $\|S_K\| < 1/n$ for every $K \subset I \setminus J_n$. Without loss of generality, we may assume that $J_n \subset J_{n+1}$ for all n. It follows that for all $n \leq m$ we have $\|S_{J_m} - S_{J_n}\| < 1/n$, which shows that the sequence (S_{J_n}) is Cauchy; hence, since a Banach space is complete, it converges to a point x. To complete the proof, we show that the unordered sum converges to x. Given $\epsilon > 0$, we pick n such that $1/n < \epsilon/2$ and put $J^\epsilon = J_n$. If J is a finite set that contains J^ϵ, then the Cauchy criterion for the set J_n and the convergence of the sequence (S_{J_n}) to x imply that

$$\|S_J - x\| \leq \|S_J - S_{J_n}\| + \|S_{J_n} - x\| < \frac{2}{n} < \epsilon.$$

□

We may use the Cauchy criterion to give a simple necessary and sufficient condition for the unconditional convergence of a sum of orthogonal terms in a Hilbert space.

Lemma 6.23 Let $U = \{u_\alpha \mid \alpha \in I\}$ be an indexed, orthogonal subset of a Hilbert space $\mathcal{H}$. The sum $\sum_{\alpha\in I} u_\alpha$ converges unconditionally if and only if $\sum_{\alpha\in I} \|u_\alpha\|^2 < \infty$, and, in that case,

$$\left\|\sum_{\alpha\in I} u_\alpha\right\|^2 = \sum_{\alpha\in I} \|u_\alpha\|^2. \tag{6.12}$$

Proof. For any finite set J we have

$$\left\|\sum_{\alpha\in J} u_\alpha\right\|^2 = \sum_{\alpha,\beta\in J} \langle u_\alpha, u_\beta\rangle = \sum_{\alpha\in J} \langle u_\alpha, u_\alpha\rangle = \sum_{\alpha\in J} \|u_\alpha\|^2.$$

It follows that the Cauchy criterion is satisfied for $\sum_{\alpha\in I} u_\alpha$ if and only if it is satisfied for $\sum_{\alpha\in I} \|u_\alpha\|^2$. Thus, one of the sums converges unconditionally if and only if the other does. Equation (6.12) follows because the sum is the limit of a sequence of finite partial sums and the norm is a continuous function. □

When combined with the following basic estimate, this lemma will imply that every element of a Hilbert space can be expanded with respect to an orthonormal basis.

Theorem 6.24 (Bessel's inequality) Let $U = \{u_\alpha \mid \alpha \in I\}$ be an orthonormal set in a Hilbert space $\mathcal{H}$ and $x \in \mathcal{H}$. Then:

(a) $\sum_{\alpha \in I} |\langle u_\alpha, x\rangle|^2 \leq \|x\|^2$;
(b) $x_U = \sum_{\alpha \in I} \langle u_\alpha, x\rangle u_\alpha$ is a convergent sum;
(c) $x - x_U \in U^\perp$.

Proof. We begin by computing $\|x - \sum_{\alpha \in J}\langle u_\alpha, x\rangle u_\alpha\|$ for any finite subset $J \subset I$:

$$\begin{aligned}
\left\| x - \sum_{\alpha \in J}\langle u_\alpha, x\rangle u_\alpha \right\|^2 &= \left\langle \left(x - \sum_{\alpha \in J}\langle u_\alpha, x\rangle u_\alpha \right), \left(x - \sum_{\beta \in J}\langle u_\beta, x\rangle u_\beta \right) \right\rangle \\
&= \langle x, x\rangle - \sum_{\beta \in J}\langle u_\beta, x\rangle\langle x, u_\beta\rangle - \sum_{\alpha \in J}\overline{\langle u_\alpha, x\rangle}\langle u_\alpha, x\rangle \\
&\quad + \sum_{\alpha,\beta \in J}\overline{\langle u_\alpha, x\rangle}\langle u_\beta, x\rangle\langle u_\alpha, u_\beta\rangle \\
&= \|x\|^2 - \sum_{\alpha \in J}|\langle u_\alpha, x\rangle|^2.
\end{aligned}$$

