

# Chapter 2

## Inner Product Spaces

### 2.1 Definition and Examples

In the study of vector algebra in  $\mathbb{R}^n$ , the notion of angle between two nonzero vectors is introduced by considering the inner (or dot) product. In fact, if  $x = (x_1, x_2, \dots, x_n)$  and  $y = (y_1, y_2, \dots, y_n)$  are any two vectors in the  $n$ -dimensional Euclidean space  $\mathbb{R}^n$ , then their inner product is defined by

$$(x, y) = \sum_{i=1}^n x_i y_i,$$

and this inner product is related to the norm by

$$(x, x) = \|x\|^2.$$

The familiar equation

$$(x, y) = \|x\| \|y\| \cos \theta$$

determines the angle  $\theta$  between  $x$  and  $y$ . The vectors  $x$  and  $y$  are orthogonal if  $(x, y) = 0$ . This concept of orthogonality proves useful and lends itself to the generalisation to spaces of higher dimensions.

We introduce below the abstract notion of an inner product and show how a vector space equipped with an inner product reflects properties analogous to those enjoyed by the  $n$ -dimensional Euclidean space  $\mathbb{R}^n$ .

Recall that we denote by  $\mathbb{F}$  either the field  $\mathbb{C}$  of complex numbers or the field  $\mathbb{R}$  of real numbers.

**Definition 2.1.1** Let  $H$  be a vector space over  $\mathbb{F}$ . An **inner product** on  $H$  is a function  $(\cdot, \cdot)$  from  $H \times H$  into  $\mathbb{F}$  such that for all  $x, y, z \in H$  and  $\lambda \in \mathbb{F}$ ,

- (i)  $(x, y) = \overline{(y, x)}$ ,
- (ii)  $(x + z, y) = (x, y) + (z, y)$ ,  $(\lambda x, y) = \lambda(x, y)$ ,
- (iii)  $(x, x) \geq 0$  and  $(x, x) = 0$  if, and only if,  $x = 0$ .

An **inner product space** is a vector space with an inner product on it. Axiom (ii) for an inner product space can be expressed as follows: the inner product is linear in the first variable. In axiom (i),  $\overline{(y, x)}$  denotes the complex conjugate of  $(y, x)$ . Inner product spaces are also called **pre-Hilbert spaces**.

It is left to the reader to verify that when  $\mathbb{F} = \mathbb{R}$  and  $H = \mathbb{R}^n$ , the usual inner product  $(x, y) = \sum_{i=1}^n x_i y_i$  (described above) satisfies the foregoing definition.

The following proposition contains some immediate consequences of Definition 2.1.1.

**Proposition 2.1.2** *For any  $x, y, z$  in an inner product space  $H$  and any  $\lambda \in \mathbb{F}$ , the following hold:*

- (a)  $(x, y + z) = (x, y) + (x, z)$ ;
- (b)  $(x, \lambda y) = \overline{\lambda}(x, y)$ ;
- (c)  $(0, y) = (x, 0) = 0$ ;
- (d)  $(x - y, z) = (x, z) - (y, z)$ ;
- (e)  $(x, y - z) = (x, y) - (x, z)$ ;
- (f) if  $(x, y) = (x, z)$  for all  $x$ , then  $y = z$ .

*Proof* (a) Using Definition 2.1.1(i) and (ii), we have

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

(c)  $(0, y) = (0 + 0, y) = (0, y) + (0, y)$ , on using Definition 2.1.1(ii), and hence  $(0, y) = 0$ .

(f) Suppose  $(x, y) = (x, z)$  for all  $x$ . Then

$$(x, y - z) = (x, y) - (x, z) = 0$$

for all  $x$ ; in particular,  $(y - z, y - z) = 0$  and hence  $y - z = 0$  by Definition 2.1.1(iii).

The proofs of (b), (d) and (e) are no different and are left to the reader.  $\square$

*Examples 2.1.3*

- (i) Let  $H = \mathbb{C}^n = \{x = (x_1, x_2, \dots, x_n): x_i \in \mathbb{C}, 1 \leq i \leq n\}$  be the complex vector space of  $n$ -tuples. For  $x = (x_1, x_2, \dots, x_n)$  and  $y = (y_1, y_2, \dots, y_n)$ , define

$$(x, y) = \sum_{i=1}^n x_i \overline{y_i}. \quad (2.1)$$

It is routine to check that the formula (2.1) does define an inner product on  $\mathbb{C}^n$  in the sense of Definition 2.1.1. This space is called  **$n$ -dimensional unitary space** and

is denoted by  $\mathbb{C}^n$ . Indeed, the vectors  $(1, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, 0, \dots, 0, 1)$  constitute a basis for  $\mathbb{C}^n$ .

- (ii) Let  $\ell_0$  be the vector space of all sequences  $x = \{x_n\}_{n \geq 1}$  of complex numbers, all of whose terms, from some index onwards, are zero (the index, of course, may vary with the sequence). If  $x = \{x_n\}_{n \geq 1}$  and  $y = \{y_n\}_{n \geq 1}$ , define

$$(x, y) = \sum_{n=1}^{\infty} x_n \overline{y_n}. \quad (2.2)$$

Since the sum on the right side of (2.2) is essentially finite, convergence is not an issue here. The axioms of Definition 2.1.1 are easily verified.

