1

Linear Analysis

1.1 INTRODUCTION

Fundamental to the study of many of the differential equations describing physical processes in applied physics and engineering is *linear analysis.* Linear analysis can be elegantly and logically placed in a mathematical structure called a *linear space.*

We begin this chapter with the definition of a linear space. We then begin to add structure to the linear space by introducing the concepts of inner product and norm. Our study leads us to Hilbert space and, finally, to linear operators within Hilbert space. The characteristics of these operators are basic to the ensuing development of the differential operators and differential equations found in electromagnetic theory.

Throughout this chapter, we shall be developing notions concerning vectors in a linear space. These ideas make use of both the real and complex number systems. A knowledge of the axioms and theorems governing real and complex numbers will be assumed in what follows. We shall use this information freely in the proofs involving vectors.

1.2 LINEAR SPACE

Let $a, b, c, \ldots$ be elements of a set $\mathcal{S}$. These elements are called *vectors.* Let $\alpha, \beta, \ldots$ be elements of the field of numbers $\mathbf{F}$. In particular, let $\mathbf{R}$ and

C be the field of real and complex numbers, respectively. The set $\mathcal{S}$ is a *linear space* if the following rules for addition and multiplication apply:

I. Rules for addition among vectors in $\mathcal{S}$:

  a. $(a+b)+c = a+(b+c)$
  b. There exists a zero vector $\mathbf{0}$ such that $a+\mathbf{0} = \mathbf{0}+a = a$.
  c. For every $a \in \mathcal{S}$, there exists $-a \in \mathcal{S}$ such that $a+(-a) = (-a)+a = \mathbf{0}$.
  d. $a+b = b+a$

II. Rules for multiplication of vectors in $\mathcal{S}$ by elements of $\mathbf{F}$:

  a. $\alpha(\beta a) = (\alpha\beta)a$
  b. $1a = a$
  c. $\alpha(a+b) = \alpha a + \alpha b$
  d. $(\alpha+\beta)a = \alpha a + \beta a$

EXAMPLE 1.1 Consider *Euclidean space* $\mathbf{R}_n$. Define vectors a and b in $\mathbf{R}_n$ as follows:

$$a = (\alpha_1, \alpha_2, \ldots, \alpha_n) \tag{1.1}$$

$$b = (\beta_1, \beta_2, \ldots, \beta_n) \tag{1.2}$$

where α_k and β_k, the *components* of vectors a and b, are in $\mathbf{R}$, $k = 1, 2, \ldots, n$. Define addition and multiplication as follows:

$$\begin{aligned} a+b &= (\alpha_1, \ldots, \alpha_n) + (\beta_1, \ldots, \beta_n) \\ &= (\alpha_1+\beta_1, \ldots, \alpha_n+\beta_n) \end{aligned} \tag{1.3}$$

$$\begin{aligned} \alpha a &= \alpha(\alpha_1, \ldots, \alpha_n) \\ &= (\alpha\alpha_1, \ldots, \alpha\alpha_n) \end{aligned} \tag{1.4}$$

where $\alpha \in \mathbf{R}$. If we assume prior establishment of rules for addition and multiplication in the field of real numbers, it is easy to show that $\mathbf{R}_n$ is a linear space. We must show that the rules in I and II are satisfied. For example, for addition rule d,

$$\begin{aligned} a+b &= (\alpha_1+\beta_1, \ldots, \alpha_n+\beta_n) \\ &= (\beta_1+\alpha_1, \ldots, \beta_n+\alpha_n) \\ &= b+a \end{aligned} \tag{1.5}$$

We leave the satisfaction of the remainder of the rules in this example for Problem 1.2. Note that $\mathbf{R}$ is also a linear space, where we make the identification $\mathbf{R} = \mathbf{R}_1$.

■

EXAMPLE 1.2 Consider *unitary space* $\mathbf{C}_n$. Vectors in the space are given by (1.1) and (1.2), where α_k and β_k, $k = 1, 2, \ldots, n$ are in $\mathbf{C}$. Addition and multiplication are defined by (1.3) and (1.4) where $\alpha \in \mathbf{C}$. Proof that $\mathbf{C}_n$ is a linear space follows the same lines as in Example 1.1. Note that $\mathbf{C}$ is a linear space, where we make the identification $\mathbf{C} = \mathbf{C}_1$.

■

EXAMPLE 1.3 Consider $\mathcal{C}(0,1)$, the space of real-valued functions continuous on the interval (0, 1). For f and g in $\mathcal{C}(0,1)$ and $\alpha \in \mathbf{R}$, we define addition and multiplication as follows:

$$(f + g)(\xi) = f(\xi) + g(\xi) \tag{1.6}$$

$$(\alpha f)(\xi) = \alpha f(\xi) \tag{1.7}$$

for all $\xi \in (0, 1)$. If we assume prior establishment of the rules for addition of two real-valued functions and multiplication of a real-valued function by a real scalar, it is easy to establish that $\mathcal{C}(0,1)$ is a linear space by showing that the rules in I and II are satisfied. For example, for addition rule d,

$$\begin{aligned}(f + g)(\xi) &= f(\xi) + g(\xi)\\ &= g(\xi) + f(\xi)\\ &= (g + f)(\xi)\end{aligned} \tag{1.8}$$

We leave the completion of the proof for Problem 1.3.

■

In ordinary vector analysis over two or three spatial coordinates, we are often concerned with vectors that are parallel (collinear). This concept can be generalized in an abstract linear space. Let $x_1, x_2, \ldots, x_n$ be elements of a set of vectors in $\mathcal{S}$. The vectors are *linearly dependent* if there exist $\alpha_k \in \mathbf{F}$, $k = 1, 2, \ldots, n$, not all zero, such that

$$\sum_{k=1}^{n} \alpha_k x_k = \mathbf{0} \tag{1.9}$$

If the only way to satisfy (1.9) is $\alpha_k = 0$, $k = 1, 2, \ldots, n$, then the elements x_k are *linearly independent*. The sum

$$\sum_{k=1}^{n} \alpha_k x_k$$

is called a *linear combination* of the vectors x_k.

EXAMPLE 1.4 In $\mathbf{R}_2$, let $x_1 = (1,3)$, $x_2 = (2,6)$. We test x_1 and x_2 for linear dependence. We form

$$\mathbf{0} = (0,0) = \alpha_1(1,3) + \alpha_2(2,6) = (\alpha_1 + 2\alpha_2, 3\alpha_1 + 6\alpha_2)$$

from which we conclude that

$$\alpha_1 + 2\alpha_2 = 0$$
$$3\alpha_1 + 6\alpha_2 = 0$$

These two equations are consistent and yield $\alpha_1 = -2\alpha_2$. Certainly, $\alpha_1 = \alpha_2 = 0$ satisfies this equation, but there is also an infinite number of nonzero possibilities. The vectors are therefore linearly dependent. Indeed, the reader can easily make a sketch to show that x_1 and x_2 are collinear.

■

EXAMPLE 1.5 In $\mathcal{C}(0,1)$, let a set of vectors be defined by $f_k(\xi) = \sqrt{2}\sin k\pi\xi$, $k = 1, 2, \ldots, n$. We test the vectors f_k for linear dependence. We form

$$\sum_{k=1}^{n} \alpha_k \sqrt{2}\sin k\pi\xi = 0 \tag{1.10}$$

where $\alpha_k \in \mathbf{R}$. The f_k, defined above, form an *orthonormal* set on $\xi \in (0,1)$. That is,

$$\int_0^1 f_m f_k d\xi = \begin{cases} 0, & k \neq m \\ 1, & k = m \end{cases} \tag{1.11}$$

Multiplication of both sides of (1.10) by $\sqrt{2}\sin m\pi\xi$, $m = 1, 2, \ldots, n$ and integration over (0,1) give, with the help of (1.11), $\alpha_m = 0$, $m = 1, 2, \ldots, n$. The vectors f_k are therefore linearly independent.

■

In Example 1.5, we note that the elements f_k are finite in number. We recognize them as a finite subset of the countably infinite number of elements in the Fourier sine series f_k, $k = 1, 2, \ldots$. A question arises concerning the linear independence of sets containing a countably infinite number of vectors. Let $x_1, x_2 \ldots$ be an infinite set of vectors in $\mathcal{S}$. The vectors are linearly independent if every finite subset of the vectors is linearly independent. In Example 1.5, this requirement is realized, so that the infinite set of elements present in the Fourier sine series is linearly independent.

In an abstract linear space $\mathcal{S}$, it would be helpful to have a measure of how many and what sort of vectors describe the space. A linear space $\mathcal{S}$ has *dimension n* if it possesses a set of n independent vectors and if every set of $n + 1$ vectors is dependent. If for every positive integer k we can find k independent vectors in $\mathcal{S}$, then $\mathcal{S}$ has infinite dimension. The set $x_1, x_2, \ldots, x_n$ is a *basis* for $\mathcal{S}$ provided that the vectors in the set are linearly independent, and provided that every $x \in \mathcal{S}$ can be written as a linear combination of the x_k, viz.

$$x = \sum_{k=1}^{n} \alpha_k x_k \tag{1.12}$$

The representation with respect to a given basis is unique. If it were not, then, in addition to the representation in (1.12), there would exist $\beta_k \in \mathbf{F}$, $k = 1, 2, \ldots, n$ such that

$$x = \sum_{k=1}^{n} \beta_k x_k \tag{1.13}$$

Subtraction of (1.13) from (1.12) yields

$$\mathbf{0} = \sum_{k=1}^{n} (\alpha_k - \beta_k) x_k \tag{1.14}$$

Since the x_k are linearly independent, we must have

$$\alpha_k - \beta_k = 0, \qquad k = 1, 2, \ldots, n \tag{1.15}$$

which proves uniqueness with respect to a given basis. Finally, if $\mathcal{S}$ is n-dimensional, any set of n linearly independent vectors $x_1, x_2, \ldots, x_n$ forms a basis. Indeed, let $x \in \mathcal{S}$. By the definition of dimension, the set $x, x_1, x_2, \ldots, x_n$ is linearly dependent, and therefore,

$$\alpha x + \sum_{k=1}^{n} \alpha_k x_k = \mathbf{0} \tag{1.16}$$

where we must have $\alpha \neq 0$. Dividing by α gives

$$x = \sum_{k=1}^{n} \left(\frac{-\alpha_k}{\alpha} \right) x_k \tag{1.17}$$

Therefore, the set $x_1, x_2, \ldots, x_n$ is a basis.

EXAMPLE 1.6 Consider Euclidean space $\mathbf{R}_n$. We shall show that the vectors $e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 1)$ satisfy the two requirements for a basis. First, the set $e_1, \ldots, e_n$ is independent (Problem 1.8). Second, if $a \in \mathbf{R}_n$,

$$\begin{aligned} a &= (\alpha_1, \alpha_2, \ldots, \alpha_n) \\ &= \alpha_1(1, 0, \ldots, 0) + \alpha_2(0, 1, \ldots, 0) + \cdots + \alpha_n(0, 0, \ldots, 1) \\ &= \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n \end{aligned} \tag{1.18}$$

Therefore, any vector in the space can be expressed as a linear combination of the e_k. A special case of this result is obtained by considering Euclidean space $\mathbf{R}_3$. The vectors $e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)$ are a basis. These vectors are perhaps best known as the unit vectors associated with the Cartesian coordinate system.

