5

Hilbert Spaces

The basic concepts of finite-dimensional vector spaces introduced in Chapter 1 can readily be generalized to infinite dimensions. The definition of a vector space and concepts of linear combination, linear independence, basis, subspace, span, and so forth all carry over to infinite dimensions. However, one thing is crucially different in the new situation, and this difference makes the study of infinite-dimensional vector spaces both richer and more nontrivial: In a finite-dimensional vector space we dealt with finite sums; in infinite dimensions we encounter infinite sums. Thus, we have to investigate the convergence of such sums.

5.1 The Question of Convergence

The intuitive notion of convergence acquired in calculus makes use of the idea of closeness. This, in turn, requires the notion of distance.[1] We considered such a notion in Chapter 1 in the context of a norm, and saw that the inner product had an associated norm. However, it is possible to introduce a norm on a vector space without an inner product.

One such norm, applicable to $\mathbb{C}^n$ and $\mathbb{R}^n$, was

$$\|a\|_p \equiv \left(\sum_{i=1}^{n} |\alpha_i|^p\right)^{1/p},$$

where p is an integer. The "natural" norm, i.e., that induced on $\mathbb{C}^n$ (or $\mathbb{R}^n$) by the usual inner product, corresponds to $p = 2$. The distance between two points

[1]It is possible to introduce the idea of closeness abstractly, without resort to the notion of distance, as is done in topology. However, distance, as applied in vector spaces, is as abstract as we want to get.

depends on the particular norm used. For example, consider the "point" (or vector) $|b\rangle = (0.1, 0.1, \ldots, 0.1)$ in a 1000-dimensional space ($n = 1000$). One can easily check that the distance of this vector from the origin varies considerably with p: $\|b\|_1 = 100$, $\|b\|_2 = 3.16$, $\|b\|_{10} = 0.2$. This variation may give the impression that there is no such thing as "closeness", and it all depends on how one defines the norm. This is not true, because closeness is a relative concept: One always *compares* distances. A norm with large p shrinks *all* distances of a space, and a norm with small p stretches them. Thus, although it is impossible (and meaningless) to say that "$|a\rangle$ is close to $|b\rangle$" because of the dependence of distance on p, one can always say "$|a\rangle$ is closer to $|b\rangle$ than $|c\rangle$ is to $|d\rangle$," regardless of the value of p.

Closeness is a relative concept!

Now that we have a way of telling whether vectors are close together or far apart, we can talk about limits and the convergence of sequences of vectors. Let us begin by recalling the definition of a Cauchy sequence

Cauchy sequence defined

5.1.1. Definition. *An infinite sequence of vectors $\{|a_i\rangle\}_{i=1}^{\infty}$ in a normed linear space $\mathcal{V}$ is called a* ***Cauchy sequence*** *if* $\lim_{\substack{i\to\infty \\ j\to\infty}} \|a_i - a_j\| = 0$.

A convergent sequence is necessarily Cauchy. This can be shown using the triangle inequality (see Problem 5.2). However, there may be Cauchy sequences in a given vector space that do not converge to any vector in that space (see the example below). Such a convergence requires additional properties of a vector space summarized in the following definition.

complete vector space defined

5.1.2. Definition. *A* ***complete vector space*** *$\mathcal{V}$ is a normed linear space for which every Cauchy sequence of vectors in $\mathcal{V}$ has a limit vector in $\mathcal{V}$. In other words, if $\{|a_i\rangle\}_{i=1}^{\infty}$ is a Cauchy sequence, then there exists a vector $|a\rangle \in \mathcal{V}$ such that* $\lim_{i\to\infty} \|a_i - a\| = 0$.

5.1.3. Example. 1. $\mathbb{R}$ is complete with respect to the absolute-value norm $\|\alpha\| = |\alpha|$. In other words, every Cauchy sequence of real numbers has a limit in $\mathbb{R}$. This is proved in real analysis.

2. $\mathbb{C}$ is complete with respect to the norm $\|\alpha\| = |\alpha| = \sqrt{(\mathrm{Re}\,\alpha)^2 + (\mathrm{Im}\,\alpha)^2}$. Using $|\alpha| \le |\mathrm{Re}\,\alpha| + |\mathrm{Im}\,\alpha|$, one can show that the completeness of $\mathbb{C}$ follows from that of $\mathbb{R}$. Details are left as an exercise for the reader.