- (iii) Let  $\ell^2$  denote the set of all complex sequences  $x = \{x_n\}_{n \geq 1}$  which are square summable, that is,

$$\sum_{n=1}^{\infty} |x_n|^2 < \infty.$$

The addition of vectors  $x = \{x_n\}_{n \geq 1}$  and  $y = \{y_n\}_{n \geq 1}$  and the scalar multiplication of  $x = \{x_n\}_{n \geq 1}$  by a scalar  $\lambda \in \mathbb{C}$  are defined by

$$x + y = \{x_n + y_n\}_{n \geq 1} \quad \text{and} \quad \lambda x = \{\lambda x_n\}_{n \geq 1}.$$

Since

$$|\alpha + \beta|^2 \leq 2|\alpha|^2 + 2|\beta|^2$$

for  $\alpha, \beta \in \mathbb{C}$ , it follows that

$$\sum_{n=1}^m |x_n + y_n|^2 \leq 2 \sum_{n=1}^m |x_n|^2 + 2 \sum_{n=1}^m |y_n|^2 \leq 2 \sum_{n=1}^{\infty} |x_n|^2 + 2 \sum_{n=1}^{\infty} |y_n|^2,$$

and hence

$$\sum_{n=1}^{\infty} |x_n + y_n|^2 \leq 2 \sum_{n=1}^{\infty} |x_n|^2 + 2 \sum_{n=1}^{\infty} |y_n|^2. \quad (2.3)$$

Thus, if  $x = \{x_n\}_{n \geq 1}$  and  $y = \{y_n\}_{n \geq 1}$  are in  $\ell^2$ , it follows from (2.3) that  $x + y \in \ell^2$ . Also, if  $x \in \ell^2$  and  $\lambda \in \mathbb{C}$ , then  $\sum_{n=1}^{\infty} |\lambda x_n|^2 = |\lambda|^2 \sum_{n=1}^{\infty} |x_n|^2$  shows that  $\lambda x \in \ell^2$ . Consequently,  $\ell^2$  is a vector space over  $\mathbb{C}$ .

For  $x = \{x_n\}_{n \geq 1}$  and  $y = \{y_n\}_{n \geq 1}$  in  $\ell^2$ , the series  $\sum_{n=1}^{\infty} x_n \overline{y_n}$  converges absolutely. In fact,

$$|x_n \bar{y}_n| \leq \frac{1}{2} (|x_n|^2 + |y_n|^2)$$

implies

$$\sum_{n=1}^m |x_n \bar{y}_n| \leq \frac{1}{2} \sum_{n=1}^m |x_n|^2 + \frac{1}{2} \sum_{n=1}^m |y_n|^2 \leq \frac{1}{2} \sum_{n=1}^{\infty} |x_n|^2 + \frac{1}{2} \sum_{n=1}^{\infty} |y_n|^2,$$

and hence

$$\sum_{n=1}^{\infty} |x_n \bar{y}_n| \leq \frac{1}{2} \sum_{n=1}^{\infty} |x_n|^2 + \frac{1}{2} \sum_{n=1}^{\infty} |y_n|^2.$$

Now define

$$(x, y) = \sum_{n=1}^{\infty} x_n \bar{y}_n, \quad x, y \in \ell^2. \quad (2.4)$$

It is now easy to check that the axioms for an inner product are satisfied. Thus,  $\ell^2$  with the inner product defined in (2.4) is an inner product space.

- (iv) Let  $C[a, b]$ ,  $-\infty < a < b < \infty$ , be the vector space of all continuous complex-valued functions defined on  $[a, b]$ . Define

$$(f, g) = \int_a^b f(t) \overline{g(t)} dt, \quad f, g \in C[a, b]. \quad (2.5)$$

Observe that  $(f, f) = \int_a^b |f(t)|^2 dt = 0$  implies  $f(t) = 0$  for each  $t \in [a, b]$ , in view of the continuity of  $f$ . The other axioms in Definition 2.1.1 are consequences of the properties of integrals.

- (v) Let  $C^n[a, b]$  be the vector space of all  $n$  times continuously differentiable complex-valued functions defined on  $[a, b]$ . For  $f, g \in C^n[a, b]$ , define

$$(f, g) = \sum_{i=0}^n \int_a^b f^{(i)}(t) \overline{g^{(i)}(t)} dt, \quad (2.6)$$

where  $f^{(i)}(t)$  denotes the  $i$ th derivative of  $f$ ,  $1 \leq i \leq n$ , and  $f^{(0)}(t) = f(t)$ ,  $t \in [a, b]$ . Observe that  $0 = (f, f) = \sum_{i=0}^n \int_a^b |f^{(i)}(t)|^2 dt = 0$  implies  $\sum_{i=0}^n |f^{(i)}(t)|^2 = 0$ ,  $t \in [a, b]$ , in view of the continuity of  $\sum_{i=0}^n |f^{(i)}(t)|^2$ ,  $t \in [a, b]$ . This implies  $f(t) = 0$  for each  $t \in [a, b]$ . The other axioms in Definition 2.1.1 are consequences of the properties of integrals.

- (vi) Let  $RL^2$  denote the space of rational functions (i.e. a ratio of two polynomials with complex coefficients) which are analytic on the unit circle

$$\partial D = \{z \in \mathbb{C} : |z| = 1\}$$

with the usual pointwise addition and scalar multiplication. The inner product is defined by

$$(f, g) = \frac{1}{2\pi i} \int_{\partial D} f(z) \overline{g(z)} \frac{dz}{z}, \quad (2.7)$$

where the integral is being taken in the anticlockwise direction around  $\partial D$ .

$RH^2$  is the subspace of  $RL^2$  consisting of those rational functions which are analytic on the closed unit disc  $\overline{D}$ , where

$$D = \{z \in \mathbb{C} : |z| < 1\},$$

with inner product given by (2.7).

Thus, a rational function belongs to  $RL^2$  if it has no pole of absolute value 1, and it belongs to  $RH^2$  if it has no pole of absolute value less than or equal to 1.

Clearly, (2.7) satisfies the axioms in (i), (ii) and part of (iii). We need to check that  $(f, f) > 0$  when  $f \neq 0$ . Indeed,

$$(f, f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(e^{i\theta})|^2 d\theta, \quad (2.8)$$

using the parametrisation  $z = e^{i\theta}$ ,  $-\pi < \theta \leq \pi$ . Since  $f(e^{i\theta})$  is continuous on  $[-\pi, \pi]$ , the right-hand side of (2.8) is positive unless  $f = 0$ .