Hence

$$\sum_{\alpha \in J}|\langle u_\alpha, x\rangle|^2 = \|x\|^2 - \left\| x - \sum_{\alpha \in J}\langle u_\alpha, x\rangle u_\alpha \right\|^2 \leq \|x\|^2.$$

Since $\sum_{\alpha \in I}|\langle u_\alpha, x\rangle|^2$ is a sum of nonnegative numbers that is bounded from above by $\|x\|^2$, it is Cauchy. Therefore the sum converges and satisfies (a). The convergence claimed in (b) follows from an application of Lemma 6.23.

In order to prove (c), we consider any $u_{\alpha_0} \in U$. Using the orthonormality of U and the continuity of the inner product, we find that

$$\begin{aligned}
\left\langle x - \sum_{\alpha \in I}\langle u_\alpha, x\rangle u_\alpha, u_{\alpha_0} \right\rangle &= \langle x, u_{\alpha_0}\rangle - \sum_{\alpha \in I}\overline{\langle u_\alpha, x\rangle}\langle u_\alpha, u_{\alpha_0}\rangle \\
&= \langle x, u_{\alpha_0}\rangle - \langle x, u_{\alpha_0}\rangle = 0.
\end{aligned}$$

Hence, $x - \sum_{\alpha \in I}\langle u_\alpha, x\rangle u_\alpha \in U^\perp$. □

Given a subset U of $\mathcal{H}$, we define the closed linear span $[U]$ of U by

$$[U] = \left\{ \sum_{u \in U} c_u u \;\middle|\; c_u \in \mathbb{C} \text{ and } \textstyle\sum_{u \in U} c_u u \text{ converges unconditionally} \right\}. \tag{6.13}$$

Equivalently, $[U]$ is the smallest closed linear subspace that contains U. We leave the proof of the following lemma to the reader.

Lemma 6.25 If $U = \{u_\alpha \mid \alpha \in I\}$ is an orthonormal set in a Hilbert space $\mathcal{H}$, then

$$[U] = \left\{ \sum_{\alpha \in I} c_\alpha u_\alpha \;\middle|\; c_\alpha \in \mathbb{C} \text{ such that } \textstyle\sum_{\alpha \in I}|c_\alpha|^2 < \infty \right\}.$$

By combining Theorem 6.13 and Theorem 6.24 we see that x_U, defined in part (b) of Theorem 6.24, is the unique element of $[U]$ satisfying

$$\|x - x_U\| = \min_{u \in [U]} \|x - u\|.$$

In particular, if $[U] = \mathcal{H}$, then $x_U = x$, and every $x \in \mathcal{H}$ may be expanded in terms of elements of U. The following theorem gives equivalent conditions for this property of U, called *completeness*.

Theorem 6.26 If $U = \{u_\alpha \mid \alpha \in I\}$ is an orthonormal subset of a Hilbert space $\mathcal{H}$, then the following conditions are equivalent:

(a) $\langle u_\alpha, x\rangle = 0$ for all $\alpha \in I$ implies $x = 0$;
(b) $x = \sum_{\alpha \in I} \langle u_\alpha, x\rangle u_\alpha$ for all $x \in \mathcal{H}$;
(c) $\|x\|^2 = \sum_{\alpha \in I} |\langle u_\alpha, x\rangle|^2$ for all $x \in \mathcal{H}$;
(d) $[U] = \mathcal{H}$;
(e) U is a maximal orthonormal set.