■

EXAMPLE 1.7 It would be consistent with notation if the dimension of $\mathbf{R}_n$ were, in fact, n. We now show that both requirements for dimension n are satisfied. First, since we have established an n-term basis for $\mathbf{R}_n$ in Example 1.6, the space has a set of n independent vectors. Second, we must show that *any* set of $n + 1$ vectors is dependent. Let $a_1, a_2, \ldots, a_n, a_{n+1}$ be an arbitrary set of $n + 1$ vectors in $\mathbf{R}_n$. We form the expression

$$\sum_{m=1}^{n+1} \gamma_m a_m = \mathbf{0} \tag{1.19}$$

where we must show that there exist $\gamma_m \in \mathbf{R}$, $m = 1, 2, \ldots, n + 1$, not all zero, such that (1.19) is satisfied. We express each of the members of the arbitrary set as a linear combination of the basis vectors, viz.

$$a_m = \sum_{k=1}^{n} \alpha_k^{(m)} e_k, \qquad m = 1, 2, \ldots, n + 1 \tag{1.20}$$

Substitution of (1.20) into (1.19) and interchanging the order of the summations gives

$$\sum_{k=1}^{n} \left(\sum_{m=1}^{n+1} \gamma_m \alpha_k^{(m)} \right) e_k = \mathbf{0} \tag{1.21}$$

Since the e_k are linearly independent,

$$\sum_{m=1}^{n+1} \gamma_m \alpha_k^{(m)} = \mathbf{0}, \qquad k = 1, 2, \ldots, n \tag{1.22}$$

Expression (1.22) is a homogeneous set of n linear equations in $n+1$ unknowns. The set is *underdetermined*, and as a result, always has a nontrivial solution [1]. There is therefore at least one nonzero coefficient among γ_m, $m = 1, 2, \ldots, n+1$. The result is that the arbitrary set $a_1, \ldots, a_n, a_{n+1}$ is linearly dependent and the dimension of $\mathbf{R}_n$ is n.

■

1.3 INNER PRODUCT SPACE

A linear space $\mathcal{S}$ is a *complex inner product space* if for every ordered pair (x, y) of vectors in $\mathcal{S}$, there exists a unique scalar in $\mathbf{C}$, symbolized $\langle x, y\rangle$, such that:

a. $\langle x, y\rangle = \overline{\langle y, x\rangle}$
b. $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$
c. $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle, \ \alpha \in \mathbf{C}$
d. $\langle x, x\rangle \geq 0$, with equality if and only if $x = \mathbf{0}$

In a, the overbar indicates complex conjugate. Similar to the above is the *real inner product space*, which we produce by eliminating the overbar in a and requiring in c that α be in $\mathbf{R}$. For the remainder of this section, we shall assume the complex case. We leave the reader to make the necessary specialization to the real inner product.

EXAMPLE 1.8 We show from the definition of complex inner product space that

$$\langle \mathbf{0}, y\rangle = 0 \tag{1.23}$$

Indeed, the result follows immediately if we substitute $\alpha = 0$ in rule c above.

■

EXAMPLE 1.9 Given the rules for the complex inner product in a–d, the following result holds:

$$\langle x, \alpha y\rangle = \bar{\alpha}\langle x, y\rangle \tag{1.24}$$

Indeed,

$$\begin{aligned}\langle x, \alpha y\rangle &= \overline{\langle \alpha y, x\rangle}\\ &= \overline{\alpha\langle y, x\rangle}\\ &= \bar{\alpha}\ \overline{\langle y, x\rangle}\\ &= \bar{\alpha}\langle x, y\rangle\end{aligned}$$

■

EXAMPLE 1.10 Given the rules for the inner product space, we may show that

$$\langle \sum_{k=1}^{n} \alpha_k x_k, y \rangle = \sum_{k=1}^{n} \alpha_k \langle x_k, y \rangle \tag{1.25}$$

The proof is left for Problem 1.9.

■

EXAMPLE 1.11 In the space $\mathbf{C}_n$, with a and b defined in Example 1.2, define an inner product by

$$\langle a, b \rangle = \sum_{k=1}^{n} \alpha_k \bar{\beta}_k \tag{1.26}$$

Then, $\mathbf{C}_n$ is a complex inner product space. To prove this, we must show that rules a–d for the complex inner product space are satisfied. For rule d, there are three parts to prove. First, we show that the inner product $\langle a, a \rangle$ is nonnegative. Indeed,

$$\langle a, a \rangle = \sum_{k=1}^{n} |\alpha_k|^2 \geq 0$$

Second, we show that $\langle a, a \rangle = 0$ implies that $a = \mathbf{0}$. We have

$$0 = \langle a, a \rangle = \sum_{k=1}^{n} |\alpha_k|^2$$

Since all the terms in the sum are nonnegative, $\alpha_k = 0$, $k = 1, 2, \ldots, n$, and therefore $a = \mathbf{0}$. Third, we must show that $a = \mathbf{0}$ implies $\langle a, a \rangle = 0$. We leave this for the reader. We also leave the reader to demonstrate that rules a–c for the inner product space are satisfied.

■

EXAMPLE 1.12 Let f and g be two vectors in $C(\alpha, \beta)$. Define an inner product by

$$\langle f, g \rangle = \int_{\alpha}^{\beta} f(\xi) g(\xi) d\xi \tag{1.27}$$

Then, $C(\alpha, \beta)$ is a real inner product space. We leave the proof for Problem 1.10.

■

One of the most important inequalities in linear analysis follows from the basic rules for the complex inner product space. The *Cauchy–Schwarz–Bunjakowsky inequality* is given by

$$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle} \tag{1.28}$$

Herein, we refer to (1.28) as the *CSB* inequality. To prove the CSB inequality, we first note that for $|\langle x, y\rangle| = 0$, there is nothing to prove. We may therefore assume $y \neq 0$, with the result $\langle y, y\rangle \neq 0$, and define

$$\alpha = \frac{\langle x, y\rangle}{\langle y, y\rangle}$$

from which we have the result

$$\begin{aligned}\frac{|\langle x, y\rangle|^2}{\langle y, y\rangle} &= \frac{\langle x, y\rangle\langle y, x\rangle}{\langle y, y\rangle} \\ &= \alpha\langle y, x\rangle \\ &= \bar{\alpha}\langle x, y\rangle \\ &= |\alpha|^2\langle y, y\rangle \end{aligned} \tag{1.29}$$

With the help of rule d, we form

$$\begin{aligned}0 \leq \langle x - \alpha y, x - \alpha y\rangle &= \langle x, x\rangle + |\alpha|^2\langle y, y\rangle - \bar{\alpha}\langle x, y\rangle - \alpha\langle y, x\rangle \\ &= \langle x, x\rangle - \frac{|\langle x, y\rangle|^2}{\langle y, y\rangle}\end{aligned}$$

from which the result in (1.28) follows.

Two concepts used throughout this book involve the notions of orthogonality and orthonormality. The concepts are generalizations of the ideas introduced in Example 1.5. Two vectors x and y are *orthogonal* if

$$\langle x, y\rangle = 0 \tag{1.30}$$

The set z_k, $k = 1, 2, \ldots$ is an orthogonal set if, for all members of the set,

$$\langle z_i, z_j\rangle = 0, \qquad i \neq j \tag{1.31}$$

The set is an *orthonormal* set if

$$\langle z_i, z_j\rangle = \delta_{ij} \tag{1.32}$$

where

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \tag{1.33}$$

An orthogonal set is called *proper* if it does not contain the zero vector. We can show that a proper orthogonal set of vectors is linearly independent. Indeed, we form

$$\sum_{k=1}^{n} \alpha_k z_k = \mathbf{0}$$

Taking the inner product of both sides with z_i gives

$$\langle \sum_{k=1}^{n} \alpha_k z_k, z_i \rangle = \langle \mathbf{0}, z_i \rangle$$

Using (1.25) and (1.31), we obtain

$$\alpha_i \langle z_i, z_i \rangle = 0$$

from which we conclude that $\alpha_i = 0, i = 1, 2, \ldots, n$ and the set is linearly independent. Further, if the index n is arbitrary, the countably infinite set $z_k, k = 1, 2, \ldots$, is linearly independent.

1.4 NORMED LINEAR SPACE

A linear space $\mathcal{S}$ is a *normed linear space* if, for every vector $x \in \mathcal{S}$, there is assigned a unique number $\|x\| \in \mathbf{R}$ such that the following rules apply:

a. $\|x\| \geq 0$, with equality if and only if $x = \mathbf{0}$

b. $\|\alpha x\| = |\alpha| \, \|x\|, \alpha \in \mathbf{F}$

c. $\|x_1 + x_2\| \leq \|x_1\| + \|x_2\|$ (triangle inequality)

Although there are many possible definitions of norms, we use exclusively the *norm induced by the inner product*, defined by

$$\|x\| = \sqrt{\langle x, x \rangle} \tag{1.34}$$

Using (1.34), we find that the CSB inequality in (1.28) can be written

$$|\langle x, y \rangle| \leq \|x\| \, \|y\| \tag{1.35}$$

It is easy to show that the norm defined by (1.34) meets the requirements in rules a, b, and c above. We leave the reader to show that a and b are satisfied. For c, for x and y in $\mathcal{S}$, we have

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle \\ &= \langle x, x \rangle + \langle y, y \rangle + \langle x, y \rangle + \langle y, x \rangle \\ &= \|x\|^2 + \|y\|^2 + 2\mathrm{Re}\langle x, y \rangle \end{aligned}$$

Since the real part of a complex number is less than or equal to its magnitude,

$$\|x + y\|^2 \leq \|x\|^2 + \|y\|^2 + 2|\langle x, y \rangle|$$

Using the CSB inequality, we obtain

$$\|x + y\|^2 \le \|x\|^2 + \|y\|^2 + 2\|x\| \, \|y\|$$

Taking the square root of both sides yields the result in c.

EXAMPLE 1.13 From the basic rules for the norm and the definition in (1.34), we can show that

$$\|x + y\|^2 + \|x - y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right) \tag{1.36}$$

Indeed,

$$\begin{aligned} \|x + y\|^2 + \|x - y\|^2 &= \langle x + y, x + y\rangle + \langle x - y, x - y\rangle \\ &= 2\langle x, x\rangle + 2\langle y, y\rangle \end{aligned}$$

from which the result in (1.36) follows.

■

EXAMPLE 1.14 For unitary space $\mathbf{C}_n$, with inner product defined by (1.26), the norm of a vector a in the space is easily found to be

$$\|a\| = \sqrt{\sum_{k=1}^{n} |\alpha_k|^2} \tag{1.37}$$

Unitary space is therefore a normed linear space.

■

EXAMPLE 1.15 For the real linear space $C(\alpha, \beta)$, with inner product defined by (1.27), the norm of a vector f in the space is

$$\|f\| = \sqrt{\int_\alpha^\beta f^2(\xi)d\xi} \tag{1.38}$$

The space $C(\alpha, \beta)$ is therefore a normed linear space.

■

One of the useful consequences of the normed linear space is that it provides a measure of the "closeness" of one vector to another. We note from rule a that $\|x - y\| = 0$ if and only if $x = y$. Therefore, closeness can be indicated by the relation $\|x - y\| < \epsilon$. This observation brings us to the notion of convergence. Among the many forms of convergence, there are two forms whose relationship is crucial to placing firm "boundaries" on the linear space. The type of boundary we seek is one that assures that the limit of a sequence in the linear space also is contained in the space.

In a normed linear space $\mathcal{S}$, a sequence of vectors $\{x_k\}_{k=1}^{\infty}$ *converges* to a vector $x \in \mathcal{S}$ if, given an $\epsilon > 0$, there exists a number N such that $\|x - x_k\| < \epsilon$ whenever $k > N$. We write $x_k \to x$ or

$$\lim_{k\to\infty} x_k = x \tag{1.39}$$

Note that if $x_k \to x$, $\|x - x_k\| \to 0$.

Fundamental to studies of approximation of one vector by another vector, to be studied later in this chapter, is the notion of *continuity of the inner product.* We show that if $\{x_k\}_{k=1}^{\infty}$ is a sequence in $\mathcal{S}$ converging to $x \in \mathcal{S}$, then

$$\langle x_k, h\rangle \to \langle x, h\rangle \tag{1.40}$$

where h is any vector in $\mathcal{S}$. To prove (1.40), it is sufficient to show that

$$\langle x_k, h\rangle - \langle x, h\rangle \to 0$$

or

$$\langle x_k - x, h\rangle \to 0 \tag{1.41}$$

By the form of the CSB inequality in (1.35), we have

$$|\langle x_k - x, h\rangle|^2 \le \|x_k - x\|^2\|h\|^2$$

But, since $x_k \to x$,

$$\|x_k - x\| \to 0$$

so that (1.41) is verified. We remark that another useful way of writing (1.40) is as follows:

$$\lim_{k\to\infty} \langle x_k, h\rangle = \langle \lim_{k\to\infty} x_k, h\rangle \tag{1.42}$$

This relationship indicates that, given x_k, the order of application of the limit and the inner product with h can be interchanged.