- (vii) A **trigonometric polynomial** is a finite sum of the form

$$f(x) = a_0 + \sum_{n=1}^k a_n e^{i\lambda_n x}, \quad x \in [-\pi, \pi],$$

where  $k \in \mathbb{N}$ ,  $a_0, a_1, \dots, a_k \in \mathbb{C}$  and  $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{N}$ . It is clear that every trigonometric polynomial is of period  $2\pi$ . The space  $TP$  of trigonometric polynomials is a vector space over  $\mathbb{C}$  with respect to pointwise addition and scalar multiplication. If we define the inner product by

$$(f, g) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \overline{g(t)} dt, \quad (2.9)$$

then  $TP$  becomes an inner product space.

**Problem Set 2.1**

2.1.P1. For which values of  $\alpha \in \mathbb{C}$  does the sequence  $\{n^{-\alpha}\}_{n \geq 1}$  belong to  $\ell^2$ ?

2.1.P2. [Notations as in Example 2.1.3(vi)] Calculate the inner product of functions

$$f(z) = \frac{1}{z - \alpha} \quad \text{and} \quad g(z) = \frac{1}{z - \beta}, \quad \text{where} \quad |\alpha| < 1, |\beta| < 1.$$

2.1.P3. [Notations as in Example 2.1.3(vi)] Let  $k_\alpha \in RH^2$  be defined by  $k_\alpha(z) = (1 - \bar{\alpha}z)^{-1}$ , where  $|\alpha| \neq 1$ . Show that for  $f \in RH^2$ ,

$$(f, k_\alpha) = \begin{cases} f(\alpha) & \text{if } |\alpha| < 1 \\ 0 & \text{if } |\alpha| > 1 \end{cases}$$

2.1.P4. [Notations as in Example 2.1.3(vi)] Let  $f \in RH^2$  and  $\alpha \in D$ . Show that

$$|f(\alpha)| \leq \frac{1}{\sqrt{1 - |\alpha|^2}} \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(e^{i\theta})|^2 d\theta \right\}^{\frac{1}{2}}.$$

**2.2 Norm of a Vector**

Let  $X$  be a vector (linear) space over  $\mathbb{F}$ .

**Definition 2.2.1** A **norm**  $\|\cdot\|$  is a function from  $X$  into the nonnegative reals  $\mathbb{R}^+$  satisfying

- (i)  $\|x\| = 0$  if, and only if,  $x = 0$ ,
- (ii)  $\|\lambda x\| = |\lambda| \cdot \|x\|$  for each  $\lambda \in \mathbb{F}$  and  $x \in X$ ,
- (iii)  $\|x + y\| \leq \|x\| + \|y\|$  for all  $x, y \in X$ . [**triangle inequality**]

We emphasise that, by definition,  $\|x\| \geq 0$  for all  $x \in X$ .

If  $X$  is a linear space and  $\|\cdot\|$  is a norm defined on  $X$ , then  $d(x, y) = \|x - y\|$  indeed gives rise to a metric as a consequence of the foregoing Definition 2.2.1. The details are as follows.

That the distance  $d(x, y)$  from a vector  $x$  to a vector  $y$  in  $H$  is strictly positive (that is,  $d(x, y) \geq 0$  and equality holds if, and only if,  $x = y$ ) follows from (i). The fact that  $d(x, y) = d(y, x)$  follows from

$$\|x - y\| = \|(y - x)\| = |-1|\|y - x\| = \|y - x\|,$$

in view of (ii). Also

$$d(x, z) = \|x - z\| = \|x - y + y - z\| \leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z)$$

for all  $x, y$  and  $z$ . The reader will observe that (iii) has been used in proving the preceding inequality.

A linear space  $X$  equipped with a norm  $\|\cdot\|$  is called a **normed linear space**. If the metric space  $(X, d)$ , where  $d(x, y) = \|x - y\|$ ,  $x, y \in X$ , is complete, then the normed linear space is said to be **complete** and is called a **Banach space**. These spaces are named after the great Polish mathematician Stefan Banach.  $\mathbb{R}^n$ , the real space of  $n$ -tuples  $x = (x_1, x_2, \dots, x_n)$  with each of the norms  $\|x\|_1 = \sum_{i=1}^n |x_i|$ ,  $\|x\|_2 =$

$\left(\sum_{i=1}^n |x_i|^2\right)^{\frac{1}{2}}$ ,  $\|x\|_\infty = \sup_i |x_i|$  is a Banach space. That  $\|x\|_1$ ,  $\|x\|_2$  and  $\|x\|_\infty$  are norms can be verified, see [30]. So is the space  $\mathbb{C}^n$  of complex  $n$ -tuples. The space  $(\mathbb{C}^n, \|\cdot\|_2)$  is complete [see Example 2.3.4(i)]. That  $\mathbb{C}^n$  with  $\|\cdot\|_1$  and  $\mathbb{C}^n$  with  $\|\cdot\|_\infty$  are complete follows from the inequalities  $\|\cdot\|_\infty \leq \|\cdot\|_2 \leq \|\cdot\|_1 \leq n\|\cdot\|_\infty$ , see [30].

Hilbert spaces are Banach spaces whose norms are derived from an inner product as detailed below.

**Definition 2.2.2** In an inner product space  $H$ , the **norm** (or **length**) of a vector  $x \in H$ , denoted by  $\|x\|$ , is the nonnegative real number as defined by

$$\|x\| = \sqrt{(x, x)},$$

and is called the norm **induced by** the inner product on  $H$ .

We shall see below that this satisfies the conditions for being a norm as laid out in Definition 2.2.1.