Proof. We prove that (a) implies (b), (b) implies (c), (c) implies (d), (d) implies (e), and (e) implies (a). The condition in (a) states that $U^\perp = \{0\}$. Part (c) of Theorem 6.24 then implies (b). The fact that (b) implies (c) follows from Lemma 6.23. To prove that (c) implies (d), we observe that (c) implies that $U^\perp = \{0\}$, which implies that $[U]^\perp = \{0\}$, so $[U] = \mathcal{H}$. Condition (e) means that if V is a subset of $\mathcal{H}$ that contains U and is strictly larger than U, then V is not orthonormal. To prove that (d) implies (e), we note from (d) that any $v \in \mathcal{H}$ is of the form $v = \sum_{\alpha \in I} c_\alpha u_\alpha$, where $c_\alpha = \langle u_\alpha, v\rangle$. Therefore, if $v \perp U$ then $c_\alpha = 0$ for all α, and hence $v = 0$, so $U \cup \{v\}$ is not orthonormal. Finally, (e) implies (a), since (a) is just a reformulation of (e). □

In view of this theorem, we can make the following definition.

Definition 6.27 An orthonormal subset $U = \{u_\alpha \mid \alpha \in I\}$ of a Hilbert space $\mathcal{H}$ is *complete* if it satisfies any of the equivalent conditions (a)–(e) in Theorem 6.26. A complete orthonormal subset of $\mathcal{H}$ is called an *orthonormal basis* of $\mathcal{H}$.

Condition (a) is often the easiest condition to verify. Condition (b) is the property that is used most often. Condition (c) is called *Parseval's identity*. Condition (d) simply expresses completeness of the basis, and condition (e) will be used in the proof of the existence of an orthonormal basis in an arbitrary Hilbert space (see Theorem 6.29).

The following generalization of Parseval's identity shows that a Hilbert space $\mathcal{H}$ with orthonormal basis $\{u_\alpha \mid \alpha \in I\}$ is isomorphic to the sequence space $\ell^2(I)$.

Theorem 6.28 (Parseval's identity) Suppose that $U = \{u_\alpha \mid \alpha \in I\}$ is an orthonormal basis of a Hilbert space $\mathcal{H}$. If $x = \sum_{\alpha \in I} a_\alpha u_\alpha$ and $y = \sum_{\alpha \in I} b_\alpha u_\alpha$, where $a_\alpha = \langle u_\alpha, x\rangle$ and $b_\alpha = \langle u_\alpha, y\rangle$, then

$$\langle x, y\rangle = \sum_{\alpha\in I} \overline{a_\alpha}\, b_\alpha .$$

To show that every Hilbert space has an orthonormal basis, we use *Zorn's lemma*, which states that a nonempty partially ordered set with the property that every totally ordered subset has an upper bound has a maximal element [49].

Theorem 6.29 Every Hilbert space $\mathcal{H}$ has an orthonormal basis. If U is an orthonormal set, then $\mathcal{H}$ has an orthonormal basis containing U.

Proof. If $\mathcal{H} = \{0\}$, then the statement is trivially true with $U = \emptyset$, so we assume that $\mathcal{H} \neq \{0\}$. We introduce a partial ordering $\leq$ on orthonormal subsets of $\mathcal{H}$ by inclusion, so that $U \leq V$ if and only if $U \subset V$. If $\{U_\alpha \mid \alpha \in \mathcal{A}\}$ is a totally ordered family of orthonormal sets, meaning that for any $\alpha, \beta \in \mathcal{A}$ we have either $U_\alpha \leq U_\beta$ or $U_\beta \leq U_\alpha$, then $\bigcup_{\alpha\in\mathcal{A}} U_\alpha$ is an orthonormal set and is an upper bound, in the sense of inclusion, of the family $\{U_\alpha \mid \alpha \in \mathcal{A}\}$. Zorn's Lemma implies that the family of all orthonormal sets in $\mathcal{H}$ has a maximal element. This element satisfies (e) in Theorem 6.26, and hence is a basis. To prove that any orthonormal set U can be extended to an orthonormal basis of $\mathcal{H}$, we apply the same argument to the family of all orthonormal sets containing U. □

The existence of orthonormal bases would not be useful if we did not have a means of constructing them. The Gram-Schmidt orthonormalization procedure is an algorithm for the construction of an orthonormal basis from any countable linearly independent set whose linear span is dense in $\mathcal{H}$.