In $\mathcal{S}$, a sequence $\{x_k\}_{k=1}^{\infty}$ *converges in the Cauchy sense* if, given an $\epsilon > 0$, there exists a number N such that $\|x_m - x_n\| < \epsilon$ whenever $\min(m, n) > N$. We write

$$\lim_{m,n\to\infty} \|x_m - x_n\| = 0 \tag{1.43}$$

We can show that convergence implies Cauchy convergence. Indeed, let $x \in S$ be defined as the limit of a sequence, as in (1.39). Then, by the triangle inequality,

$$\|x_m - x_n\| \leq \|x_m - x\| + \|x - x_n\|$$

Since $x_k \to x$, there exists a number N such that for $n > N$

$$\|x - x_n\| \leq \frac{\epsilon}{2}$$

and therefore, for $\min(m, n) > N$,

$$\|x_m - x_n\| \leq \epsilon$$

which proves the assertion. Unfortunately, the converse is not always true. The interpretation is that it is possible for two members of the sequence to become arbitrarily close without the sequence itself approaching a limit in S. A normed linear space is said to be *complete* if every Cauchy sequence in the space converges to a vector in the space. The concept of completeness is an important one in what is to follow. Although it is beyond the scope of this book to include a detailed treatment, we shall give a brief discussion.

In real analysis, the space of *rational numbers* is defined [2] as those numbers that can be written as p/q, where p and q are integers. It is a standard exercise [3],[4] to produce a sequence of rational numbers that has the Cauchy property and yet fails to converge in the space. (We consider an example in Problem 1.15.) This incompleteness is caused by the fact that in between two rational numbers, no matter how close, is an infinite number of irrational numbers; often, a sequence of rationals can converge to an irrational. The solution to this problem is a procedure due to Cantor [5] whereby the irrationals are appended to the rationals in such a manner so as to produce a complete linear space called the space of real numbers **R**. We shall assume henceforth that **R** is complete, and direct the reader to the literature in real analysis for details.

EXAMPLE 1.16 We can show that Euclidean space $\mathbf{R}_n$ is complete. For vectors a and b in the space, defined by (1.1) and (1.2), we define an inner product by

$$\langle a, b \rangle = \sum_{k=1}^{n} \alpha_k \beta_k \tag{1.44}$$

Let a_m, $m = 1, 2, \ldots$ be a Cauchy sequence in $\mathbf{R}_n$, where

$$a_m = \left(\alpha_1^{(m)}, \alpha_2^{(m)}, \ldots, \alpha_n^{(m)}\right)$$

Then,

$$\|a_m - a_p\| = \left\{ \sum_{k=1}^{n} \left[\alpha_k^{(m)} - \alpha_k^{(p)} \right]^2 \right\}^{\frac{1}{2}} \leq \epsilon$$

for $\min(m, p) > N$. Since all the terms in the sum are nonnegative, we must have

$$\left| \alpha_k^{(m)} - \alpha_k^{(p)} \right| \leq \epsilon, \qquad k = 1, 2, \ldots, n$$

for $\min(m, p) > N$. Since the space of real numbers **R** is complete, then as $m \to \infty$,

$$\alpha_k^{(m)} \to \alpha_k, \qquad k = 1, 2, \ldots, n$$

and therefore $a_m \to a$.

■

EXAMPLE 1.17 We can show that the normed linear space $C(\alpha, \beta)$ with norm given by (1.38) is incomplete. We shall consider a well-known [6],[7] Cauchy sequence that fails to converge to a vector in the space. Without loss of generality, let (α, β) be (-1,1). Consider the sequence

$$f_k(\xi) = \begin{cases} 0, & -1 \leq \xi \leq 0 \\ k\xi, & 0 \leq \xi \leq \frac{1}{k} \\ 1, & \frac{1}{k} \leq \xi \leq 1 \end{cases} \tag{1.45}$$

where $k = 1, 2, \ldots$. This sequence is continuous, and therefore is in the linear space C(-1,1). We show that this sequence is Cauchy. We form the difference between two members of the sequence. For $m > k$,

$$f_m - f_k = \begin{cases} 0, & -1 \leq \xi \leq 0 \\ (m-k)\xi, & 0 \leq \xi \leq \frac{1}{m} \\ (1 - k\xi), & \frac{1}{m} \leq \xi \leq \frac{1}{k} \\ 0, & \frac{1}{k} \leq \xi \leq 1 \end{cases}$$

We display the two sequence members f_k and f_m in Fig. 1-1. Note that the difference $f_m - f_k$ is always less than unity. It follows that unity is an upper bound on $(f_m - f_k)^2$. We therefore must have

$$\|f_m - f_k\|^2 = \int_{-1}^{1} (f_m(\xi) - f_k(\xi))^2 d\xi \leq \frac{1}{k}$$

Although the above result has been obtained for $m > k$, an interchange of m and k gives the general result

$$\|f_m - f_k\|^2 \leq \max\left(\frac{1}{m}, \frac{1}{k}\right)$$

$$\lim_{m,k\to\infty} \|f_m - f_k\| = 0$$

which proves that the sequence is Cauchy. However, it is apparent that the sequence f_k converges to the Heaviside function $H(\xi)$, defined by

$$H(\xi) = \begin{cases} 0, & \xi < 0 \\ 1, & \xi > 0 \end{cases} \tag{1.46}$$

But, $H(\xi) \notin C(-1,1)$. Therefore, the space $C(\alpha, \beta)$ is not complete.

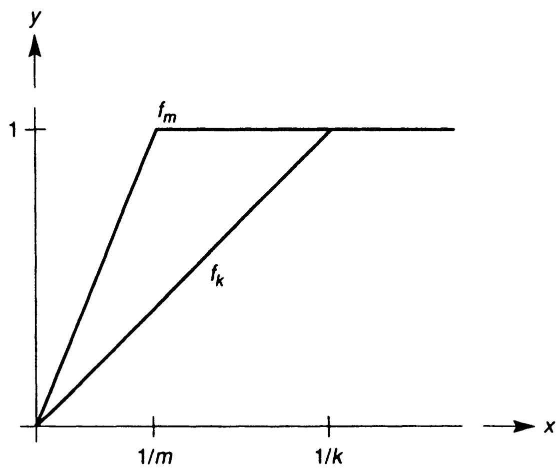

Fig. 1-1 Two members f_k and f_m of sequence in (1.45) for the case $m > k$.

■

1.5 HILBERT SPACE

A linear space is a *Hilbert space* if it is complete in the norm induced by the inner product. Therefore, in any Hilbert space, Cauchy convergence implies convergence. From Example 1.16, Euclidean space $\mathbf{R}_n$ is complete in the norm induced by the inner product in (1.44). Therefore, $\mathbf{R}_n$ is a Hilbert space. In a similar manner, it can be shown that unitary space $\mathbf{C}_n$ is complete. However, from Example 1.17, $C(\alpha, \beta)$ is incomplete. In a manner similar to the completion of the space of rational numbers, C can be completed. The result is $\mathcal{L}_2(\alpha, \beta)$, the Hilbert space of real functions $f(\xi)$ square integrable on the interval (α, β), viz.

$$\int_\alpha^\beta f^2(\xi)d\xi < \infty \tag{1.47}$$

In (1.47), the integration is to be understood in the Lebesgue sense [8]. Although the Lebesgue theory is essential to the understanding of the proof of completeness, in this book such proofs will be omitted. For discussions of the issues involved, the reader is directed to [9],[10].

In linear analysis, we are often concerned with subsets of vectors in a linear space. One such subset is called a *linear manifold.* If $\mathcal{S}$ is a linear space and $\alpha, \beta \in \mathbf{F}$, then $\mathcal{M}$ is a linear manifold in $\mathcal{S}$, provided that $\alpha x + \beta y$ are in $\mathcal{M}$ whenever x and y are in $\mathcal{M}$. It is easy to show that $\mathcal{M}$ is a linear space. The proof is left for Problem 1.16.

EXAMPLE 1.18 In $\mathbf{R}_2$, the set of all vectors in the first quadrant is *not* a linear manifold. Indeed, for x and y in the first quadrant, it is easy to find $\alpha, \beta \in \mathbf{R}$ such that $\alpha x + \beta y$ is not in the first quadrant.

■

EXAMPLE 1.19 Let x_k, $k = 1, 2, \ldots, n$ be a linearly independent sequence of vectors in the Hilbert space $\mathcal{H}$. Define $\mathcal{M}$ to be the set of all linear combinations of the n vectors, viz.

$$\sum_{k=1}^{n} \alpha_k x_k$$

Then, $\mathcal{M}$ is a linear manifold. We leave the proof for Problem 1.17. $\mathcal{M}$ is called the linear manifold *generated* by x_k, $k = 1, 2, \ldots, n$.

■

The results in Example 1.19 raise an interesting issue. Can the same linear manifold be generated by more than one sequence of vectors? We shall show that indeed this is the case. We consider the *Gram–Schmidt orthogonalization process.* Let $\{x_1, \ldots, x_n\}$ be a linearly independent sequence of vectors generating the linear manifold $\mathcal{M} \subset \mathcal{H}$. The Gram–Schmidt process is a constructive procedure for generating an orthonormal sequence $\{e_1, \ldots, e_n\}$ from the independent sequence. To begin, let

$$z_1 = x_1 \tag{1.48}$$

and

$$e_1 = \frac{z_1}{\|z_1\|} \tag{1.49}$$

The next member of the sequence is generated by

$$z_2 = x_2 - \langle x_2, e_1 \rangle e_1 \tag{1.50}$$

and

$$e_2 = \frac{z_2}{\|z_2\|} \tag{1.51}$$

For the third member, we form

$$z_3 = x_3 - \langle x_3, e_1 \rangle e_1 - \langle x_3, e_2 \rangle e_2 \tag{1.52}$$

and

$$e_3 = \frac{z_3}{\|z_3\|} \tag{1.53}$$

This process continues until the final member of the sequence is produced by

$$z_n = x_n - \sum_{k=1}^{n-1} \langle x_n, e_k \rangle e_k \tag{1.54}$$

and

$$e_n = \frac{z_n}{\|z_n\|} \tag{1.55}$$

We leave the reader to show that the sequence $\{e_1, \ldots, e_n\}$ possesses the orthonormal property. In addition, each e_k, $k = 1, \ldots, n$ is a linear combination of $x_1, \ldots, x_k$. We conclude that any linear combination

$$\sum_{k=1}^{n} \alpha_k e_k$$

is also a linear combination

$$\sum_{k=1}^{n} \beta_k x_k$$

The original sequence and the orthonormal sequence obtained from it therefore generate the same linear manifold.

EXAMPLE 1.20 Given the sequence $\{1, \tau, \tau^2, \ldots\} \in \mathcal{L}_2(-1, 1)$, we use the Gram–Schmidt procedure to produce an orthonormal sequence. Indeed, we define an inner product by

$$\langle f, g \rangle = \int_{-1}^{1} f(\tau) g(\tau) d\tau$$

Then,

$$
\begin{aligned}
z_1(\tau) &= 1 \\
e_1(\tau) &= 1/\|1\| = 1/\sqrt{2} \\
z_2(\tau) &= \tau - \langle \tau, 1/\sqrt{2} \rangle (1/\sqrt{2}) = \tau \\
e_2(\tau) &= \sqrt{3/2}\,\tau \\
z_3(\tau) &= \tau^2 - 1/3 \\
e_3(\tau) &= \sqrt{45/8}\left(\tau^2 - 1/3\right)
\end{aligned}
$$

This process continues for as many terms in the orthonormal sequence as we wish to calculate. We remark that the members of the sequence so produced are proportional to the orthogonal sequence of *Legendre polynomials*, whose first few members are

$$
\begin{aligned}
P_0(\tau) &= 1 \\
P_1(\tau) &= \tau \\
P_2(\tau) &= \frac{1}{2}(3\tau^2 - 1) \\
P_3(\tau) &= \frac{1}{2}(5\tau^3 - 3\tau) \\
P_4(\tau) &= \frac{1}{8}(35\tau^4 - 30\tau^2 + 3) \\
P_5(\tau) &= \frac{1}{8}(63\tau^5 - 70\tau^3 + 15\tau)
\end{aligned}
$$

The Legendre functions, orthogonal but not orthonormal, are constructed in such a way that $P_n(\pm 1) = \pm 1$.