The norm of an element  $x = (\alpha_1, \alpha_2, \dots, \alpha_n)$  in the unitary space  $\mathbb{C}^n$  is

$$\|x\| = \left(\sum_{i=1}^n |\alpha_i|^2\right)^{\frac{1}{2}},$$

and that of an  $x = \{\alpha_i\}_{i \geq 1}$  in  $\ell^2$  is

$$\|x\| = \left(\sum_{i=1}^{\infty} |\alpha_i|^2\right)^{\frac{1}{2}}.$$

The norm of an element  $f \in C[a, b]$  [respectively,  $f \in C^n[a, b]$ ] is

$$\|f\| = \left( \int_a^b |f(t)|^2 dt \right)^{\frac{1}{2}} \left[ \text{resp.} \left( \sum_{i=0}^n \int_a^b |f^{(i)}(t)|^2 dt \right)^{\frac{1}{2}} \right].$$

The norm of an element  $f$  in  $RL^2$  or in  $RH^2$  is

$$\|f\| = \left( \frac{1}{2\pi i} \int_{\partial D} |f(z)|^2 \frac{dz}{z} \right)^{\frac{1}{2}} = \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(e^{i\theta})|^2 d\theta \right)^{\frac{1}{2}}.$$

**Proposition 2.2.3** *In an inner product space  $H$ ,  $\|\cdot\|$  has the following properties: for  $x, y \in H$  and  $\lambda \in \mathbb{F}$ ,*

- (a)  $\|x\| \geq 0$  and  $\|x\| = 0$  if, and only if,  $x = 0$ ;
- (b)  $\|\lambda x\| = |\lambda| \|x\|$ ;
- (c) **(Parallelogram Law)**

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

- (d) **(Polarisation Identity in case  $\mathbb{F} = \mathbb{C}$ )**

$$4(x, y) = \|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2.$$

*Proof* (a) is immediate from Definition 2.1.1(c) while (b) follows from

$$\|\lambda x\|^2 = (\lambda x, \lambda x) = |\lambda|^2 (x, x) = |\lambda|^2 \|x\|^2.$$

For  $x, y \in H$ , we have

$$\|x + y\|^2 = (x + y, x + y) = \|x\|^2 + \|y\|^2 + (x, y) + (y, x). \quad (2.10)$$

In the identity (2.10), replace  $y$  by  $-y$  to obtain

$$\|x - y\|^2 = (x - y, x - y) = \|x\|^2 + \|y\|^2 - (x, y) - (y, x). \quad (2.11)$$

Adding (2.10) and (2.11), we get

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

This proves the Parallelogram Law.

In the identity (2.10), replace  $y$  by  $-y$ ,  $iy$  and  $-iy$ :

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - (x, y) - (y, x). \quad (2.12)$$

$$\|x + iy\|^2 = \|x\|^2 + \|y\|^2 - i(x, y) + i(y, x). \quad (2.13)$$

$$\|x - iy\|^2 = \|x\|^2 + \|y\|^2 + i(x, y) - i(y, x). \quad (2.14)$$

Multiply both sides of (2.12) by  $-1$ , (2.13) by  $i$  and (2.14) by  $-i$  and add to (2.10) to obtain the following:

$$\|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2 = 4(x, y).$$

This completes the proof of the Polarisation Identity.  $\square$

*Remark* In Proposition 2.2.3, the assertions (a)–(c) are valid in real as well as complex inner product spaces, but (d) holds only in a complex inner product space.

**Theorem 2.2.4** (Cauchy–Schwarz Inequality) *Let  $H$  be an inner product space and let  $\|x\|$  denote the norm of  $x \in H$ . Then*

$$|(x, y)| \leq \|x\| \|y\| \quad (2.15)$$

for  $x, y \in H$ , and equality holds if, and only if,  $x$  and  $y$  are linearly dependent.

*Proof* Choose a real number  $\theta$  such that  $e^{i\theta}(x, y) = |(x, y)|$ . Let  $\lambda = \alpha e^{i\theta}$ , where  $\alpha \in \mathbb{R}$ . Then

$$(x - \bar{\lambda}y, x - \bar{\lambda}y) = \|x\|^2 + \|\bar{\lambda}y\|^2 - \lambda(x, y) - \bar{\lambda}(y, x). \quad (2.16)$$

The expression on the left side of (2.16) is real and nonnegative. Hence,

$$\|x\|^2 + \alpha^2 \|y\|^2 - 2\alpha |(x, y)| \geq 0 \quad (2.17)$$

for every real  $\alpha$ . If  $\|y\| = 0$ , then we must have  $|(x, y)| = 0$ , for otherwise (2.17) will be false for large positive values of  $\alpha$ . If  $\|y\| > 0$ , take  $\alpha = |(x, y)|/\|y\|^2$  in (2.17) and obtain

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

If  $x$  and  $y$  are linearly dependent, then we may write  $y = kx$  or  $x = ky$  for some  $k \in \mathbb{F}$ . Then

$$\begin{aligned} |(x, y)| &= |(x, kx)| = |k| \|x\|^2 \\ &= \|kx\| \|x\| = \|y\| \|x\|, \end{aligned}$$

that is, equality holds in (2.15).

On the other hand, suppose that  $|(x, y)| = \|x\|\|y\|$ . If  $\|y\| = 0$ , then  $y = 0$  and  $x$  and  $y$  are linearly dependent. If  $\|y\| \neq 0$ , then

$$\begin{aligned} \left( x - \frac{(x, y)}{\|y\|^2} y, x - \frac{(x, y)}{\|y\|^2} y \right) &= \|x\|^2 + \frac{|(x, y)|^2}{\|y\|^2} - 2\Re \left( x, \frac{(x, y)}{\|y\|^2} y \right) \\ &= \|x\|^2 + \frac{|(x, y)|^2}{\|y\|^2} - 2 \frac{|(x, y)|^2}{\|y\|^2} \\ &= \|x\|^2 - \frac{|(x, y)|^2}{\|y\|^2}. \end{aligned}$$

Hence,  $x - \frac{(x, y)}{\|y\|^2} y = 0$ ; that is,  $x$  and  $y$  are linearly dependent.  $\square$

*Remark* The above proof of the Cauchy–Schwarz Inequality is valid in the real case as well.

Applying Theorem 2.2.4 in specific spaces such as  $\mathbb{C}^n$ ,  $\ell^2$  and  $C[a, b]$ , the following corollary results.