3. The set of rational numbers $\mathbb{Q}$ is *not* complete with respect to the absolute-value norm. In fact, $\{(1 + 1/k)^k\}_{k=1}^{\infty}$ is a sequence of rational numbers that is Cauchy but does not converge to a rational number; it converges to e, the base of the natural logarithm, which is known to be an irrational number. ■

Let $\{|a_i\rangle\}_{i=1}^{\infty}$ be a Cauchy sequence of vectors in a finite-dimensional vector space $\mathcal{V}_N$. Choose an orthonormal basis $\{|e_k\rangle\}_{k=1}^{N}$ in $\mathcal{V}_N$ such that[2] $|a_i\rangle =$

[2]Recall that one can always define an inner product on a finite-dimensional vector space. So, the existence of orthonormal bases is guaranteed.

$\sum_{k=1}^{N} \alpha_k^{(i)} |e_k\rangle$ and $|a_j\rangle = \sum_{k=1}^{N} \alpha_k^{(j)} |e_k\rangle$. Then

$$\|a_i - a_j\|^2 = \langle a_i - a_j | a_i - a_j \rangle = \left\| \sum_{k=1}^{N} (\alpha_k^{(i)} - \alpha_k^{(j)}) |e_k\rangle \right\|^2$$
$$= \sum_{k,l=1}^{N} (\alpha_k^{(i)} - \alpha_k^{(j)})^* (\alpha_l^{(i)} - \alpha_l^{(j)}) \langle e_k | e_l \rangle = \sum_{k=1}^{N} |\alpha_k^{(i)} - \alpha_k^{(j)}|^2.$$

The LHS goes to zero, because the sequence is assumed Cauchy. Furthermore, all terms on the RHS are positive. Thus, they too must go to zero as $i, j \to \infty$. By the completeness of $\mathbb{C}$, there must exist $\alpha_k \in \mathbb{C}$ such that $\lim_{n\to\infty} \alpha_k^{(n)} = \alpha_k$ for $k = 1, 2, \ldots, N$. Now consider $|a\rangle \in \mathcal{V}_N$ given by $|a\rangle = \sum_{k=1}^{N} \alpha_k |e_k\rangle$. We claim that $|a\rangle$ is the limit of the above sequence of vectors in $\mathcal{V}_N$. Indeed,

$$\lim_{i\to\infty} \|a_i - a\|^2 = \lim_{i\to\infty} \sum_{k=1}^{N} |\alpha_k^{(i)} - \alpha_k|^2 = \sum_{k=1}^{N} \lim_{i\to\infty} |\alpha_k^{(i)} - \alpha_k|^2 = 0$$

We have proved the following:

all finite-dimensional vector spaces are complete

5.1.4. Proposition. *Every Cauchy sequence in a finite-dimensional inner product space over $\mathbb{C}$ (or $\mathbb{R}$) is convergent. In other words, every finite-dimensional complex (or real) inner product space is complete with respect to the norm induced by its inner product.*

The next example shows how important the word "finite" is.

5.1.5. Example. Consider $\{f_k\}_{k=1}^{\infty}$, the infinite sequence of *continuous* functions defined in the interval $[-1, +1]$ by

$$f_k(x) = \begin{cases} 1 & \text{if } 1/k \le x \le 1, \\ (kx+1)/2 & \text{if } -1/k \le x \le 1/k, \\ 0 & \text{if } -1 \le x \le -1/k. \end{cases}$$

This sequence belongs to $\mathcal{C}^0(-1, 1)$, the inner product space of continuous functions with its usual inner product: $\langle f | g \rangle = \int_{-1}^{1} f^*(x) g(x)\, dx$. It is straightforward to verify that $\|f_k - f_j\|^2 = \int_{-1}^{1} |f_k(x) - f_j(x)|^2 dx \xrightarrow[k,j\to\infty]{} 0$. Therefore, the sequence is Cauchy. However, the limit of this sequence is (see Figure 5.1)

$$f(x) = \begin{cases} 1 & \text{if } 0 < x < 1, \\ 0 & \text{if } -1 < x < 0, \end{cases}$$

which is discontinuous at $x = 0$ and therefore does not belong to the space in which the original sequence lies. ■

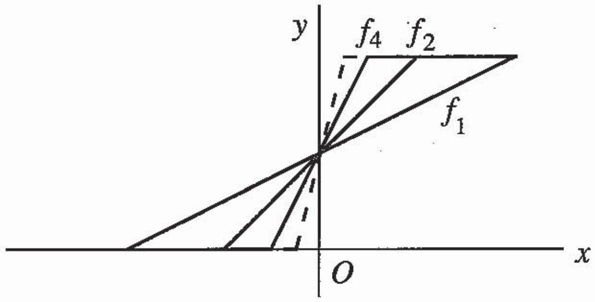


Figure 5.1 The limit of the sequence of the *continuous* functions f_k is a discontinuous function that is 1 for $x > 0$ and 0 for $x < 0$.