■

We next discuss a characteristic associated with linear manifolds that plays a central role in approximation theory. A linear manifold $\mathcal{M}$ is said to be *closed* if it contains the limits of all sequences that can be constructed from the members of $\mathcal{M}$. It is easy to demonstrate that not all linear manifolds are closed. For example, the space $\mathcal{C}(\alpha, \beta)$ is a linear manifold since a linear combination of two continuous functions is a continuous function. However, in Example 1.17, we have given a sequence of vectors in $\mathcal{C}$ that fails to converge to a vector in $\mathcal{C}$. The linear manifold $\mathcal{C}$ is therefore not closed.

An interesting result occurs if a closed linear manifold is contained in a Hilbert space. Specifically, if $\mathcal{H}$ is a Hilbert space and $\mathcal{M}$ is a closed linear manifold in $\mathcal{H}$, then $\mathcal{M}$ is a Hilbert space. Indeed, let $\{x_k\}_{k=1}^{\infty}$ be a Cauchy sequence in $\mathcal{M}$. Then, since $\mathcal{M}$ is contained in the Hilbert space

$\mathcal{H}$, $x_k \to x \in \mathcal{H}$. But $\mathcal{M}$ is closed, and therefore $x \in \mathcal{M}$. We conclude that $\mathcal{M}$ is a Hilbert space.

1.6 BEST APPROXIMATION

Within the structure of the Hilbert space, it is possible to generalize the concepts of approximation of vectors and functions. Let x be a vector in a Hilbert space $\mathcal{H}$ and let $\{z_k\}_{k=1}^m$ be an orthonormal set in $\mathcal{H}$. We form the sum

$$x_m = \sum_{k=1}^{m} \alpha_k z_k \tag{1.56}$$

This sum generates a linear manifold $\mathcal{M} \subset \mathcal{H}$. Different members of the linear manifold are produced by assigning various values to the sequence of coefficients $\{\alpha_k\}_{k=1}^m$. We should like to determine what choice of coefficients results in x_m being the "best" approximation to x. Specifically, let us make x_m "close" to x by adjusting the coefficients to minimize $\|x - x_m\|$. We expand the square of the norm as follows:

$$\begin{aligned}
\|x - x_m\|^2 &= \langle x - x_m, x - x_m \rangle \\
&= \langle x, x \rangle + \langle x_m, x_m \rangle - \langle x_m, x \rangle - \langle x, x_m \rangle \\
&= \|x\|^2 + \sum_{k=1}^{m} |\alpha_k|^2 - \sum_{k=1}^{m} \alpha_k \overline{\langle x, z_k \rangle} - \sum_{k=1}^{m} \overline{\alpha_k} \langle x, z_k \rangle
\end{aligned}$$

Completing the square, we obtain

$$\|x - x_m\|^2 = \|x\|^2 + \sum_{k=1}^{m} (\alpha_k - \langle x, z_k \rangle) \overline{(\alpha_k - \langle x, z_k \rangle)} - \sum_{k=1}^{m} |\langle x, z_k \rangle|^2 \tag{1.57}$$

Since the sum in (1.57) containing the coefficients α_k is nonnegative, the norm-squared (and hence the norm) is minimized by the choice

$$\alpha_k = \langle x, z_k \rangle, \qquad k = 1, 2, \ldots, m \tag{1.58}$$

Expression (1.58) defines the *Fourier coefficients* associated with orthonormal expansions. Note that once we have made the selection given in (1.58), we can define an error vector e_m by

$$\begin{aligned}
e_m &= x - x_m \\
&= x - \sum_{k=1}^{m} \langle x, z_k \rangle z_k
\end{aligned} \tag{1.59}$$

Taking the inner product with any member z_j of the orthonormal sequence, we find that

$$\langle e_m, z_j\rangle = \langle x, z_j\rangle - \sum_{k=1}^{m}\langle x, z_k\rangle\langle z_k, z_j\rangle = 0, \qquad j = 1, 2, \ldots, m \quad (1.60)$$

Since x_m is a linear combination of members of the sequence $\{z_j\}_{j=1}^{m}$, we must have

$$\langle e_m, x_m\rangle = 0 \qquad (1.61)$$

We summarize these results as follows:

a. The vector $x \in \mathcal{H}$ has been decomposed into a vector $x_m \in \mathcal{M} \subset \mathcal{H}$ plus an error vector e_m, viz.

$$x = x_m + e_m \qquad (1.62)$$

b. The error vector e_m is orthogonal to the approximation vector x_m.

EXAMPLE 1.21 We consider the Fourier sine series in Example 1.5. Let $f(\xi) \in \mathcal{L}_2(0, 1)$ with inner product

$$\langle f, g\rangle = \int_0^1 f(\xi)g(\xi)d\xi \qquad (1.63)$$

Let $\{\sqrt{2}\sin k\pi\xi\}_{k=1}^{m}$ be an orthonormal set generating a linear manifold $\mathcal{M} \subset \mathcal{H}$. Then, by (1.56) and (1.58), the "best" approximating function $f_m \in \mathcal{M}$ to $f(\xi)$ is given by

$$f_m(\xi) = \sum_{k=1}^{m}\alpha_k\sqrt{2}\sin k\pi\xi \qquad (1.64)$$

where

$$\alpha_k = \int_0^1 f(\eta)\sqrt{2}\sin k\pi\eta d\eta \qquad (1.65)$$

which is the classical Fourier result.

■

The above results for the approximation of a vector $x \in \mathcal{H}$ by a vector $x_m \in \mathcal{M} \subset \mathcal{H}$ can be generalized. We shall need the concept of a manifold that is orthogonal to a given manifold. If $\mathcal{M}$ is a linear manifold, then the vector $e \in \mathcal{H}$ is a member of a set $\mathcal{M}^{\perp}$ if it is orthogonal to every vector in $\mathcal{M}$. The set $\mathcal{M}^{\perp}$ is a linear manifold since linear combinations of vectors in $\mathcal{M}^{\perp}$ are also orthogonal to vectors in $\mathcal{M}$. In fact, $\mathcal{M}^{\perp}$ is also closed.

Indeed, let $\{e_k\}_{k=1}^{\infty}$ be a sequence in $\mathcal{M}^{\perp}$ converging to a vector e in $\mathcal{H}$. We have

$$\langle e_k, x\rangle = 0$$

for all $x \in \mathcal{M}$. But, by continuity of the inner product,

$$\lim_{k\to\infty} \langle e_k, x\rangle = \langle e, x\rangle = 0$$

so that $e \in \mathcal{M}^{\perp}$, which is therefore closed. The closed linear manifold $\mathcal{M}^{\perp}$ is called the *orthogonal complement* to $\mathcal{M}$.

We now can produce a result called the *Projection Theorem.* Let x be any vector in the Hilbert space $\mathcal{H}$, and let $\mathcal{M} \subset \mathcal{H}$ be a *closed* linear manifold. Then, there is a unique vector $y_0 \in \mathcal{M} \subset \mathcal{H}$ closest to x in the sense that $\|x - y_0\| \leq \|x - y\|$ for all y in $\mathcal{M}$. Furthermore, the necessary and sufficient condition that y_0 is the unique minimizing vector is that $e = x - y_0$ is in $\mathcal{M}^{\perp}$. The proof of this important theorem is deferred to Appendix A.1 at the end of the chapter. The vector y_0 is called the *projection* of x onto $\mathcal{M}$. The vector e is called the projection of x onto $\mathcal{M}^{\perp}$. The ideas inherent to the projection theorem have a well-known interpretation in two- and three-dimensional vector spaces, as in the following example.

EXAMPLE 1.22 Let $a = (\alpha_1, \alpha_2)$ be any vector in $\mathbf{R}_2$. Let $\mathcal{M}$ be the set of all vectors b in $\mathbf{R}_2$ with second component equal to zero, viz.

$$b = (\beta_1, 0)$$

Since linear combinations of vectors in $\mathcal{M}$ are also in $\mathcal{M}$, the set $\mathcal{M}$ is a linear manifold. In addition, the manifold is closed. Indeed, let $b_1, b_2, \ldots$ be a sequence in $\mathcal{M}$ converging to $b \in \mathcal{H}$, where

$$b_k = \left(\beta_1^{(k)}, 0\right)$$

Since $b_k \to b \in \mathcal{H}$, $\beta_1^{(k)} \to \beta_1$. Therefore, $b \in \mathcal{M}$. By the projection theorem, among all vectors $b \in \mathcal{M}$, the vector $\hat{b}$ closest to a can be obtained from

$$\langle a - \hat{b}, b\rangle = 0$$

where

$$\hat{b} = (\hat{\beta}_1, 0)$$

Solving this equation, employing the usual definition of inner product for $\mathbf{R}_2$, we obtain

$$(\alpha_1 - \hat{\beta}_1)\beta_1 = 0$$

Since β_1 is arbitrary, $\hat{\beta}_1 = \alpha_1$. This result can be visualized (Fig. 1-2) by drawing the vector a and noting that the vector b lies along the x-axis. The "best" b is then obtained by dropping a perpendicular from the tip of a to the x-axis. The result is a vector $\hat{b}$, called the projection of a onto the x-axis. Note that the error vector e is such that it is orthogonal to any vector along the x-axis and that $a = \hat{b} + e$, as required by the projection theorem.

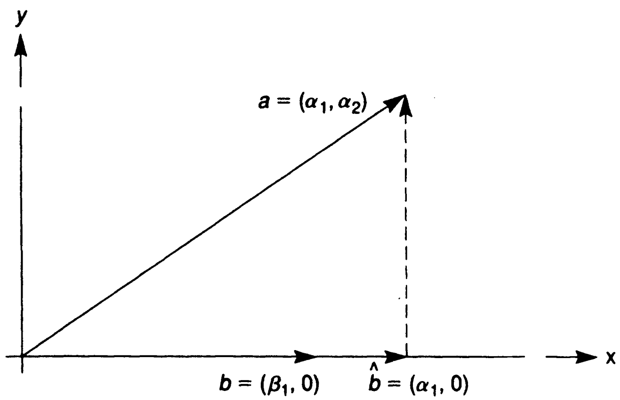

Fig. 1-2 Illustration of the projection theorem in $\mathbf{R}_2$, as given in Example 1.22.

■

In (1.56)–(1.58), we introduced best approximation in terms of orthonormal expansion functions generating a linear manifold. With the aid of the projection theorem, we next generalize the concept of best approximation to include expansion functions that are linearly independent but not necessarily orthogonal. Let $y \in \mathcal{H}$, and let $\{y_j\}_{j=1}^{M}$ be a linearly independent sequence of vectors in $\mathcal{H}$. We form the sum

$$\hat{y} = \sum_{j=1}^{M} \alpha_j y_j \tag{1.66}$$

We wish to approximate y with $\hat{y}$ by suitable choice of the coefficients α_j. We have already indicated in Example 1.19 (Problem 1.17) that linear combinations of the type in (1.66) form a linear manifold $\mathcal{M}$. In fact, since the limit of sequences of vectors in $\mathcal{M}$ must necessarily be in $\mathcal{M}$, the manifold is closed and therefore meets the requirements of the projection theorem. Since $y_j \in \mathcal{M}$, $j = 1, 2, \ldots, M$, the projection theorem gives

$$\langle y - \hat{y}, y_k \rangle = 0, \qquad k = 1, 2, \ldots, M \tag{1.67}$$

Substituting (1.66) and rearranging, we have

$$\sum_{j=1}^{M} \alpha_j \langle y_j, y_k \rangle = \langle y, y_k \rangle, \qquad k = 1, 2, \ldots, M \tag{1.68}$$

We write (1.68) in matrix form as follows:

$$\begin{bmatrix} \langle y_1, y_1 \rangle & \langle y_2, y_1 \rangle & \cdots & \langle y_M, y_1 \rangle \\ \langle y_1, y_2 \rangle & \langle y_2, y_2 \rangle & \cdots & \langle y_M, y_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle y_1, y_M \rangle & \langle y_2, y_M \rangle & \cdots & \langle y_M, y_M \rangle \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_M \end{bmatrix} = \begin{bmatrix} \langle y, y_1 \rangle \\ \langle y, y_2 \rangle \\ \vdots \\ \langle y, y_M \rangle \end{bmatrix} \tag{1.69}$$

Inversion of this matrix yields the coefficients α_j, $j = 1, 2, \ldots, M$. These coefficients then determine $\hat{y}$ in (1.66). The square matrix on the left-hand side of (1.69) is the *transpose* of a matrix called the *Gram matrix*. In addition to its appearance in best approximation, it also finds use in proofs of linear independence (Problem 1.21). Note that the result in (1.69) is a generalization of the Fourier coefficient result in (1.58). Indeed, for cases where the independent sequence y_j, $j = 1, 2, \ldots, M$ is orthogonal, the matrix in (1.69) diagonalizes. Inversion then produces the Fourier coefficient result.