Let V be a countable set of linearly independent vectors in a Hilbert space $\mathcal{H}$. The Gram-Schmidt orthonormalization procedure is a method of constructing an orthonormal set U such that $[U] = [V]$, where the closed linear span $[V]$ of V is defined in (6.13). We denote the elements of V by v_n. The orthonormal set $U = \{u_n\}$ is then constructed inductively by setting $u_1 = v_1/\|v_1\|$, and

$$u_{n+1} = c_{n+1}\left(v_{n+1} - \sum_{k=1}^{n} \langle u_k, v_{n+1}\rangle u_k\right)$$

for all $n \geq 1$. Here $c_{n+1} \in \mathbb{C}$ is chosen so that $\|u_{n+1}\| = 1$. It is straightforward to check that $[\{v_1, \ldots, v_n\}] = [\{u_1, \ldots, u_n\}]$ for all $n \geq 1$, and hence that

$$[V] = \overline{\bigcup_n [\{v_1, \ldots, v_n\}]} = \overline{\bigcup_n [\{u_1, \ldots, u_n\}]} = [U].$$

Example 6.30 Let $(a, b) \subset \mathbb{R}$ be a finite or infinite interval and $w : (a, b) \to \mathbb{R}$ a continuous function such that $w(x) > 0$ for $a < x < b$. We define a weighted inner product on

$$C_w([a, b]) = \left\{ f : [a, b] \to \mathbb{C} \mid f \text{ continuous and } \int_a^b w(x)|f(x)|^2\, dx < \infty \right\}$$

by

$$\langle f, g \rangle = \int_a^b w(x)\overline{f(x)}g(x)\,dx.$$

Let $L_w^2([a,b])$ be the Hilbert space obtained by the completion of $C_w([a,b])$ with respect to the norm derived from this inner product. The Gram-Schmidt procedure applied to the set of monomials $\{x^n \mid n \geq 0\}$ gives an orthonormal basis of polynomials for this Hilbert space. The simplest case is that of the space $L^2([-1,1])$, with the usual unweighted inner product, which leads to the Legendre polynomials (see Exercise 6.12). The Tchebyschev polynomials are obtained from Gram-Schmidt orthonormalization of the monomials in $L_w^2([-1,1])$ where $w(x) = \left(1 - x^2\right)^{1/2}$ (see Exercise 6.13). The Hermite polynomials are obtained by Gram-Schmidt orthonormalization of the monomials in the space $L_w^2(\mathbb{R})$ with the Gaussian weight function $w(x) = e^{-x^2/2}$ (see Exercise 6.14). For a description of other polynomials that arise in this way, such as the Jacobi and Laguerre polynomials, see [5].

6.4 Hilbert spaces in applications

In this section, we describe several applications in which Hilbert spaces arise naturally.

The first is quantum mechanics. The introduction of quantum mechanics in the 1920s represents one of the most profound shifts in history of our understanding of the physical world. The theory developed at a feverish pace, and people hardly had time to pause to think about the mathematical structures they were inventing and using. Only later was it realized, by von Neumann, that Hilbert spaces are the natural setting for quantum mechanics.

One of the simplest quantum mechanical systems consists of a particle, such as an electron, confined to move in a straight line between two parallel walls: the "particle in a box." Quantum effects are important when the kinetic energy of the particle is comparable with $E = \hbar^2/(2mL^2)$, where m is the mass of the particle, $\hbar$ is Planck's constant, and L is the distance between the walls. Planck's constant has the dimensions of action, or energy times time, so E has the dimensions of energy.