**Corollary 2.2.5**

(a) If  $x_1, x_2, \dots, x_n$  and  $y_1, y_2, \dots, y_n$  are complex numbers,

$$\left| \sum_{i=1}^n \alpha_i \bar{\beta}_i \right| \leq \left( \sum_{i=1}^n |\alpha_i|^2 \right)^{\frac{1}{2}} \left( \sum_{i=1}^n |\beta_i|^2 \right)^{\frac{1}{2}}.$$

(b) If  $\{\alpha_i\}_{i \geq 1}$  and  $\{\beta_i\}_{i \geq 1}$  are square summable sequences of complex numbers,

$$\left| \sum_{i=1}^{\infty} \alpha_i \bar{\beta}_i \right| \leq \left( \sum_{i=1}^{\infty} |\alpha_i|^2 \right)^{\frac{1}{2}} \left( \sum_{i=1}^{\infty} |\beta_i|^2 \right)^{\frac{1}{2}}.$$

(c) If  $f, g \in C[a, b]$ , then

$$\left| \int_a^b f(t) \overline{g(t)} dt \right| \leq \left( \int_a^b |f(t)|^2 dt \right)^{\frac{1}{2}} \left( \int_a^b |g(t)|^2 dt \right)^{\frac{1}{2}}.$$

*In each case, equality holds if, and only if, the vectors involved are linearly dependent.*

**Theorem 2.2.6** (Triangle inequality) *In an inner product space  $H$ ,*

$$\|x + y\| \leq \|x\| + \|y\| \quad (2.18)$$

*for all  $x, y \in H$ .*

*Proof* For  $x, y \in H$ ,

$$\begin{aligned} \|x + y\|^2 &= (x + y, x + y) \\ &= \|x\|^2 + \|y\|^2 + (x, y) + (y, x) \\ &= \|x\|^2 + \|y\|^2 + 2\Re(x, y) \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|, \end{aligned}$$

using the Cauchy–Schwarz Inequality (2.15). Thus,

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

which implies

$$\|x + y\| \leq \|x\| + \|y\|.$$

□

Applying Theorem 2.2.6 to specific inner product spaces, such as  $\ell^2$ ,  $RL^2$  and  $C[a, b]$ , the following inequalities are obtained:

**Corollary 2.2.7**

(a) *If  $\{x_i\}_{i \geq 1}$  and  $\{y_i\}_{i \geq 1}$  are in  $\ell^2$ , then*

$$\left( \sum_{i=1}^{\infty} |x_i + y_i|^2 \right)^{\frac{1}{2}} \leq \left( \sum_{i=1}^{\infty} |x_i|^2 \right)^{\frac{1}{2}} + \left( \sum_{i=1}^{\infty} |y_i|^2 \right)^{\frac{1}{2}}.$$

(b) *If  $f, g$  are in  $RL^2$ , then*

$$\left( \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(e^{i\theta}) + g(e^{i\theta})|^2 d\theta \right)^{\frac{1}{2}} \leq \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(e^{i\theta})|^2 d\theta \right)^{\frac{1}{2}} + \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} |g(e^{i\theta})|^2 d\theta \right)^{\frac{1}{2}}.$$

(c) If  $f, g$  are in  $C[a, b]$ , then

$$\left( \int_a^b |f(t) + g(t)|^2 dt \right)^{\frac{1}{2}} \leq \left( \int_a^b |f(t)|^2 dt \right)^{\frac{1}{2}} + \left( \int_a^b |g(t)|^2 dt \right)^{\frac{1}{2}}.$$

**Corollary 2.2.8** In an inner product space  $H$ ,

$$||x|| - ||y|| \leq ||x - y|| \quad (2.19)$$

for all  $x, y \in H$ .

*Proof* For  $x, y \in H$ ,

$$||x|| = ||x - y + y|| \leq ||x - y|| + ||y||$$

by Theorem 2.2.6. This implies

$$||x|| - ||y|| \leq ||x - y||. \quad (2.20)$$

On interchanging the roles of  $x$  and  $y$ , we get

$$||y|| - ||x|| \leq ||y - x|| = ||x - y||. \quad (2.21)$$

The inequality (2.19) follows upon combining (2.20) and (2.21).  $\square$

### Problem Set 2.2

2.2.P1. Show that for  $x, y$  and  $z \in X$ , an inner product space, the following Apollonius Identity holds:

$$||x - z||^2 + ||y - z||^2 = \frac{1}{2} ||x - y||^2 + 2 \left\| z - \frac{1}{2}(x + y) \right\|^2.$$

2.2.P2. Show that the formula  $(A, B) = \text{trace}(B^*A)$  defines an inner product on the space  $\mathbb{C}^{n \times n}$  of  $n \times n$  complex matrices, where  $n \in \mathbb{N}$  and  $B^*$  denotes the conjugate transpose of  $B$ . Determine  $||I_n||$ , where  $I_n$  is the identity matrix.

2.2.P3. Show that (i)  $x = \left\{ \frac{1}{n} \right\}_{n=1}^\infty \in \ell^2$  and determine  $||x||$ ; (ii)  $x = \left\{ 2^{-n/2} \right\}_{n=1}^\infty \in \ell^2$  and determine  $||x||$ .

2.2.P4. Prove that, for any function  $f \in C[0, \pi]$ ,

$$\left| \int_0^\pi f(t) \sin t \, dt \right| \leq \sqrt{\frac{\pi}{2}} \left[ \int_0^\pi |f(t)|^2 \, dt \right]^{\frac{1}{2}},$$

and describe the nonzero functions for which equality holds.

2.2.P5. Suppose  $\mu(X) = 1$  and  $f, g$  are positive measurable functions on  $X$  such that  $fg \geq 1$ . Prove that  $\int_X f \, d\mu \int_X g \, d\mu \geq 1$ .