One of the classic problems of algebra is the approximation of a function by a polynomial. This problem is easily cast as best approximation in the following example.

EXAMPLE 1.23 In the Hilbert space $\mathcal{L}_2(0, 1)$, consider the approximation of a function $f(\tau)$ by a polynomial. Let $\{\tau^{n-1}\}_{n=1}^{N}$ be a sequence in $\mathcal{L}_2(0, 1)$. The sequence is linearly independent. Indeed, by the fundamental theorem of algebra, the equation

$$\sum_{n=1}^{N} \beta_n \tau^{n-1} = 0$$

can have at most $N - 1$ roots. Therefore, the only solution valid for all $\tau \in (0, 1)$ is $\beta_n = 0, n = 1, 2, \ldots, N$. We wish to approximate $f(\tau)$ by

$$\hat{f}(\tau) = \sum_{n=1}^{N} \alpha_n \tau^{n-1} \tag{1.70}$$

All possible such linear combinations form a closed linear manifold so that the projection theorem applies. Comparing (1.66), we identify

$$y_n = \tau^{n-1} \tag{1.71}$$

Define an inner product for $\mathcal{L}_2(0, 1)$ by

$$\int_0^1 f(\tau)g(\tau)d\tau \tag{1.72}$$

We then have

$$\langle y_m, y_n \rangle = \int_0^1 \tau^{m-1}\tau^{n-1}d\tau = \frac{1}{m+n-1} \tag{1.73}$$

$$\langle y, y_m \rangle = \int_0^1 \tau^{m-1}f(\tau)d\tau \tag{1.74}$$

Substitution of (1.73) and (1.74) into (1.69) gives

$$\begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{N} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{N+1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{N} & \frac{1}{N+1} & \cdots & \frac{1}{2N-1} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix} = \begin{bmatrix} \int_0^1 f(\tau)d\tau \\ \int_0^1 \tau f(\tau)d\tau \\ \vdots \\ \int_0^1 \tau^{N-1} f(\tau)d\tau \end{bmatrix} \tag{1.75}$$

Inversion of this matrix equation yields the best approximation.

■

1.7 OPERATORS IN HILBERT SPACE

Consider the following transformation in $\mathbf{R}_3$:

$$\begin{aligned} \zeta_1 &= \alpha_{11}\xi_1 + \alpha_{12}\xi_2 + \alpha_{13}\xi_3 \\ \zeta_2 &= \alpha_{21}\xi_1 + \alpha_{22}\xi_2 + \alpha_{23}\xi_3 \\ \zeta_3 &= \alpha_{31}\xi_1 + \alpha_{32}\xi_2 + \alpha_{33}\xi_3 \end{aligned}$$

Using the usual matrix notation, we let

$$z = [\,\zeta_1 \quad \zeta_2 \quad \zeta_3\,]^T$$

$$x = [\,\xi_1 \quad \xi_2 \quad \xi_3\,]^T$$

where T indicates matrix transpose. We then have

$$Ax = z$$

where

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix}$$

The solution is given formally by

$$x = A^{-1}z$$

The matrix operation is linear. Indeed, given x_1, x_2, z_1, and z_2, we have by ordinary matrix methods

$$A(\alpha x_1 + \beta x_2) = \alpha A x_1 + \beta A x_2 = \alpha z_1 + \beta z_2$$

The concepts of linearity and inversion for matrices can be generalized to linear operators in a Hilbert space.

An *operator* L is a mapping that assigns to a vector $x \in \mathcal{S}$ another vector $Lx \in \mathcal{S}$. We write

$$Lx = y \tag{1.76}$$

The *domain* of the operator L is the set of vectors x for which the mapping is defined. The *range* of the operator L is the set of vectors y resulting from the mapping. The operator L is *linear* if the mapping is such that for any x_1 and x_2 in the domain of L, the vector $\alpha_1 x_1 + \alpha_2 x_2$ is also in the domain and

$$L\,(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 L x_1 + \alpha_2 L x_2 \tag{1.77}$$

A linear operator L with domain $\mathcal{D}_L \subset \mathcal{H}$ is *bounded* if there exists a real number γ such that

$$\|Lu\| \le \gamma \|u\| \tag{1.78}$$

for all $u \in \mathcal{D}_L$.

EXAMPLE 1.24 Let $\mathbf{R}_\infty$ be the space of all vectors consisting of a countably infinite set of real numbers (components), viz.

$$a = (\alpha_1, \alpha_2, \alpha_3 \ldots) \tag{1.79}$$

where $\alpha_k \in \mathbf{R}$. If $b = (\beta_1, \beta_2, \beta_3, \ldots)$, define an inner product for the space by

$$\langle a, b\rangle = \sum_{k=1}^{\infty} \alpha_k \beta_k \tag{1.80}$$

Let the norm for the space be induced by the inner product. We restrict $\mathbf{R}_\infty$ to those vectors with finite norm. Define the *right shift* operator A_R in $\mathbf{R}_\infty$ by

$$A_R a = (0, \alpha_1, \alpha_2, \ldots)$$

The right shift operator A_R is linear. The proof is easy and is omitted. In addition, A_R is bounded. Indeed,

$$\|A_R a\| = \sqrt{\sum_{k=1}^{\infty} \alpha_k^2} = \|a\|$$

Therefore, the operator A_R is bounded in $\mathbf{R}_\infty$. Indeed, the least upper bound on γ is unity.

■

EXAMPLE 1.25 On the complex Hilbert space $\mathcal{L}_2(0, 1)$, we consider the following integral equation:

$$\int_0^1 u(\xi')k(\xi, \xi')d\xi' = f(\xi) \tag{1.81}$$

This equation can be written in operator notation as follows:

$$Lu = f$$

where L is the linear operator given by

$$L = \int_0^1 (\cdot)k(\xi, \xi')d\xi' \tag{1.82}$$

We shall show that the operator L is bounded if

$$\int_0^1 \int_0^1 |k(\xi, \xi')|^2 d\xi d\xi' < \infty \tag{1.83}$$

This property of the *kernel* $k(\xi, \xi')$ is called the *Hilbert–Schmidt* property, and the operator L it generates is called a *Hilbert–Schmidt* operator. To show that L is bounded, we form

$$\|Lu\|^2 = \int_0^1 |f(\xi)|^2 d\xi$$

where

$$\begin{aligned}|f(\xi)|^2 &= \left|\int_0^1 u(\xi')k(\xi, \xi')d\xi'\right|^2 \\ &\le \int_0^1 |u(\xi')|^2 d\xi' \int_0^1 |k(\xi, \xi')|^2 d\xi' \\ &= \|u\|^2 \int_0^1 |k(\xi, \xi')|^2 d\xi'\end{aligned}$$

It follows that

$$\|Lu\|^2 \le \|u\|^2 \left[\int_0^1 \int_0^1 |k(\xi, \xi')|^2 d\xi d\xi'\right]$$

and finally,

$$\|Lu\| \le M\|u\|$$

where M is the bound on the double integral.

■

A linear operator L with domain $\mathcal{D}_L \subset \mathcal{H}$ is *continuous* if given an $\epsilon > 0$, there exists a $\delta > 0$ such that, for every $u_0 \in \mathcal{D}_L$, $\|Lu_0 - Lu\| < \epsilon$, for all $u \in \mathcal{D}_L$ satisfying $\|u_0 - u\| < \delta$. We can interpret this definition to mean that when an operator is continuous, Lu_0 is close to Lu whenever u_0 is close to u. There is an important theorem on interchange of operators and limits that follows immediately from the above definition. A linear operator L with domain $\mathcal{D}_L \subset \mathcal{H}$ is continuous if and only if for every sequence $\{u_n\}_{n=1}^{\infty} \in \mathcal{D}_L$ converging to $u_0 \in \mathcal{D}_L$,

$$Lu_0 = L \lim_{n\to\infty} u_n = \lim_{n\to\infty} Lu_n \tag{1.84}$$

The proof is in two parts. First, we suppose that L is continuous and $\epsilon > 0$ is given. We may select a δ according to the definition of continuity and suppose that $\|u_0 - u_n\| < \delta$. Since u_n is a member of a converging sequence, $\|u_0 - u_n\| < \delta$ for all $n \geq N$. Therefore,

$$\|Lu_0 - Lu_n\| < \epsilon, \qquad n \geq N$$

and

$$Lu_0 = \lim_{n\to\infty} Lu_n$$

This first part of the proof shows that, if an operator is continuous, the operator and the limit can be interchanged. In the second part of the proof, we must show that if the operator and limit can be interchanged, the operator is continuous. This part is not essential to our development and is omitted. The interested reader is referred to [12].

We now give a theorem linking the boundedness and continuity of operators. A linear operator L with domain $\mathcal{D}_L \subset \mathcal{H}$ is bounded if and only if it is continuous. The proof is in two parts. In the first part, we show that if the operator is bounded, it is continuous. Indeed, if L is bounded and $u_0 \in \mathcal{D}_L$,

$$\begin{aligned}\|Lu_0 - Lu\| &= \|L(u_0 - u)\| \\ &\leq \gamma \|u_0 - u\|\end{aligned}$$

for all $u \in \mathcal{D}_L$. Then, given any $\epsilon > 0$, it is easy to find a $\delta > 0$ such that

$$\|Lu_0 - Lu\| < \epsilon$$

whenever

$$\|u_0 - u\| < \delta$$

Indeed, the choice $\delta = \epsilon/\gamma$ is sufficient. In the second part of the proof, we must show that if an operator is continuous, it is bounded. This part is not essential to our development and is omitted. The interested reader is referred to [13].

It is straightforward to show that the differential operator $L = d/d\xi$ is unbounded. The proof is by contradiction. We suppose that $d/d\xi$ is bounded. Then, it is continuous. Therefore, for any $u_n \to u$, we must have

$$Lu = L \lim_{n\to\infty} u_n = \lim_{n\to\infty} Lu_n$$

But, if we choose u_n as a member of the sequence

$$u_n = \frac{\cos n\pi\xi}{n}, \qquad n = 1, 2, \ldots$$

we have

$$\lim_{n\to\infty} u_n = 0$$

and therefore,

$$L \lim_{n\to\infty} u_n = 0$$

But,

$$\lim_{n\to\infty} Lu_n = \lim_{n\to\infty} (-\pi \sin n\pi\xi)$$

and this limit is undefined. We therefore have arrived at a contradiction, and we conclude that $d/d\xi$ is unbounded.