We see that infinite-dimensional vector spaces are not generally complete. It is a nontrivial task to show whether or not a given infinite-dimensional vector space is complete.

Any vector space (finite- or infinite-dimensional) contains all finite linear combinations of the form $\sum_{i=1}^{n} \alpha_i |a_i\rangle$ when it contains all the $|a_i\rangle$'s. This follows from the very definition of a vector space. However, the situation is different when n goes to infinity. For the vector space to contain the infinite sum, firstly, the meaning of such a sum has to be clarified, i.e., a norm and an associated convergence criterion needs to be put in place. Secondly, the vector space has to be complete with respect to that norm. A complete normed vector space is called a **Banach space**. We shall not deal with a general Banach space, but only with those spaces whose norms arise naturally from an inner product. This leads to the following definition:

Banach space

Hilbert space defined

5.1.6. Definition. *A complete inner product space, commonly denoted by $\mathcal{H}$, is called a* ***Hilbert space****.*

Thus, all finite-dimensional real or complex vector spaces are Hilbert spaces. However, when we speak of a Hilbert space, we shall usually assume that it is infinite-dimensional.

It is convenient to use orthonormal vectors in studying Hilbert spaces. So, let us consider an infinite sequence $\{|e_i\rangle\}_{i=1}^{\infty}$ of orthonormal vectors all belonging to a Hilbert space $\mathcal{H}$. Next, take any vector $|f\rangle \in \mathcal{H}$, construct the complex numbers $f_i = \langle e_i | f\rangle$, and form the sequence of vectors[3]

$$|f_n\rangle = \sum_{i=1}^{n} f_i |e_i\rangle \qquad \text{for } n = 1, 2, \ldots \tag{5.1}$$

[3] We can consider $|f_n\rangle$ as an "approximation" to $|f\rangle$, because both share the same components along the same set of orthonormal vectors. The sequence of orthonormal vectors acts very much as a basis. However, to be a basis, an extra condition must be met. We shall discuss this condition shortly.

For the pair of vectors $|f\rangle$ and $|f_n\rangle$, the Schwarz inequality gives

$$|\langle f|f_n\rangle|^2 \le \langle f|f\rangle\langle f_n|f_n\rangle = \langle f|f\rangle\left(\sum_{i=1}^{n}|f_i|^2\right), \tag{5.2}$$

where Equation (5.1) has been used to evaluate $\langle f_n|f_n\rangle$. On the other hand, taking the inner product of (5.1) with $\langle f|$ yields

$$\langle f|f_n\rangle = \sum_{i=1}^{n} f_i\langle f|e_i\rangle = \sum_{i=1}^{n} f_i f_i^* = \sum_{i=1}^{n}|f_i|^2.$$

Parseval inequality

Substitution of this in Equation (5.2) yields the **Parseval inequality**:

$$\sum_{i=1}^{n}|f_i|^2 \le \langle f|f\rangle. \tag{5.3}$$

This conclusion is true for arbitrarily large n and can be stated as follows:

Bessel inequality

5.1.7. Proposition. *Let $\{|e_i\rangle\}_{i=1}^{\infty}$ be an infinite set of orthonormal vectors in a Hilbert space, $\mathcal{H}$. Let $|f\rangle \in \mathcal{H}$ and define complex numbers $f_i = \langle e_i|f\rangle$. Then the* ***Bessel inequality*** *holds: $\sum_{i=1}^{\infty}|f_i|^2 \le \langle f|f\rangle$.*

The Bessel inequality shows that the vector

$$\sum_{i=1}^{\infty} f_i|e_i\rangle \equiv \lim_{n\to\infty}\sum_{i=1}^{n} f_i|e_i\rangle$$

converges; that is, it has a finite norm. However, the inequality does not say whether the vector converges to $|f\rangle$. To make such a statement we need completeness:

complete orthonormal sequence of vectors

5.1.8. Definition. *A sequence of orthonormal vectors $\{|e_i\rangle\}_{i=1}^{\infty}$ in a Hilbert space $\mathcal{H}$ is called* ***complete*** *if the only vector in $\mathcal{H}$ that is orthogonal to all the $|e_i\rangle$ is the zero vector.*

This completeness property is the extra condition alluded to (in the footnote) above, and is what is required to make a basis.