In quantum mechanics, the state of the particle at each instant in time t is described by an element $\psi(\cdot, t) \in L^2([0, L])$, that is, a vector in the Hilbert space of square-integrable, complex-valued functions on the interval $[0, L]$. The function ψ is called the wavefunction of the particle. This description contrasts with classical, Newtonian mechanics, where the state of the particle is described by just two numbers: the position $0 \leq x \leq L$ and the velocity $v \in \mathbb{R}$. The physical interpretation of the wavefunction is that $|\psi|^2$ is a probability density. If the position x of the particle is measured at some time t, then the probability of observing the particle in some interval $[a, b]$, where $0 \leq a < b \leq L$, is given by

$$\Pr\left[\text{particle is in the interval } [a,b] \text{ at time } t\right] = \frac{\int_a^b |\psi(x,t)|^2\, dx}{\int_0^L |\psi(x,t)|^2\, dx}.$$

The dynamics of the quantum mechanical particle is described by a partial differential equation for the wavefunction, called the Schrödinger equation. For the particle in a box, the Schrödinger equation is

$$i\hbar\psi_t = -\frac{\hbar^2}{2m}\psi_{xx}, \qquad x \in [0, L],\ t \in \mathbb{R}, \tag{6.14}$$

with the boundary conditions $\psi(0, t) = \psi(L, t) = 0$ for all $t \in \mathbb{R}$.

A second way in which L^2-spaces arise naturally is as "energy" spaces. The quantity

$$\int |f(x)|^2\, dx \tag{6.15}$$

often represents the total energy of a physical system, or some other fundamental quantity, and one often wants to restrict attention to systems for which this quantity is finite. For example, in fluid mechanics, if $\mathbf{u}(\mathbf{x})$ is the velocity of a fluid at the point $\mathbf{x}$, then

$$\int_V |\mathbf{u}(\mathbf{x})|^2\, d\mathbf{x},$$

where $|\cdot|$ denotes the Euclidean norm of a vector, is proportional to the kinetic energy of the fluid in V. This energy should be finite for any region V with finite volume. An electromagnetic field is described by two vector fields, the electric field $\mathbf{E}$ and the magnetic field $\mathbf{B}$. In suitable units, the energy of the electromagnetic field in a region V is given by

$$\int_V \left\{|\mathbf{E}(\mathbf{x})|^2 + |\mathbf{B}(\mathbf{x})|^2\right\}\, d\mathbf{x}.$$

The requirement of finite energy leads naturally to the requirement that $\mathbf{E}$ and $\mathbf{B}$ belong to appropriate L^2-spaces.

A third area in which Hilbert spaces arise naturally is in probability theory. As we discuss in greater detail in Chapter 12, a random experiment is modeled mathematically by a space Ω, called the sample space, and a probability measure P on Ω. Each point $\omega \in \Omega$ corresponds to a possible outcome of the experiment. An event A is a measurable subset of Ω. The probability measure P associates with each event A a probability $P(A)$, where $0 \leq P(A) \leq 1$ and $P(\Omega) = 1$.

A *random variable* X is a measurable function $X : \Omega \to \mathbb{C}$, which associates a number $X(\omega)$ with each possible outcome $\omega \in \Omega$. The *expected value* $\mathbb{E}X$ of a random variable X is the mean, or integral, of the random variable X with respect to the probability measure P,

$$\mathbb{E}X = \int_{\Omega} X(\omega)\, dP(\omega).$$

A random variable X is said to be *second-order* if

$$\mathbb{E}|X|^2 < \infty.$$

The set of second-order random variables forms a Hilbert space with respect to the inner product

$$\langle X, Y\rangle = \mathbb{E}\left[\overline{X}Y\right],$$

where we identify random variables that are equal almost surely. Here, "almost surely" is the probabilistic terminology for "almost everywhere," so that two random variables are equal almost surely if they are equal on a subset of Ω which has probability one. The space of second-order random variables may be identified with the space $L^2(\Omega, P)$ of square-integrable functions on (Ω, P), with the inner product

$$\langle X, Y\rangle = \int_{\Omega} \overline{X(\omega)} Y(\omega)\, dP(\omega).$$

The Cauchy-Schwarz inequality and the fact that $\mathbb{E}1 = 1$ imply that a second-order random variable has finite mean, since

$$|\mathbb{E}X| = |\langle 1, X\rangle| \leq \mathbb{E}\left[|X|^2\right]^{1/2}.$$

Thus, the Hilbert space of second-order random variables consists of the random variables with finite mean and finite variance, where the *variance* $\operatorname{Var} X$ of a random variable X is defined by

$$\operatorname{Var} X = \mathbb{E}\left[|X - \mathbb{E}X|^2\right].$$

Two random variables X, Y are *uncorrelated* if

$$\mathbb{E}[XY] = \mathbb{E}X\, \mathbb{E}Y.$$

In particular, two random variables with zero mean are uncorrelated if and only if they are orthogonal.