Given the concepts of continuity and boundedness of a linear operator, we can show that a bounded linear operator is uniquely determined by a matrix. Indeed, let $\{z_k\}_{k=1}^{\infty}$ be a basis for $\mathcal{H}$. Let L be a bounded linear operator with

$$Lu = f$$

We expand u in the basis as follows:

$$u = \lim_{n\to\infty} \sum_{k=1}^{n} \alpha_k z_k \tag{1.85}$$

Since boundedness implies continuity,

$$Lu = L \lim_{n\to\infty} \sum_{k=1}^{n} \alpha_k z_k = \lim_{n\to\infty} L \sum_{k=1}^{n} \alpha_k z_k$$

By the linearity of the operator L, we then have

$$\lim_{n\to\infty} \sum_{k=1}^{n} \alpha_k L z_k = f \tag{1.86}$$

We take the inner product of both sides of (1.86) with a member of the basis set to give

$$\langle \lim_{n\to\infty} \sum_{k=1}^{n} \alpha_k L z_k, z_j \rangle = \langle f, z_j \rangle, \qquad j = 1, 2, \ldots$$

By continuity of the inner product and the rules for inner products, we obtain

$$\lim_{n\to\infty} \sum_{k=1}^{n} \alpha_k \langle L z_k, z_j \rangle = \langle f, z_j \rangle, \qquad j = 1, 2, \ldots \tag{1.87}$$

Equation (1.87) is a matrix equation that can be written in standard matrix notation as follows:

$$\begin{bmatrix} \langle Lz_1, z_1 \rangle & \langle Lz_2, z_1 \rangle & \cdots \\ \langle Lz_1, z_2 \rangle & \langle Lz_2, z_2 \rangle & \cdots \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} \langle f, z_1 \rangle \\ \langle f, z_2 \rangle \\ \vdots \end{bmatrix} \tag{1.88}$$

If this matrix can be inverted to give the coefficients $\alpha_1, \alpha_2, \ldots$, substitution into (1.85) completes the determination of u.

EXAMPLE 1.26 On the real Hilbert space $\mathcal{L}_2(0, 1)$, consider the integral equation

$$f(\xi) = -\mu u(\xi) + \int_0^1 k(\xi, \xi') u(\xi') d\xi' \tag{1.89}$$

where

$$k(\xi, \xi') = \begin{cases} \xi(1 - \xi'), & 0 \le \xi \le \xi' \\ \xi'(1 - \xi), & \xi' \le \xi \le 1 \end{cases} \tag{1.90}$$

In operator notation,

$$(L - \mu I)u = f, \qquad \xi \in (0, 1) \tag{1.91}$$

where I is the identity operator. The operator $L - \mu I$ is bounded. We leave the proof for Problem 1.26. We wish to obtain the matrix representation and thereby solve the integral equation. We define the inner product for the space as in (1.63). For basis functions, we choose

$$z_n = \sin n\pi\xi, \qquad n = 1, 2, \ldots \tag{1.92}$$

Then, operating on any member of the basis set, we obtain

$$(L - \mu I)z_n = -\mu \sin n\pi\xi + \int_0^1 k(\xi, \xi') \sin n\pi\xi' d\xi' \tag{1.93}$$

But, using (1.90), we find that

$$\int_0^1 k(\xi,\xi')\sin n\pi\xi' d\xi' = (1-\xi)\int_0^\xi \xi' \sin n\pi\xi' d\xi' + \xi\int_\xi^1 (1-\xi')\sin n\pi\xi' d\xi'$$

After some elementary integrations, we obtain the general matrix element in the square matrix in (1.88), viz.

$$\begin{aligned}\langle (L-\mu I)z_n, z_m\rangle &= \left[\frac{1}{(n\pi)^2}-\mu\right]\langle z_n, z_m\rangle \\ &= \frac{1}{2}\left[\frac{1}{(n\pi)^2}-\mu\right]\delta_{nm} \qquad (1.94)\end{aligned}$$

where the inner product is the usual inner product for $\mathcal{L}_2(0,1)$ and δ_{nm} has been defined in (1.33). The matrix representation in (1.88) therefore diagonalizes, and the inversion yields

$$\alpha_k = \frac{2}{\frac{1}{(k\pi)^2}-\mu}\langle f, z_k\rangle, \qquad k = 1, 2, \ldots \qquad (1.95)$$

Substitution of (1.95) into (1.85) yields the final result, viz.

$$u(\xi) = 2\sum_{k=1}^{\infty}\frac{\int_0^1 f(\xi')\sin k\pi\xi' d\xi'}{\frac{1}{(k\pi)^2}-\mu}\sin k\pi\xi \qquad (1.96)$$

■

In the above example of representation of an operator by a matrix, the choice of the basis functions resulted in diagonalization of the matrix and, therefore, trivial matrix inversion. There are many operators, however, that do not have properties that result in this diagonalization. These concepts are better understood after a study of operator properties and resulting Green's functions and spectral representations in the next two chapters.

An important collection of operators for which there are established convergence criteria are nonnegative, positive, and positive-definite operators. The reader is cautioned that there is little uniformity of notation concerning these operators in the literature. For the purposes herein, an operator L is *nonnegative* if $\langle Lx, x\rangle \geq 0$, for all $x \in \mathcal{D}_L$. An operator is *positive* if $\langle Lx, x\rangle > 0$, for all $x \neq 0$ in $\mathcal{D}_L$. An operator is *positive-definite* if $\langle Lx, x\rangle \geq c^2\|x\|^2$, for $c > 0$ and $x \in \mathcal{D}_L$. An operator L is *symmetric* if $\langle Lx, x\rangle = \langle x, Lx\rangle$. It is easy to show that nonnegative, positive, and positive-definite operators are symmetric. In fact, any operator having the property that $\langle Lx, x\rangle$ is real is symmetric. Indeed,

$$\langle x, Lx\rangle = \overline{\langle Lx, x\rangle} = \langle Lx, x\rangle$$

A special inner product and norm [17], associated with positive and positive-definite operators, are useful in relating convergence criteria. Define the *energy inner product* with respect to the operator L by

$$[x, y] = \langle Lx, y \rangle \tag{1.97}$$

With this inner product definition, $\mathcal{D}_L$ becomes a Hilbert space $\mathcal{H}_L$. The associated *energy norm* in $\mathcal{H}_L$ is given by

$$\left| x \right| = \sqrt{\langle Lx, x \rangle} \tag{1.98}$$

We emphasize that the operator L must be positive for (1.98) to satisfy the basic definitions of a norm. Indeed, the energy inner product and norm defined in (1.97) and (1.98) must be shown in each case to satisfy the rules for inner products and norms. For positive-definite operators, we can prove the following important relationship between norms:

$$\|x\| \leq \frac{1}{c} \left| x \right| \tag{1.99}$$

Indeed,

$$\left| x \right|^2 = \langle Lx, x \rangle \geq c^2 \|x\|^2$$

Therefore,

$$\|x\|^2 \leq \frac{1}{c^2} \left| x \right|^2$$

Taking the square root of both sides yields the desired result.

Among the many forms of convergence criteria, there are several types that are particularly useful in numerical methods in electromagnetics. For a sequence $\{u_n\} \subset \mathcal{H}$, u_n *converges to* u is written

$$u_n \rightarrow u \tag{1.100}$$

and means that

$$\lim_{n \to \infty} \|u_n - u\| = 0 \tag{1.101}$$

The statement u_n *converges in energy to* u is written

$$u_n \xrightarrow{e} u \tag{1.102}$$

and means that

$$\lim_{n \to \infty} \left| u_n - u \right| = 0 \tag{1.103}$$

The statement u_n *converges weakly to* u is written

$$u_n \xrightarrow{w} u \tag{1.104}$$

and means that for every $g \in \mathcal{H}$

$$\lim_{n\to\infty} |\langle u_n - u, g\rangle| = 0 \tag{1.105}$$

It is straightforward to show the following relationships among the types of convergence:

A. If $\|Lu_n\|$ is bounded, convergence implies convergence in energy.

B. Convergence implies weak convergence.

C. Convergence in energy implies $Lu_n \xrightarrow{w} f$. The weak convergence is for those g, defined by (1.105), in $\mathcal{H}_L$. If, however, $\|Lu_n\|$ is bounded, then $Lu_n \xrightarrow{w} f$ in $\mathcal{H}$.

D. If L is positive-definite, convergence in energy implies convergence.

We first prove Property A. We have

$$\begin{aligned} |u - u_n|^2 &= |\langle L(u_n - u), u_n - u\rangle| \le \|L(u_n - u)\| \, \|u_n - u\| \\ &= \|Lu_n - Lu\| \, \|u_n - u\| \le (\|Lu_n\| + \|Lu\|) \, \|u_n - u\| \end{aligned}$$

Since, by hypothesis, $\|Lu_n\|$ is bounded and $u_n \to u$, a limiting operation gives

$$\lim_{n\to\infty} |u_n - u|^2 = 0$$

To prove Property B, we use the CSB inequality to give

$$|\langle u_n - u, g\rangle| \le \|u_n - u\| \, \|g\|$$

for any g in $\mathcal{H}$. Taking the limit yields the desired result, viz.

$$\lim_{n\to\infty} |\langle u_n - u, g\rangle| = 0$$

To prove Property C, we have

$$|\langle Lu_n - f, g\rangle| = |\langle L(u_n - u), g\rangle| = |[u_n - u, g]| \le |u_n - u| |g|$$

where we have used the CSB inequality on Hilbert space $\mathcal{H}_L$. By hypothesis, we have convergence in energy. Therefore,

$$\lim_{n\to\infty} |\langle Lu_n - f, g\rangle| = 0$$

for $g \in \mathcal{H}_L$. This procedure proves the first half of Property C. The proof of the second half is based on the Hilbert space $\mathcal{H}_L$ being *dense* in $\mathcal{H}$ and is omitted. (See [17, p. 24–25].) To prove Property D, we write

$$\|u_n - u\| \leq \frac{1}{c}\left| u_n - u \right|$$

Taking the limit as $n \to \infty$ yields the desired result.

1.8 METHOD OF MOMENTS

The purpose of this section is to introduce the Method of Moments in a general way and develop various special cases. Emphasis is on convergence and error minimization.

If L is a linear operator, an approximate solution to $Lu = f$ is given by the following procedure. For L an operator in $\mathcal{H}$, consider

$$Lu - f = 0 \tag{1.106}$$

where $u \in \mathcal{D}_L$, $f \in \mathcal{R}_L$. Define the linearly independent sets $\{\phi_k\}_{k=1}^n \subset \mathcal{D}_L$ and $\{w_k\}_{k=1}^n$, where ϕ_k and w_k are called *expansion* functions and *weighting* functions, respectively. Define a sequence of approximants to u by

$$u_n = \sum_{k=1}^{n} \alpha_k \phi_k, \qquad n = 1, 2, \ldots \tag{1.107}$$

A matrix equation is formed in (1.106) by the condition that, upon replacement of u by u_n, the left side shall be orthogonal to the sequence $\{w_k\}$. We have

$$\langle Lu_n - f, w_m \rangle = 0, \qquad m = 1, 2, \ldots, n \tag{1.108}$$

Substitution of (1.107) into (1.108) and use of (1.25) gives the matrix equation of the *Method of Moments* [18],[19], viz.

$$\sum_{k=1}^{n} \alpha_k \langle L\phi_k, w_m \rangle = \langle f, w_m \rangle, \qquad m = 1, 2, \ldots, n \tag{1.109}$$

Note that the *exact* operator equation (1.106) in a Hilbert space $\mathcal{H}$ has been transformed into an *approximate* operator equation on Hilbert space $\mathbf{C}_n$, viz.

$$Ax = b \tag{1.110}$$

where, in usual matrix form,

$$x = (\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n)^T \tag{1.111}$$

$$b = (\langle f, w_1 \rangle \quad \langle f, w_2 \rangle \quad \cdots \quad \langle f, w_n \rangle)^T \tag{1.112}$$

$$A = [a_{mk}] \tag{1.113}$$

where T denotes *transpose* and a_{mk} are the individual matrix elements, given by

$$a_{mk} = \langle L\phi_k, w_m \rangle \tag{1.114}$$

We note the following interesting result. If the operator L is bounded, if $w_k = \phi_k$, and if the sequence ϕ_k, $k = 1, 2, \ldots$ is a basis for the Hilbert space, then taking the limit as $n \to \infty$ in (1.109) reproduces the result in (1.87). We therefore view the Method of Moments given by (1.109) as an extension to the matrix representation result for bounded operators in (1.87). There remains, however, a basic question concerning the convergence of (1.107) when the sequence $\{\alpha_k\}$ is determined by solution of the matrix equation (1.110).