5.1.9. Proposition. *Let $\{|e_i\rangle\}_{i=1}^{\infty}$ be an orthonormal sequence in $\mathcal{H}$. Then the following statements are equivalent:*

1. $\{|e_i\rangle\}_{i=1}^{\infty}$ *is complete.*
2. $|f\rangle = \sum_{i=1}^{\infty}|e_i\rangle\langle e_i|f\rangle \quad \forall\ |f\rangle \in \mathcal{H}$.
3. $\sum_{i=1}^{\infty}|e_i\rangle\langle e_i| = \mathbf{1}$.
4. $\langle f|g\rangle = \sum_{i=1}^{\infty}\langle f|e_i\rangle\langle e_i|g\rangle \quad \forall\ |f\rangle, |g\rangle \in \mathcal{H}$.

5. $\|f\|^2 = \sum_{i=1}^{\infty} |\langle e_i| f\rangle|^2 \quad \forall\ |f\rangle \in \mathcal{H}$.

Proof. We shall prove the implications $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 5 \Rightarrow 1$.
$1 \Rightarrow 2$: It is sufficient to show that the vector $|\psi\rangle \equiv |f\rangle - \sum_{i=1}^{\infty} |e_i\rangle \langle e_i| f\rangle$ is orthogonal to all the $|e_j\rangle$:

$$\langle e_j| \psi\rangle = \langle e_j| f\rangle - \sum_{i=1}^{\infty} \overbrace{\langle e_j| e_i\rangle}^{\delta_{ij}} \langle e_i| f\rangle = 0.$$

$2 \Rightarrow 3$: Since $|f\rangle = \mathbf{1}\,|f\rangle = \sum_{i=1}^{\infty}(|e_i\rangle\langle e_i|)\,|f\rangle$ is true for all $|f\rangle \in \mathcal{H}$, we must have $\mathbf{1} = \sum_{i=1}^{\infty} |e_i\rangle \langle e_i|$.
$3 \Rightarrow 4$: $\langle f| g\rangle = \langle f|\mathbf{1}|g\rangle = \langle f|\left(\sum_{i=1}^{\infty} |e_i\rangle\langle e_i|\right)|g\rangle = \sum_{i=1}^{\infty} \langle f| e_i\rangle \langle e_i| g\rangle$.
$4 \Rightarrow 5$: Let $|g\rangle = |f\rangle$ in statement 4 and recall that $\langle f| e_i\rangle = \langle e_i| f\rangle^*$.
$5 \Rightarrow 1$: Let $|f\rangle$ be orthogonal to all the $|e_i\rangle$. Then all the terms in the sum are zero implying that $\|f\|^2 = 0$, which in turn gives $|f\rangle = 0$, because only the zero vector has a zero norm. □

Parseval equality; generalized Fourier coefficients

The equality

$$\|f\|^2 = \langle f| f\rangle = \sum_{i=1}^{\infty} |\langle e_i| f\rangle|^2 = \sum_{i=1}^{\infty} |f_i|^2, \qquad f_i = \langle e_i| f\rangle, \tag{5.4}$$

is called the **Parseval equality**, and the complex numbers f_i are called **generalized Fourier coefficients**. The relation

completeness relation

$$\mathbf{1} = \sum_{i=1}^{\infty} |e_i\rangle \langle e_i| \tag{5.5}$$

is called the **completeness relation.**

basis for Hilbert spaces

5.1.10. Definition. *A complete orthonormal sequence $\{|e_i\rangle\}_{i=1}^{\infty}$ in a Hilbert space $\mathcal{H}$ is called a **basis** of $\mathcal{H}$.*