2.2.P6. Let  $X_1 = (C[0, 1], \|\cdot\|_\infty)$ , where  $\|x\|_\infty = \sup_{0 \leq t \leq 1} |x(t)|$ , and let  $X_2 = (C[0, 1], \|\cdot\|_2)$ , where

$$\|x\|_2 = (x, x)^{\frac{1}{2}}, \quad (x, y) = \int_0^1 x(t) \overline{y(t)} \, dt.$$

Show that the identity mapping  $id: X_1 \rightarrow X_2$  is continuous, but its inverse  $id: X_2 \rightarrow X_1$  is not.

2.2.P7. Let  $X$  be an inner product space. If  $\|x\| = \|y\| = \frac{1}{2} \|x + y\|$ , then show that  $x = y$ . The result fails to hold if  $X = \mathbb{R}^n$  or  $\mathbb{C}^n$  with norm  $\|\cdot\|_1$  when  $n > 1$ .

2.2.P8. Let  $X$  be a vector space over  $\mathbb{C}$ . Let  $(\cdot, \cdot)$  be a complex-valued function of two variables  $(x, y): X \times X \rightarrow \mathbb{C}$  which has the following properties:

- (a)  $(\alpha x_1 + \beta x_2, y) = \alpha(x_1, y) + \beta(x_2, y)$
- (b)  $(x, y) = \overline{(y, x)}$
- (c)  $(x, x) \geq 0$  and  $(x, x)$  may be zero for nonzero  $x$ .

Prove that the Cauchy–Schwarz Inequality still holds but without the rider about when equality holds.

2.2.P9. [Notations as in Example 2.1.3(vi)] Let  $g(z) = \frac{1}{(z-\alpha)(z-\beta)}$ , where  $\alpha$  and  $\beta$  are distinct points in  $D$ . Using the Residue Theorem, show that

$$\|g\| = \frac{(1 - |\alpha\beta|^2)^{\frac{1}{2}}}{(1 - |\alpha|^2)^{\frac{1}{2}}(1 - |\beta|^2)^{\frac{1}{2}}|1 - \overline{\alpha}\beta|}.$$

2.2.P10. [Notations as in Example 2.1.3(vi)] Prove that for  $\alpha \in D$ ,

$$F = \{f \in RH^2 : f(\alpha) = 0\}$$

is a closed linear subspace of  $RH^2$ .

## 2.3 Inner Product Spaces as Metric Spaces

We have seen in Proposition 2.2.3 and Theorem 2.2.6 that the norm induced by an inner product in  $H$  satisfies the following conditions of Definition 2.2.1:

- (i)  $\|x\| \geq 0$  and  $\|x\| = 0$  if, and only if,  $x = 0$ ,
- (ii)  $\|\lambda x\| = |\lambda|\|x\|$  for all  $\lambda \in \mathbb{C}$  and  $x \in H$ ,
- (iii)  $\|x + y\| \leq \|x\| + \|y\|$  for all  $x, y \in H$ .

*Remark 2.3.1* The inner product in  $H$  together with the metric  $d(x, y) = \|x - y\|$  from the norm induced by the inner product is a metric space. As in any metric space,  $d(x, y)$  is called the distance from  $x$  to  $y$  or between  $x$  and  $y$ .

We shall henceforth feel free to use, for inner product spaces, all metric concepts such as open and closed sets, convergence, continuity, uniform continuity, Cauchy sequence, completeness, dense sets and separability. Below we translate the general metric space concepts defined in Sect. 1.2 into inner product space terms.

It follows from (2.19) that the map  $x \rightarrow \|x\|$  defined in  $H$  is continuous. In fact, it is uniformly continuous in view of (2.19).

*Remarks 2.3.2*

- (i) A sequence  $\{x_n\}_{n \geq 1}$  in a normed space or in an inner product space is Cauchy if for every  $\varepsilon > 0$ , there exists an integer  $n_0$  such that

$$\|x_n - x_m\| < \varepsilon \text{ whenever } n, m \geq n_0.$$

- (ii) Every convergent sequence is Cauchy; the converse is, however, not true; in fact, let  $\{x_n\}_{n \geq 1}$ , where  $x_n = (1, \frac{1}{2}, \dots, \frac{1}{n}, 0, \dots)$ ,  $n = 1, 2, \dots$  be a sequence in the inner product space  $\ell_0$  (see Example 2.1.3(ii)). Then the sequence  $\{x_n\}_{n \geq 1}$  is Cauchy because

$$\|x_{m+p} - x_m\| = \left( \sum_{k=m+1}^{m+p} \frac{1}{k^2} \right)^{\frac{1}{2}}$$

can be made arbitrarily small by choosing  $m$  sufficiently large. However, the sequence does not converge to an element of the space. Assume the contrary, that is suppose  $x_n \rightarrow x$ , where  $x = (\lambda_1, \lambda_2, \dots, \lambda_N, 0, 0, \dots)$ . If  $n \geq N$ , then

$$\begin{aligned} \|x_n - x\|^2 &= \sum_{k=1}^n \left| \frac{1}{k} - \lambda_k \right|^2 + \sum_{k=n+1}^{\infty} |\lambda_k|^2 \\ &= \sum_{k=1}^n \left| \frac{1}{k} - \lambda_k \right|^2. \end{aligned}$$

On letting  $n \rightarrow \infty$ , we obtain  $\sum_{k=1}^{\infty} \left| \frac{1}{k} - \lambda_k \right|^2 = 0$ , which implies  $\lambda_k = \frac{1}{k}$  for each  $k$ , contradicting the fact that  $x$  has finitely many nonzero terms.

- (iii) The open [respectively, closed] ball with centre  $x_0$  and radius  $\varepsilon$  is the set  $\{x \in H: \|x - x_0\| < \varepsilon\}$  [respectively,  $\{x \in H: \|x - x_0\| \leq \varepsilon\}$ ]. In view of Proposition 1.2.10, a subset  $K \subseteq H$  is bounded if, and only if, there exists an  $M > 0$  such that  $K \subseteq \{x: \|x\| \leq M\}$ .