In the special case where the expansion functions are identical to the weighting functions, the result is *Galerkin's Method* [17], viz.

$$\sum_{k=1}^{n} \alpha_k \langle L\phi_k, \phi_m \rangle = \langle f, \phi_m \rangle, \qquad m = 1, 2, \ldots, n \tag{1.115}$$

If nothing is known about the mathematical properties of the operator L other than its linearity, nothing in general can be said concerning the convergence of the approximants u_n to the solution u. Unfortunately, most of the interesting and practical problems in electromagnetics involve operators that are neither positive nor positive-definite. Therefore, most of the large body of solutions to electromagnetic problems by the Method of Moments lack any sort of mathematical convergence criteria.

If, however, the operator L is positive, we may define the Hilbert space $\mathcal{H}_L$ with the norm given by (1.98). We then write (1.115) as follows:

$$\sum_{k=1}^{n} \alpha_k [\phi_k, \phi_m] = [u, \phi_m], \qquad m = 1, 2, \ldots, n \tag{1.116}$$

We further assume that the sequence $\{\phi_k\}$ is complete in $\mathcal{H}_L$. We may show that, under these circumstances, Galerkin's Method results in convergence in energy. Indeed, since the complete sequence $\{\phi_k\}$ defines $\mathcal{H}_L$, we may apply the Gram–Schmidt orthonormalization procedure. Assuming that the ϕ_m are orthonormal in (1.116), we obtain

$$\alpha_k = [u, \phi_k] \tag{1.117}$$

Substitution into (1.107) gives

$$u_n = \sum_{k=1}^{n} [u, \phi_k]\phi_k \tag{1.118}$$

which is the Fourier series expansion in $\mathcal{H}_L$ of u_n with Fourier coefficients given by (1.117). Therefore,

$$\lim_{n\to\infty} \left| u_n - u \right| = 0 \tag{1.119}$$

By Property C, the result in (1.119) implies that $Lu_n \xrightarrow{w} f$. Unfortunately, nothing can be said about the nearness of u_n to u. If, however, L is positive-definite, Property D states that the approximants converge, viz.

$$\lim_{n\to\infty} \|u_n - u\| = 0 \tag{1.120}$$

In the Galerkin procedure, if the operator L is positive and the sequence $\{\phi_k\}$ is complete in $\mathcal{H}_L$, the method is called the *Rayleigh–Ritz* method. For a classical treatment of the Rayleigh–Ritz method, the reader should consult [17],[20].

For the more general operators often encountered in electromagnetics, a positive operator can be produced by the following procedure. Consider

$$Lu = f \tag{1.121}$$

Let the *adjoint operator* L^* be defined by

$$\langle Lu, v\rangle \;=\; \langle u, L^*v\rangle \tag{1.122}$$

for $u \in \mathcal{D}_L$, $v \in \mathcal{D}_{L^*}$. Then, if the adjoint L^* exists, operating on both sides of (1.121) with L^* produces

$$L^*Lu = L^*f \tag{1.123}$$

for any $f \in \mathcal{D}_{L*}$. Provided that $Lu = 0$ has none but the trivial solution, it is easy to show that the operator L^*L is positive. Indeed, $\langle L^*Lu, u\rangle = \|Lu\|^2 > 0$, unless $Lu = 0$. But, $Lu = 0$ implies $u = 0$.

The Method of Moments applied to L^*L gives

$$\sum_{k=1}^{n} \alpha_k \langle L^*L\phi_k, w_m\rangle = \langle L^*f, w_m\rangle, \qquad m = 1, 2, \ldots, n \tag{1.124}$$

The Galerkin specialization follows immediately, viz.

$$\sum_{k=1}^{n} \alpha_k \langle L^* L\phi_k, \phi_m \rangle = \langle L^* f, \phi_m \rangle, \qquad m = 1, 2, \ldots, n \tag{1.125}$$

Since L^*L is positive, if the sequence $\{\phi_k\}$ is complete in $\mathcal{D}_{L^*L}$, (1.125) is the Rayleigh–Ritz method and convergence in energy $u_n \xrightarrow{e} u$ is assured, viz.

$$\lim_{n\to\infty} \left| u_n - u \right| = 0 \tag{1.126}$$

where the energy norm is with respect to the operator L^*L. By properties of the adjoint, (1.125) can also be written

$$\sum_{k=1}^{n} \alpha_k \langle L\phi_k, L\phi_m \rangle = \langle f, L\phi_m \rangle, \qquad m = 1, 2, \ldots, n \tag{1.127}$$

which is the result in the *Method of Least Squares*, more usually derived [20] by minimization of

$$\|Lu_n - f\|^2$$

It is easy to show that (1.126) implies that

$$\lim_{n\to\infty} \|Lu_n - f\| = 0 \tag{1.128}$$

so that $Lu_n \to f$. Unless the operator L^*L is positive-definite, nothing can be said concerning the convergence of u_n to u.

A.1 APPENDIX—PROOF OF PROJECTION THEOREM

In this Appendix, we prove the Projection Theorem, considered in Section 1.6. We restate the theorem here for convenience. Let x be any vector in the Hilbert space $\mathcal{H}$, and let $\mathcal{M} \subset \mathcal{H}$ be a *closed* linear manifold. Then, there is a unique vector $y_0 \in \mathcal{M} \subset \mathcal{H}$ closest to x in the sense that

$$\delta = \inf_{y\in\mathcal{M}} \|x - y\| = \|x - y_0\| \tag{A.1}$$

where *inf* is the *greatest lower bound*, or *infimum*. In other words, y_0 is closest to x in the sense that $\|x - y_0\| \leq \|x - y\|$ for all y in $\mathcal{M}$. Furthermore, the necessary and sufficient condition that y_0 is the unique minimizing vector is that $e = x - y_0$ is in $\mathcal{M}^\perp$. The vector y_0 is called the *projection* of x onto $\mathcal{M}$. The vector e is called the projection of x onto $\mathcal{M}^\perp$.

In proving the Projection Theorem, we begin by noting that the first equality in (A.1) makes sense. Indeed, $\|x - y\|$ is bounded below by zero, and therefore has a greatest lower bound. We next show that there exists at least one vector y_0 closest to x. We begin by asserting that there exists a vector $y_n \in \mathcal{M}$ such that by the definition of infimum,

$$\delta \le \|x - y_n\| < \delta + \frac{1}{n} \tag{A.2}$$

Taking the limit as $n \to \infty$, we find that

$$\lim_{n\to\infty} \|x - y_n\| = \delta \tag{A.3}$$

Therefore, we can always define a sequence $\{y_n\} \in \mathcal{M}$ such that $\|x - y_n\|$ converges to δ. In (1.36), if we replace x by $x - y_n$ and y by $x - y_m$, we obtain [21], after some rearrangement,

$$\|y_n - y_m\|^2 = 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\|x - \frac{1}{2}(y_n + y_m)\|^2 \tag{A.4}$$

Since $\mathcal{M}$ is a linear manifold, $(y_n + y_m)/2 \in \mathcal{M}$, and we may assert that

$$\|x - \frac{1}{2}(y_n + y_m)\| \ge \delta$$

Therefore,

$$\|y_n - y_m\|^2 \le 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\delta^2 \tag{A.5}$$

In the limit as $m, n \to \infty$, the right side goes to $2\delta^2 + 2\delta^2 - 4\delta^2 = 0$, and we conclude that the sequence $\{y_n\}$ is Cauchy. Since $\mathcal{H}$ is a Hilbert space and $\mathcal{M}$ is closed, $\mathcal{M}$ is a Hilbert space and Cauchy convergence implies convergence. Therefore, $y_n \to y_0 \in \mathcal{M}$.

We next show that y_0 is unique [22]. Suppose it is not unique. Then, we must have at least two solutions y_0 and $\hat{y}_0$ satisfying $\|x - y_0\| = \|x - \hat{y}_0\| = \delta$. Then,

$$\begin{aligned}\|y_0 - \hat{y}_0\|^2 &= 2\|x - y_0\|^2 + 2\|x - \hat{y}_0\|^2 - 4\|x - \frac{1}{2}(y_0 + \hat{y}_0)\|^2 \\ &\le 2\delta^2 + 2\delta^2 - 4\delta^2 = 0\end{aligned} \tag{A.6}$$

Therefore, $y_0 = \hat{y}_0$.

Finally, we show that $e = x - y_0 \in \mathcal{M}^{\perp}$. We must show that e is orthogonal to every vector in $\mathcal{M}$. Suppose that there exists a vector $z \in \mathcal{M}$ that is not orthogonal to e. Then, we would have [23]

$$\langle e, z\rangle = \langle x - y_0, z\rangle = A \neq 0, \qquad z \in \mathcal{M} \tag{A.7}$$

We define a vector $z_0 \in \mathcal{M}$ such that

$$z_0 = y_0 + \frac{A}{\|z\|^2} z \tag{A.8}$$

Then,

$$\begin{aligned}
\|x - z_0\|^2 &= \langle x - y_0 - \frac{A}{\|z\|^2} z, x - y_0 - \frac{A}{\|z\|^2} z\rangle \\
&= \|x - y_0\|^2 - \frac{\overline{A}}{\|z\|^2}\langle x - y_0, z\rangle - \frac{A}{\|z\|^2}\langle z, x - y_0\rangle + \frac{|A|^2}{\|z\|^2} \\
&= \|x - y_0\|^2 - \frac{|A|^2}{\|z\|^2}
\end{aligned} \tag{A.9}$$

Therefore,

$$\|x - z_0\| < \|x - y_0\| \tag{A.10}$$

which, by (A.1), is impossible.

PROBLEMS

1.1. Using the rules defining a linear space, show that $0a = \mathbf{0}$ and $-1a = -a$.

1.2. Show that $\mathbf{R}_n$ is a linear space.

1.3. Show that $C(0,1)$ is a linear space.

1.4. As an extension to Example 1.4, in $\mathbf{R}_2$, let $x_1 = (1, 3)$, $x_2 = (2, 6.00000001)$. Show that x_1 and x_2 are linearly independent. Comment on what might occur in solving this problem on a computer with eight-digit accuracy. (This problem is indicative of the difficulties that can arise in establishing linear independence in numerical experiments in finite length arithmetic.)

1.5. If $x_1, x_2, \ldots, x_n$ is a linearly dependent set, show that at least one member of the set can be written as a linear combination of the other members.

1.6. Show that if $\mathbf{0}$ is a member of the set $x_1, x_2, \ldots, x_n$, the set is linearly dependent.

1.7. Show that if a set containing n vectors is linearly dependent, and if m additional vectors are added to the set, the resulting set of $n + m$ vectors is linearly dependent.

1.8. Show that in $\mathbf{R}_n$ the set of vectors

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

is linearly independent. Is the same conclusion valid in $\mathbf{C}_n$? Is the set of vectors a basis for $\mathbf{C}_n$?

1.9. Given the basic definition of an inner product space, show that

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

1.10. Show that $C(\alpha, \beta)$ with inner product defined by (1.27) is a real inner product space.

1.11. Consider the linear space of real continuous twice differentiable functions over the interval (0, 1). As a candidate for an inner product for the space, consider

$$\langle f, g\rangle = \int_0^1 f''(\tau)g''(\tau)d\tau$$

where f and g are members of the space and "primes" indicate differentiation. Determine whether $\langle f, g\rangle$ is a legitimate inner product.

1.12. Prove the following corollary to the CSB inequality in (1.35):

$$|\langle x, y\rangle| = \|x\| \, \|y\|$$

if and only if x and y are linearly dependent.