5.2 The Space of Square-Integrable Functions

Chapter 1 showed that the collection of all continuous functions defined on an interval $[a, b]$ forms a linear vector space. Example 5.1.5 showed that this space is not complete. Can we enlarge this space to make it complete? Since we are interested in an inner product as well, and since a natural inner product for functions is defined in terms of integrals, we want to make sure that our functions are integrable. However, integrability does not require continuity, it only requires *piecewise* continuity. In this section we shall discuss conditions under which the space of functions becomes complete. An important class of functions has already been mentioned in Chapter 1. These functions satisfy the inner product given by

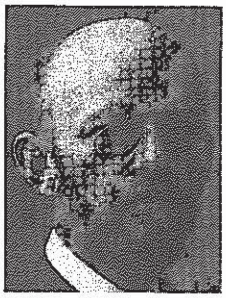
$$\langle g|f\rangle = \int_a^b g^*(x)f(x)w(x)\,dx.$$

square-integrable functions

If $g(x) = f(x)$, we obtain

$$\langle f|f\rangle = \int_a^b |f(x)|^2 w(x)\,dx. \tag{5.6}$$

Functions for which such an integral is defined are said to be **square-integrable.**



David Hilbert (1862–1943), the greatest mathematician of this century, received his Ph.D. from the University of Königsberg and was a member of the staff there from 1886 to 1895. In 1895 he was appointed to the chair of mathematics at the University of Göttingen, where he continued to teach for the rest of his life.

Hilbert is one of that rare breed of late 19th-century mathematicians whose spectrum of expertise covered a wide range, with formal set theory at one end and mathematical physics at the other. He did superb work in geometry, algebraic geometry, algebraic number theory, integral equations, and operator theory. The seminal two-volume book *Methoden der mathematische Physik* by R. Courant, still one of the best books on the subject, was greatly influenced by Hilbert.

Hilbert's work in geometry had the greatest influence in that area since Euclid. A systematic study of the axioms of Euclidean geometry led Hilbert to propose 21 such axioms, and he analyzed their significance. He published *Grundlagen der Geometrie* in 1899, putting geometry on a formal axiomatic foundation. His famous 23 Paris problems challenged (and still today challenge) mathematicians to solve fundamental questions.

It was late in his career that Hilbert turned to the subject for which he is most famous among physicists. A lecture by Erik Holmgren in 1901 on Fredholm's work on integral equations, which had already been published in Sweden, aroused Hilbert's interest in the subject. David Hilbert, having established himself as the leading mathematician of his time by his work on algebraic numbers, algebraic invariants, and the foundations of geometry, now turned his attention to **integral equations.** He says that an investigation of the subject showed him that it was important for the theory of definite integrals, for the development of arbitrary functions in series (of special functions or trigonometric functions), for the theory of linear differential equations, for potential theory, and for the calculus of variations. He wrote a series of six papers from 1904 to 1910 and reproduced them in his book *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen* (1912). During the latter part of this work he applied integral equations to problems of mathematical physics.

It is said that Hilbert discovered the correct field equation for general relativity in 1915 (one year before Einstein) using the variational principle, but never claimed priority.

Hilbert claimed that he worked best out-of-doors. He accordingly attached an 18-foot blackboard to his neighbor's wall and built a covered walkway there so that he could work outside in any weather. He would intermittently interrupt his pacing and his blackboard computations with a few turns around the rest of the yard on his bicycle, or he would pull some weeds, or do some garden trimming. Once, when a visitor called, the maid sent him to the backyard and advised that if the master wasn't readily visible at the blackboard to look for him up in one of the trees.

Highly gifted and highly versatile, David Hilbert radiated over mathematics a catching optimism and a stimulating vitality that can only be called "the spirit of Hilbert." Engraved on a stone marker set over Hilbert's grave in Göttingen are the master's own optimistic words: "Wir müssen wissen. Wir werden wissen." ("We must know. We shall know.")