6.5 References

The material of this chapter's introduction to Hilbert space is covered in Chapter 4 of Rudin [49], and also in Simmons [50]. Halmos [20] contains a large number of problems on Hilbert spaces, together with hints and solutions. For an introduction to probability theory, see Grimmett and Stirzaker [17].

6.6 Exercises

Exercise 6.1 Prove that a closed, convex subset of a Hilbert space has a unique point of minimum norm.

Exercise 6.2 Consider $C([0,1])$ with the sup-norm. Let

$$N = \left\{ f \in C([0,1]) \;\middle|\; \int_0^1 f(x)\,dx = 0 \right\}$$

be the closed linear subspace of $C([0,1])$ of functions with zero mean. Let

$$X = \left\{ f \in C([0,1]) \;\middle|\; f(0) = 0 \right\}$$

and define $M = N \cap X$, meaning that

$$M = \left\{ f \in C([0,1]) \;\middle|\; f(0) = 0, \quad \int_0^1 f(x)\,dx = 0 \right\}.$$

(a) If $u \in C([0,1])$, prove that

$$d(u, N) = \inf_{n \in N} \|u - n\| = |\overline{u}|,$$

where $|\overline{u}| = \int_0^1 u(x)\,dx$ is the mean of u, so the infimum is attained when $n = u - \overline{u} \in N$.

(b) If $u(x) = x \in X$, show that

$$d(x, M) = \inf_{m \in M} \|u - m\| = 1/2,$$

but that the infimum is not attained for any $m \in M$.

Exercise 6.3 If A is a subset of a Hilbert space, prove that

$$A^\perp = \overline{A}^\perp,$$

where $\overline{A}$ is the closure of A. If $\mathcal{M}$ is a linear subspace of a Hilbert space, prove that

$$\mathcal{M}^{\perp\perp} = \overline{\mathcal{M}}.$$

Exercise 6.4 Suppose that $\mathcal{H}_1$ and $\mathcal{H}_2$ are two Hilbert spaces. We define

$$\mathcal{H}_1 \oplus \mathcal{H}_2 = \{(x_1, x_2) \mid x_1 \in \mathcal{H}_1,\ x_2 \in \mathcal{H}_2\}$$

with the inner product

$$\langle (x_1, x_2), (y_1, y_2) \rangle_{\mathcal{H}_1 \oplus \mathcal{H}_2} = \langle x_1, y_1 \rangle_{\mathcal{H}_1} + \langle x_2, y_2 \rangle_{\mathcal{H}_2}.$$

Prove that $\mathcal{H}_1 \oplus \mathcal{H}_2$ is a Hilbert space. Find the orthogonal complement of the subspace $\{(x_1, 0) \mid x_1 \in \mathcal{H}_1\}$.

Exercise 6.5 Suppose that $\{\mathcal{H}_n \mid n \in \mathbb{N}\}$ is a set of orthogonal closed subspaces of a Hilbert space $\mathcal{H}$. We define the infinite direct sum

$$\bigoplus_{n=1}^{\infty} \mathcal{H}_n = \left\{ \sum_{n=1}^{\infty} x_n \;\middle|\; x_n \in \mathcal{H}_n \text{ and } \sum_{n=1}^{\infty} \|x_n\|^2 < \infty \right\}.$$

Prove that $\bigoplus_{n=1}^{\infty} \mathcal{H}_n$ is a closed linear subspace of $\mathcal{H}$.

Exercise 6.6 Prove that the vectors in an orthogonal set are linearly independent.