1.13. Prove the following identity:

$$| \, \|x\| - \|y\| \, | \leq \|x - y\|$$

1.14. Consider a complex inner product space with norm induced by the inner product. If x and y are members of the space, prove that

$$\langle x, y\rangle - \langle y, x\rangle = \frac{i}{2}\left(\|x + iy\|^2 - \|x - iy\|^2\right)$$

and

$$\langle x, y\rangle + \langle y, x\rangle = \frac{1}{2}\left(\|x + y\|^2 - \|x - y\|^2\right)$$

1.15. Given the following sequence in the space of rational numbers:

$$x_n = \sum_{k=1}^{n} \frac{1}{(k-1)!}$$

First, show that the sequence is Cauchy; next, show that the sequence does not converge in the space. (Indeed, it converges to e; the details can be found in [4].)

1.16. Show that if $\mathcal{S}$ is a linear space, a linear manifold $\mathcal{M} \subset \mathcal{S}$ is also a linear space.

1.17. Let x_k, $k = 1, 2, \ldots, n$ be a linearly independent sequence of vectors in the Hilbert space $\mathcal{H}$. Define $\mathcal{M}$ to be the set of all linear combinations of the n vectors. Prove that $\mathcal{M}$ is a linear manifold.

1.18. Let $\mathbf{R}_\infty$ be the space of all vectors consisting of a countably infinite set of real numbers (components), viz.

$$a = (\alpha_1, \alpha_2, \ldots)$$

where $\alpha_k \in \mathbf{R}$. Let $\mathcal{M}$ be the set of vectors in $\mathbf{R}_\infty$ with only a finite number of the countably infinite number of components different from zero. Show that $\mathcal{M}$ is a linear manifold. Show that $\mathcal{M}$ is not closed. *Hint*: Consider the concept of *closed* as applied specifically to the sequence of vectors

$$x_n = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, 0, 0, \ldots\right)$$

1.19. The Legendre functions $P_n(\xi)$, $n = 0, 1, 2, \ldots$, are orthogonal on $\xi \in (-1, 1)$, but they are *not* orthonormal. Create a sequence of *orthonormalized* Legendre functions $\hat{P}_n(\xi)$, $n = 0, 1, 2, \ldots$.

1.20. Given that in $\mathbf{R}_3$, $x_1 = (1, 2, 0)$, $x_2 = (0, 1, 2)$, $x_3 = (1, 0, 1)$.

- **(a)** Prove that $\{x_1, x_2, x_3\}$ is a linearly independent set of vectors.
- **(b)** From the linearly independent set, produce *the first two members* of the associate orthonormal set using the Gram–Schmidt procedure.

1.21. Show that the determinant of the Gram matrix in (1.69) is nonzero if and only if the sequence of vectors $\{y_k\}_{k=1}^{M}$ is linearly independent [11].

1.22. Let $\mathbf{R}_\infty$ be the space described in Problem 1.18. If $b = (\beta_1, \beta_2, \ldots)$, define an inner product for the space by

$$\langle a, b\rangle = \sum_{k=1}^{\infty} \alpha_k \beta_k$$

Let the norm for the space be induced by the inner product. We restrict $\mathbf{R}_\infty$ to those vectors with finite norm. Define the operator A in $\mathbf{R}_\infty$ by

$$Aa = \left(\alpha_1, \frac{1}{2}\alpha_2, \frac{1}{3}\alpha_3, \ldots\right)$$

Test the operator A for boundedness.

1.23. On the Hilbert space $\mathcal{L}_2(-\alpha, \alpha)$, with inner product

$$\langle f, g\rangle = \int_{-\alpha}^{\alpha} f(\xi)g(\xi)\frac{d\xi}{\sqrt{\alpha^2 - \xi^2}}$$

consider the following integral equation:

$$\int_{-\alpha}^{\alpha} u(\xi) \ln|\eta - \xi| \frac{d\xi}{\sqrt{\alpha^2 - \xi^2}} = f(\eta)$$

This integral equation occurs in diffraction by a slit in a perfectly conducting screen. It can be shown [14] that the operator

$$L = \int_{-\alpha}^{\alpha} (\,\cdot\,) \ln|\eta - \xi| \frac{d\xi}{\sqrt{\alpha^2 - \xi^2}}$$

is bounded. Solve the integral equation by using the Chebyshev polynomials $T_n(\xi/\alpha)$ as a basis for $\mathcal{L}_2(-\alpha, \alpha)$ and obtaining the matrix representation for L. *Hint*: The following are useful integral relations [15]:

$$\int_{-\alpha}^{\alpha} \frac{T_n(\eta/\alpha) \ln|\xi - \eta| d\eta}{\sqrt{\alpha^2 - \eta^2}} = \begin{cases} -\pi \ln(2/\alpha) T_0(\xi/\alpha), & n = 0 \\ -\dfrac{\pi}{n} T_n(\xi/\alpha), & n > 0 \end{cases}$$

$$\int_{-1}^{1} T_m(\xi) T_n(\xi) \frac{d\xi}{\sqrt{1 - \xi^2}} = \begin{cases} 0, & m \neq n \\ \frac{\pi}{2}, & m = n \neq 0 \\ \pi, & m = n = 0 \end{cases}$$

1.24. Let $L = d/d\xi$, and consider the sequence of partial sums

$$u_n = \sum_{k=1}^{n} \frac{1}{k} \cos k\pi\xi$$

It is well-known [16] that

$$\lim_{n\to\infty} u_n = -\ln\left(2 \sin\frac{\pi\xi}{2}\right)$$

Show that $\lim_{n\to\infty} Lu_n$ is undefined. The problem is that L is unbounded. This result is an example of the fact that a Fourier series cannot always be differentiated term by term.

1.25. Let $L = d/d\xi$, and consider the sequence of partial sums

$$u_n = \sum_{k=1}^{n} \frac{1}{k^2} \cos k\pi\xi$$

Using well-known series summation results [16], show that, although L is unbounded, in this case the operator and limit *can* be interchanged.

1.26. Show that the operator in (1.91) with kernel defined in (1.90) is bounded.

1.27. Consider Hilbert space $\mathcal{L}_2(-1, 1)$ with inner product

$$\langle u, v\rangle = \int_{-1}^{1} u(\xi)v(\xi)d\xi$$

where all functions are real-valued. Consider the following function $f(\xi)$:

$$f(\xi) = \xi - 4\xi^3$$

Construct a function $g(\xi)$ orthogonal to $f(\xi)$. Adjust $g(\xi)$ such that it has unit norm

$$\|g(\xi)\| = 1$$

1.28. Given that, in $\mathbf{R}_3$, $x_1 = (1, 1, 0), x_2 = (0, 1, 1), x_3 = (1, 0, 1)$.

(a) Prove that $\{x_1, x_2, x_3\}$ is a linearly independent set of vectors.

(b) Using the Gram–Schmidt procedure, produce the first two members $\{e_1, e_2\}$ of the orthonormal set obtained from this linearly independent set.

1.29. Given Hilbert space $\mathcal{L}_2(0, 2\pi)$ with inner product

$$\langle u, v\rangle = \int_{0}^{2\pi} u(\xi)\overline{v(\xi)}d\xi$$

where members of the space are complex functions.

(a) Show that the sequence

$$\{e^{in\xi}\}_{n=-\infty}^{\infty}$$

is an orthogonal sequence.

(b) Produce an orthonormal (O.N.) sequence from the orthogonal sequence.

(c) Using the members of the O.N. sequence contained on $-N \le n \le N$, where N is a positive integer, find the best approximation to

$$f(\xi) = \sin\left(\frac{\xi - 2a}{2}\right), \qquad a \in \mathbf{R}$$

in the sense given in Section 1.6, *Best Approximation*.

1.30. Consider the real Hilbert space $\mathcal{L}_2(-1, 1)$. For $f(\xi), g(\xi) \in \mathcal{L}_2(-1, 1)$, define an inner product

$$\langle f, g\rangle = \int_{-1}^{1} f(\xi)g(\xi)\frac{d\xi}{\sqrt{1 - \xi^2}}$$

Determine whether this definition results in a legitimate inner product.

1.31. Consider the real Hilbert space $\mathcal{L}_2(0, 1)$ with inner product

$$\langle f, g \rangle = \int_0^1 f(\xi)g(\xi)d\xi$$

where $f(\xi), g(\xi) \in \mathcal{L}_2(0, 1)$. Suppose that

$$f(\xi) = 1 - \frac{\xi}{2}$$

(a) By the method of best approximation, approximate $f(\xi)$ by $\hat{f}(\xi)$, where $\hat{f}(\xi)$ is a linear combination constructed from the orthonormal set $\{\sqrt{\epsilon_k} \cos k\pi\xi\}_{k=0}^{1}$ in $\mathcal{L}_2(0, 1)$. In the orthonormal set, ϵ_k is 1 for $k = 0$ and 2 for $k \neq 0$.

(b) Calculate the norm of the error $\|f(\xi) - \hat{f}(\xi)\|$.

1.32. Consider Euclidean space $\mathbf{R}_4$. Define vectors a and b in $\mathbf{R}_4$ by

$$a = (\alpha_1, \cdots, \alpha_4)$$
$$b = (\beta_1, \cdots, \beta_4)$$

Define an inner product for the space by

$$\langle a, b \rangle = \sum_{k=1}^{4} \alpha_k \beta_k$$

Consider those vectors a in $\mathbf{R}_4$ restricted by

$$\alpha_1 - \alpha_2 = \alpha_3 - \alpha_4 = 0$$

(a) Show that all vectors with this restriction form a linear manifold $\mathcal{M}$.

(b) Find all vectors b that are members of $\mathcal{M}^{\perp}$.

REFERENCES

[1] Stewart, G.W. (1973), *Introduction to Matrix Computations.* New York: Academic Press, 54–56.

[2] Hardy, G.H. (1967), *A Course in Pure Mathematics.* London: Cambridge University Press, 1–2.

[3] De Lillo, N.J. (1982), *Advanced Calculus with Applications.* New York: Macmillan, 32.

[4] Stakgold, I. (1967), *Boundary Value Problems of Mathematical Physics, Vol. 1.* New York: Macmillan, 101–102.

[5] Mac Duffee, C.C. (1940), *Introduction to Abstract Algebra.* New York: Wiley, chapter VI.

[6] Greenberg, M.D. (1978), *Foundations of Applied Mathematics*. Englewood Cliffs, NJ: Prentice-Hall, 319–320.

[7] Helmberg, G. (1975), *Introduction to Spectral Theory in Hilbert Space*. Amsterdam: North-Holland, 20–21.

[8] Ibid., Appendix B.

[9] Shilov, G.E. (1961), *An Introduction to the Theory of Linear Spaces*. Englewood Cliffs, NJ: Prentice-Hall, sect. 83–85.

[10] Stakgold, I. (1979), *Green's Functions and Boundary Value Problems*. New York: Wiley-Interscience, 36–41.

[11] Luenberger, D.G. (1969), *Optimization by Vector Space Methods*. New York: Wiley, 56–57.

[12] op.cit. Helmberg, 73.

[13] Akhiezer, N.I., and I.M. Glazman (1961), *Theory of Linear Operators in Hilbert Space*. New York: Frederick Ungar, 33,39.

[14] Dudley, D.G. (1985), Error minimization and convergence in numerical methods, *Electromagnetics* 5: 89–97.

[15] Butler, C.M., and D.R. Wilton (1980), General analysis of narrow strips and slots, *IEEE Trans. Antennas Propagat.* AP-28: 42–48.

[16] Abramowitz, M., and I.A. Stegun (Eds.) (1964), *Handbook of Mathematical Functions*, National Bureau of Standards, Applied Mathematics Series, 55, Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402, 1005.

[17] Mikhlin, S.G. (1965), *The Problem of the Minimum of a Quadratic Functional*. San Francisco: Holden-Day.

[18] Harrington, R.F. (1968), *Field Computation by Moment Methods*. New York: Macmillan.

[19] Harrington, R.F. (1967), "Matrix methods for field problems," *Proc. IEEE* 55: 136–149.

[20] op.cit. Stakgold (1967), vol. 2, sect. 8.10.

[21] Krall, A.M. (1973), *Linear Methods of Applied Analysis*. Reading, MA: Addison-Wesley, 181–182.

[22] Ibid., 182.

[23] op.cit. Akhiezer and Glazman, 10.