The space of square-integrable functions over the interval $[a, b]$ is denoted by $\mathcal{L}_w^2(a, b)$. In this notation $\mathcal{L}$ stands for *Lebesgue*, who generalized the notion of the ordinary Riemann integral to cases for which the integrand could be highly discontinuous; 2 stands for the power of $f(x)$ in the integral; a and b denote the limits of integration; and w refers to the weight function (a strictly positive real-valued function). When $w(x) = 1$, we use the notation $\mathcal{L}^2(a, b)$. The significance of $\mathcal{L}_w^2(a, b)$ lies in the following theorem (for a proof, see [Reed 80, Chapter III]):

$\mathcal{L}_w^2(a, b)$ is complete

5.2.1. Theorem. (Riesz-Fischer theorem) *The space $\mathcal{L}_w^2(a, b)$ is complete.*

A complete infinite-dimensional inner product space was earlier defined to be a Hilbert space. The following theorem shows that the number of Hilbert spaces is severely restricted. (For a proof, see [Frie 82, p. 216].)

all Hilbert spaces are alike

5.2.2. Theorem. *All infinite-dimensional complete inner product spaces are isomorphic to $\mathcal{L}_w^2(a, b)$.*

$\mathcal{L}_w^2(a, b)$ is defined in terms of functions that satisfy Equation (5.6). Yet an inner product involves integrals of the form $\int_a^b g^*(x)f(x)w(x)\,dx$. Are such integrals well-defined and finite? Using the Schwarz inequality, which holds for any inner product space, finite or infinite, one can show that the integral is defined. The isomorphism of Theorem 5.2.2 makes the Hilbert space more tangible, because it identifies the space with a space of functions, objects that are more familiar than abstract vectors. Nonetheless, a faceless function is very little improvement over an abstract vector. What is desirable is a set of concrete functions with which we can calculate. The following theorem provides such functions (for a proof, see [Simm 83, pp. 154–161]).

5.2.3. Theorem. (Stone-Weierstrass approximation theorem) *The sequence of functions (monomials) $\{x^k\}$, where $k = 0, 1, 2, \ldots$, forms a basis of $\mathcal{L}_w^2(a, b)$.*

Thus, any function f can be written as $f(x) = \sum_{k=0}^{\infty} \alpha_k x^k$. Note that the $\{x^k\}$ are not orthonormal but are linearly independent. If we wish to obtain an orthonormal—or simply orthogonal—linear combination of these vectors, we can use the Gram-Schmidt process. The result will be certain polynomials, denoted by $C_n(x)$, that are orthogonal to one another and span $\mathcal{L}_w^2(a, b)$.

Such orthogonal polynomials satisfy very useful **recurrence relations**, which we now derive. In the following discussion $p_{\le k}(x)$ denotes a generic polynomial of degree less than or equal to k. For example, $3x^5 - 4x^2 + 5$, $2x + 1$, $-2.4x^4 + 3x^3 - x^2 + 6$, and 2 are all denoted by $p_{\le 5}(x)$ or $p_{\le 8}(x)$ or $p_{\le 59}(x)$ because they all have degrees less than or equal to 5, 8, and 59. Since a polynomial of degree less than n can be written as a linear combination of $C_k(x)$ with $k < n$, we have the obvious property

$$\int_a^b C_n(x)p_{\le n-1}(x)w(x)\,dx = 0. \tag{5.7}$$

Let k_m and k_m' denote, respectively, the coefficients of x^m and x^{m-1} in $C_m(x)$, and let

$$h_m = \int_a^b [C_m(x)]^2 w(x)\,dx. \tag{5.8}$$

The polynomial $C_{n+1}(x) - (k_{n+1}/k_n)xC_n(x)$ has degree less than or equal to n, and therefore can be expanded as a linear combination of the $C_j(x)$:

$$C_{n+1}(x) - \frac{k_{n+1}}{k_n}xC_n(x) = \sum_{j=0}^{n} a_j C_j(x). \tag{5.9}$$

Take the inner product of both sides of this equation with $C_m(x)$:

$$\int_a^b C_{n+1}(x)C_m(x)w(x)\,dx - \frac{k_{n+1}}{k_n}\int_a^b xC_n(x)C_m(x)w(x)\,dx$$
$$= \sum_{j=0}^{n} a_j \int_a^b C_j(x)C_m(x)w(x)\,dx.$$

The first integral on the LHS vanishes as long as $m \le n$; the second integral vanishes if $m \le n - 2$ [if $m \le n - 2$, then $xC_m(x)$ is a polynomial of degree $n - 1$]. Thus, we have

$$\sum_{j=0}^{n} a_j \int_a^b C_j(x)C_m(x)w(x)\,dx = 0 \qquad \text{for } m \le n-2.$$