**Definition 2.3.3** An inner product space  $H$  is said to be **complete** if every Cauchy sequence in  $H$  converges. That is, if  $\{x_n\}_{n \geq 1}$  is a sequence in  $H$  satisfying  $\|x_n - x_m\| \rightarrow 0$  as  $n, m \rightarrow \infty$ , there exists an  $x \in H$  such that  $\|x_n - x\| \rightarrow 0$  as  $n \rightarrow \infty$ . An inner product space which is complete is called a **Hilbert space**.

Every Hilbert space is a Banach space. The norm in a Hilbert space is derived from the inner product.

#### Examples 2.3.4

- (i) The inner product space  $H = \mathbb{C}^n$  with the metric given by

$$d(x, y) = \|x - y\| = \left( \sum_{i=1}^n |x_i - y_i|^2 \right)^{\frac{1}{2}}, \quad (2.22)$$

where  $x = (x_1, x_2, \dots, x_n)$  and  $y = (y_1, y_2, \dots, y_n)$  are in  $\mathbb{C}^n$ , is a Hilbert space with metric as above or with inner product as in (i) of Examples 2.1.3.

We need to check that  $\mathbb{C}^n$ , with the metric defined in (2.22), is complete [see (i) of Examples 2.1.3]. Let  $\{x^{(m)}\}_{m \geq 1} = (x_1^{(m)}, x_2^{(m)}, \dots, x_n^{(m)})$  denote a Cauchy sequence in  $\mathbb{C}^n$ , i.e.  $d(x^{(m)}, x^{(m')}) \rightarrow 0$  as  $m, m' \rightarrow \infty$ . Then for a given  $\varepsilon > 0$  there exists an integer  $n_0(\varepsilon)$  such that

$$\left( \sum_{k=1}^n \left| x_k^{(m)} - x_k^{(m')} \right|^2 \right)^{\frac{1}{2}} < \varepsilon \quad \text{for all } m, m' \geq n_0(\varepsilon). \quad (2.23)$$

Hence  $\left| x_k^{(m)} - x_k^{(m')} \right| < \varepsilon$  for all  $m, m' \geq n_0(\varepsilon)$  and all  $k = 1, 2, \dots, n$ . Upon fixing  $k$  and using the Cauchy Principle of Convergence, it follows that  $\{x_k^{(m)}\}_{m \geq 1}$  converges to a limit  $x_k$ . Let  $x = (x_1, x_2, \dots, x_n)$  and  $m \geq n_0(\varepsilon)$ . It follows from (2.23) that

$$\sum_{k=1}^n \left| x_k^{(m)} - x_k^{(m')} \right|^2 < \varepsilon^2 \quad (2.24)$$

for all  $m' \geq n_0(\varepsilon)$ . Letting  $m' \rightarrow \infty$  in (2.24), we have

$$\sum_{k=1}^n \left| x_k^{(m)} - x_k \right|^2 \leq \varepsilon^2$$

for all  $m \geq n_0(\varepsilon)$ . That is,  $d(x^{(m)}, x) \rightarrow 0$  in  $\mathbb{C}^n$ .

(ii) The inner product space  $H = \ell^2$  (see Example 2.1.3(iii)) is a Hilbert space.

We shall show that  $\ell^2$  with the metric

$$d(x, y) = \|x - y\| = \left( \sum_{k=1}^{\infty} |x_k - y_k|^2 \right)^{\frac{1}{2}} \quad (2.25)$$

is complete. Let  $\{x^{(m)}\}_{m \geq 1} = (x_1^{(m)}, x_2^{(m)}, \dots)$  denote a Cauchy sequence in  $\ell^2$ . Then for a given  $\varepsilon > 0$  there exists an integer  $n_0(\varepsilon)$  such that

$$\left( \sum_{k=1}^{\infty} \left| x_k^{(m)} - x_k^{(m')} \right|^2 \right)^{\frac{1}{2}} < \varepsilon \quad \text{for all } m, m' \geq n_0(\varepsilon). \quad (2.26)$$

This implies  $\left| x_k^{(m)} - x_k^{(m')} \right| < \varepsilon$  for all  $m, m' \geq n_0(\varepsilon)$ , i.e. for each  $k$ , the sequence  $\{x_k^{(m)}\}_{m \geq 1}$  is a Cauchy sequence of complex numbers. So by the Cauchy Principle of Convergence,  $\lim_{m \rightarrow \infty} x_k^{(m)} = x_k$ , say. Let  $x$  be the sequence  $(x_1, x_2, \dots)$ . It will be shown that  $x \in \ell^2$  and  $\lim_{m \rightarrow \infty} x^{(m)} = x$ . From (2.26), we have

$$\sum_{k=1}^N \left| x_k^{(m)} - x_k^{(m')} \right|^2 < \varepsilon^2 \quad (2.27)$$

for any positive integer  $N$ , provided  $m, m' \geq n_0(\varepsilon)$ . Letting  $m' \rightarrow \infty$  in (2.27), we obtain

$$\sum_{k=1}^N \left| x_k^{(m)} - x_k \right|^2 \leq \varepsilon^2$$

for any positive integer  $N$  and all  $m \geq n_0(\varepsilon)$ . The sequence  $\left\{ \sum_{k=1}^N \left| x_k^{(m)} - x_k \right|^2 \right\}_{N \geq 1}$  is a monotonically increasing sequence of nonnegative real numbers and is bounded above and, therefore, has a finite limit  $\sum_{k=1}^{\infty} \left| x_k^{(m)} - x_k \right|^2$  which is less than or equal to  $\varepsilon^2$ . Hence

$$\left( \sum_{k=1}^{\infty} |x_k^{(m)} - x_k| \right)^{\frac{1}{2}} \leq \varepsilon \quad \text{for all } m, m' \geq n_0(\varepsilon). \quad (2.28)$$

Observe that

$$\left( \sum_{k=1}^{\infty} |x_k|^2 \right)^{\frac{1}{2}} \leq \left( \sum_{k=1}^{\infty} |x_k^{(m)} - x_k|^2 \right)^{\frac{1}{2}} + \left( \sum_{k=1}^{\infty} |x_k^{(m)}|^2 \right)^{\frac{1}{2}},$$

using Corollary 2.2.7(a) and consequently  $x \in \ell^2$ . Moreover,  $\lim_m x^{(m)} = x$  in  $\ell^2$  by (2.28).