Exercise 6.7 Let $\{x_\alpha\}_{\alpha \in I}$ be a family of nonnegative real numbers. Prove that

$$\sum_{\alpha \in I} x_\alpha = \sup \left\{ \sum_{\alpha \in J} x_\alpha \;\middle|\; J \subset I \text{ and } J \text{ is finite} \right\}.$$

Exercise 6.8 Let $\{x_n \mid n \in \mathbb{N}\}$ be an orthonormal set in a Hilbert space. Show that the sum $\sum_{n=1}^{\infty} x_n/n$ converges unconditionally but not absolutely.

Exercise 6.9 Prove Lemma 6.25.

Exercise 6.10 Prove that a Hilbert space is a separable metric space if and only if it has a countable orthonormal basis.

Exercise 6.11 Prove that if $\mathcal{M}$ is a dense linear subspace of a separable Hilbert space $\mathcal{H}$, then $\mathcal{H}$ has an orthonormal basis consisting of elements in $\mathcal{M}$. Does the same result hold for arbitrary dense subsets of $\mathcal{H}$?

Exercise 6.12 Define the *Legendre polynomials* P_n by

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} \left(x^2 - 1\right)^n.$$

(a) Compute the first few Legendre polynomials, and compare with what you get by Gram-Schmidt orthogonalization of the monomials $\{1, x, x^2, \ldots\}$ in $L^2([-1, 1])$.
(b) Show that the Legendre polynomials are orthogonal in $L^2([-1, 1])$, and that they are obtained by Gram-Schmidt orthogonalization of the monomials.
(c) Show that

$$\int_{-1}^{1} P_n(x)^2 \, dx = \frac{2}{2n+1}.$$

(d) Prove that the Legendre polynomials form an orthogonal basis of $L^2([-1, 1])$. Suppose that $f \in L^2([-1, 1])$ is given by

$$f(x) = \sum_{n=0}^{\infty} c_n P_n(x).$$

Compute c_n and say explicitly in what sense the series converges.

(e) Prove that the Legendre polynomial P_n is an eigenfunction of the differential operator

$$L = -\frac{d}{dx}\left(1-x^2\right)\frac{d}{dx}$$

with eigenvalue $\lambda_n = n(n+1)$, meaning that

$$LP_n = \lambda_n P_n.$$

Exercise 6.13 Let $\mathcal{H}$ be the Hilbert space of functions $f : [-1,1] \to \mathbb{C}$ such that

$$\int_{-1}^{1} \frac{|f(x)|^2}{\sqrt{1-x^2}}\, dx < \infty,$$

with the inner-product

$$\langle f, g\rangle = \int_{-1}^{1} \frac{\overline{f(x)}g(x)}{\sqrt{1-x^2}}\, dx.$$

Show that the *Tchebyshev polynomials*,

$$T_n(x) = \cos(n\theta) \qquad \text{where } \cos\theta = x \text{ and } 0 \le \theta \le \pi,$$

$n = 0, 1, 2, \ldots$, form an orthogonal set in $\mathcal{H}$, and

$$\|T_0\| = \sqrt{\pi}, \qquad \|T_n\| = \sqrt{\frac{\pi}{2}} \quad n \ge 1.$$

Exercise 6.14 Define the Hermite polynomials H_n by

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n}\left(e^{-x^2}\right).$$

(a) Show that

$$\varphi_n(x) = e^{-x^2/2} H_n(x)$$

is an orthogonal set in $L^2(\mathbb{R})$.

(b) Show that the nth Hermite function φ_n is an eigenfunction of the linear operator

$$H = -\frac{d^2}{dx^2} + x^2$$

with eigenvalue

$$\lambda_n = 2n + 1.$$

HINT: Let

$$A = \frac{d}{dx} + x, \qquad A^* = -\frac{d}{dx} + x.$$

Show that

$$A\varphi_n = 2n\varphi_{n-1}, \quad A^*\varphi_n = \varphi_{n+1}, \quad H = AA^* - 1.$$

In quantum mechanics, H is the *Hamiltonian operator* of a simple harmonic oscillator, and A^* and A are called *creation* and *annihilation*, or *ladder*, operators.