The integral in the sum is zero unless $j = m$, by orthogonality. Therefore, the sum reduces to

$$a_m \int_a^b [C_m(x)]^2 w(x)\,dx = 0 \qquad \text{for } m \le n-2.$$

Since the integral is nonzero, we conclude that $a_m = 0$ for $m = 0, 1, 2, \ldots, n-2$, and Equation (5.9) reduces to

$$C_{n+1}(x) - \frac{k_{n+1}}{k_n}xC_n(x) = a_{n-1}C_{n-1}(x) + a_nC_n(x). \tag{5.10}$$

It can be shown that if we define

$$\alpha_n = \frac{k_{n+1}}{k_n}, \quad \beta_n = \alpha_n \left(\frac{k'_{n+1}}{k_{n+1}} - \frac{k'_n}{k_n} \right), \quad \gamma_n = -\frac{h_n}{h_{n-1}} \frac{\alpha_n}{\alpha_{n-1}}, \tag{5.11}$$

a recurrence relation for orthogonal polynomials

then Equation (5.10) can be expressed as

$$C_{n+1}(x) = (\alpha_n x + \beta_n) C_n(x) + \gamma_n C_{n-1}(x), \tag{5.12}$$

or

$$x C_n(x) = \frac{1}{\alpha_n} C_{n+1}(x) - \frac{\beta_n}{\alpha_n} C_n(x) - \frac{\gamma_n}{\alpha_n} C_{n-1}(x). \tag{5.13}$$

Other recurrence relations, involving higher powers of x, can be obtained from the one above. For example, a recurrence relation involving x^2 can be obtained by multiplying both sides of Equation (5.13) by x and expanding each term of the RHS using that same equation. The result will be

$$\begin{aligned} x^2 C_n(x) &= \frac{1}{\alpha_n \alpha_{n+1}} C_{n+2}(x) - \left(\frac{\beta_{n+1}}{\alpha_n \alpha_{n+1}} + \frac{\beta_n}{\alpha_n^2} \right) C_{n+1}(x) \\ &\quad - \left(\frac{\gamma_{n+1}}{\alpha_n \alpha_{n+1}} - \frac{\beta_n^2}{\alpha_n^2} + \frac{\gamma_n}{\alpha_n \alpha_{n-1}} \right) C_n(x) \\ &\quad + \left(\frac{\beta_n \gamma_n}{\alpha_n^2} + \frac{\beta_{n-1} \gamma_n}{\alpha_n \alpha_{n-1}} \right) C_{n-1}(x) + \frac{\gamma_{n-1} \gamma_n}{\alpha_n \alpha_{n-1}} C_{n-2}(x). \end{aligned} \tag{5.14}$$

5.2.4. Example. As an application of the recurrence relations above, let us evaluate

$$I_1 \equiv \int_a^b x C_m(x) C_n(x) w(x)\, dx.$$

Substituting (5.13) in the integral gives

$$\begin{aligned} I_1 &= \frac{1}{\alpha_n} \int_a^b C_m(x) C_{n+1}(x) w(x)\, dx - \frac{\beta_n}{\alpha_n} \int_a^b C_m(x) C_n(x) w(x)\, dx \\ &\quad - \frac{\gamma_n}{\alpha_n} \int_a^b C_m(x) C_{n-1}(x) w(x)\, dx. \end{aligned}$$

We now use the orthogonality relations among the $C_k(x)$ to obtain

$$\begin{aligned} I_1 &= \frac{1}{\alpha_n} \delta_{m,n+1} \overbrace{\int_a^b C_m^2(x) w(x)\, dx}^{=h_m} - \frac{\beta_n}{\alpha_n} \delta_{mn} \int_a^b C_m^2(x) w(x)\, dx \\ &\quad - \frac{\gamma_n}{\alpha_n} \delta_{m,n-1} \int_a^b C_m^2(x) w(x)\, dx \\ &= \left(\frac{1}{\alpha_{m-1}} \delta_{m,n+1} - \frac{\beta_m}{\alpha_m} \delta_{mn} - \frac{\gamma_{m+1}}{\alpha_{m+1}} \delta_{m,n-1} \right) h_m, \end{aligned}$$

or

$$I_1 = \begin{cases} h_m/\alpha_{m-1} & \text{if } m = n+1, \\ -\beta_m h_m/\alpha_m & \text{if } m = n, \\ -\gamma_{m+1} h_m/\alpha_{m+1} & \text{if } m = n-1, \\ 0 & \text{otherwise.} \end{cases}$$ ■

5.2.5. Example. Let us find the orthogonal polynomials forming a basis of $\mathcal{L}^2(-1, +1)$, which we denote by $P_k(x)$, where k is the degree of the polynomial. Let $P_0(x) = 1$. To find $P_1(x)$, write $P_1(x) = ax + b$, and determine a and b in such a way that $P_1(x)$ is orthogonal to $P_0(x)$:

$$0 = \int_{-1}^{1} P_1(x)P_0(x)\,dx = \int_{-1}^{1} (ax+b)\,dx = \tfrac{1}{2}ax^2|_{-1}^{1} + 2b = 2b.$$

So one of the coefficients, b, is zero. To find the other one, we need some standardization procedure. We "standardize" $P_k(x)$ by requiring that $P_k(1) = 1\ \forall k$. For $k = 1$ this yields $a \times 1 = 1$, or $a = 1$, so that $P_1(x) = x$.

We can calculate $P_2(x)$ similarly: Write $P_2(x) = ax^2 + bx + c$, impose the condition that it be orthogonal to both $P_1(x)$ and $P_0(x)$, and enforce the standardization procedure. All this will yield

$$0 = \int_{-1}^{1} P_2(x)P_0(x)\,dx = \frac{2}{3}a + 2c, \quad 0 = \int_{-1}^{1} P_2(x)P_1(x)\,dx = \frac{2}{3}b,$$

and $P_2(1) = a+b+c = 1$. These three equations have the unique solution $a = 3/2, b = 0$, $c = -1/2$. Thus, $P_2(x) = \frac{1}{2}(3x^2 - 1)$. These are the first three Legendre polynomials, which are part of a larger group of polynomials to be discussed in Chapter 7. ■

5.2.1 Orthogonal Polynomials and Least Squares

The method of least squares is no doubt familiar to the reader. In the simplest procedure, one tries to find a linear function that most closely fits a set of data. By definition, "most closely" means that the sum of the squares of the differences between the data points and the corresponding values of the linear function is minimum. More generally, one seeks the best polynomial fit to the data.

We shall consider a related topic, namely least-square fitting of a given *function* with polynomials. Suppose $f(x)$ is a function defined on (a, b). We want to find a polynomial that most closely approximates f. Write such a polynomial as $p(x) = \sum_{k=0}^{n} a_k x^k$, where the a_k's are to be determined such that

$$S(a_0, a_1, \ldots, a_n) \equiv \int_a^b [f(x) - a_0 - a_1 x - \cdots - a_n x^n]^2\,dx$$

is a minimum. Differentiating S with respect to the a_k's and setting the result equal to zero gives

$$0 = \frac{\partial S}{\partial a_j} = \int_a^b 2(-x^j)\left[f(x) - \sum_{k=0}^{n} a_k x^k\right] dx,$$

(a) $\|f \pm g\| = \|f\| + \|g\|$.
(b) $\|f + g\|^2 + \|f - g\|^2 = 2(\|f\| + \|g\|)^2$.
(c) Using parts (a), (b), and Theorem 1.2.8, show that $\mathcal{L}^1(\mathbb{R})$ is not an inner product space. This shows that not all norms arise from an inner product.

5.6. Use Equation (5.10) to derive Equation (5.12). Hint: To find a_n, equate the coefficients of x^n on both sides of Equation (5.10). To find a_{n-1}, multiply both sides of Equation (5.10) by $C_{n-1}w(x)$ and integrate, using the definitions of k_n, k_n', and h_n.

5.7. Evaluate the integral $\int_a^b x^2 C_m(x) C_n(x) w(x)\, dx$.

Additional Reading

1. Boccara, N. *Functional Analysis*, Academic Press, 1990. An application oriented book with many abstract topics related to Hilbert spaces (e.g., Lebesgue measure) explained for a physics audience.

2. DeVito, C. *Functional Analysis and Linear Operator Theory*, Addison-Wesley, 1990.

3. Reed, M., and Simon, B. *Functional Analysis*, Academic Press, 1980. Coauthored by a mathematical physicist (B.S.), this first volume of a four-volume encyclopedic treatise on functional analysis and Hilbert spaces has many examples and problems to help the reader comprehend the rather abstract presentation.

4. Zeidler, E. *Applied Functional Analysis*, Springer-Verlag, 1995. Another application-oriented book on Hilbert spaces suitable for a physics audience.