It follows from Remark 2.3.2(ii) that  $\ell_0$  is an inner product space that is not complete.

*Remarks 2.3.5*

- (i) The inner product space  $\ell_0$  of sequences all of whose terms, from some index onwards, are zero is dense in  $\ell^2$ . In fact, let  $x = (\alpha_1, \alpha_2, \dots)$  be an element in  $\ell^2$  (not in  $\ell_0$ ) and  $\varepsilon > 0$  be given. Choose  $N$  such that

$$\sum_{k=N+1}^{\infty} |\alpha_k|^2 < \varepsilon.$$

Then the sequence  $y = (\alpha_1, \alpha_2, \dots, \alpha_N, 0, \dots)$  is in the desired inner product space and is such that

$$\|x - y\| = \left( \sum_{k=N+1}^{\infty} |\alpha_k|^2 \right)^{\frac{1}{2}} < \varepsilon.$$

This shows that each  $x \in \ell^2$  (not in  $\ell_0$ ) is a limit point of the space  $\ell_0$  of sequences all of whose terms, from some index onwards, are zero.

- (ii) It may be discerned from (i) above that  $\ell_0$  is not complete.  
 (iii) For  $j = 1, 2, \dots$ , let  $e_j = (0, \dots, 0, 1, 0, 0, \dots)$ , where 1 occurs only in the  $j$ th place and

$$E = \{k_1 e_1 + \dots + k_n e_n : n = 1, 2, \dots, \Re k_j, \Im k_j \text{ are rational}\}.$$

Since the rational numbers constitute a countable set,  $E$  is countable. We show that  $E$  is dense in  $\ell^2$ . Let  $(x_1, x_2, \dots) \in \ell^2$  and  $\varepsilon > 0$ . As  $\sum_{j=1}^{\infty} |x_j|^2$  is finite, there is some  $N$  such that

$$\sum_{j=N+1}^{\infty} |x_j|^2 < \varepsilon^2/2.$$

Since the rational numbers are dense in  $\mathbb{R}$ , there are  $k_1, \dots, k_N$  in  $\mathbb{C}$  with  $\Re k_j, \Im k_j$  rational and

$$|x_j - k_j|^2 < \varepsilon^2/2N, \quad j = 1, 2, \dots, N.$$

Consider  $y = k_1 e_1 + \dots + k_N e_N$  in  $E$ . Then

$$\|x - y\|^2 = \sum_{j=1}^N |x_j - k_j|^2 + \sum_{j=N+1}^{\infty} |x_j|^2 < \varepsilon^2/2 + \varepsilon^2/2 = \varepsilon^2.$$

Hence,  $y \in S(x, \varepsilon)$ . Thus,  $E$  is dense in  $\ell^2$ . Consequently,  $\ell^2$  is a separable metric space.

**Definition 2.3.6** Two Hilbert spaces  $H$  and  $K$  are said to be **isometrically isomorphic** if there exists a **linear isometry** between  $H$  and  $K$ , i.e. if there exists a bijective linear mapping  $A: H \rightarrow K$  such that

$$A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A(x_1) + \alpha_2 A(x_2),$$

$$(Ax_1, Ax_2) = (x_1, x_2)$$

for all  $x_1, x_2 \in H$  and scalars  $\alpha_1$  and  $\alpha_2$ .

**Theorem 2.3.7** *For every inner product space  $X$ , there is a Hilbert space  $H$  such that  $X$  is a dense linear subspace of  $H$ , and for  $x, y \in X$ , the inner product  $(x, y)$  in  $X$  and in  $H$  is the same. The space  $H$  is unique up to a linear isometry; that is, if  $X$  is a dense linear subspace of a Hilbert space  $K$ , then there is a unique linear isometry  $A: H \rightarrow K$  such that the restriction of  $A$  to  $X$  is the identity map.*

*Proof* Consider  $X$  as a metric space with the metric induced by the inner product on  $X$ , i.e. with  $d(x, y) = \|x - y\| = (x - y, x - y)^{1/2}$ . Let  $H$  be its completion. Let  $x, y \in H$  and let  $\{x_n\}_{n \geq 1}$  and  $\{y_n\}_{n \geq 1}$  be sequences in  $X$  such that  $x_n \rightarrow x$  and  $y_n \rightarrow y$ . Then for scalars  $\lambda$  and  $\mu$ , the sequence  $\{\lambda x_n + \mu y_n\}_{n \geq 1}$  is a Cauchy sequence in  $X$ . Now, if  $H$  is to be given a Hilbert space structure such that the inner product on  $H$  induces the metric on  $H$ , then

$$\lim_n (\lambda x_n + \mu y_n) = \lambda \lim_n x_n + \mu \lim_n y_n = \lambda x + \mu y.$$

Thus,  $\lambda x + \mu y$  must be defined to be the limit of the Cauchy sequence  $\{\lambda x_n + \mu y_n\}_{n \geq 1}$ . It may be checked that the addition, scalar multiplication and the limits of the Cauchy sequences are well defined. It is now easy to check that with this definition of addition and scalar multiplication,  $H$  becomes a vector space. Now define  $(x, y) = \lim_n (x_n, y_n)$ ; note that it is well defined. In fact, it is an inner product

on  $H$  whose restriction to  $X$  agrees with the given inner product in  $X$ . With this inner product,  $H$  is a Hilbert space. The uniqueness can be easily verified.  $\square$

### Problem Set 2.3

2.3.P1. Show that the space  $(C[0, 1], \|\cdot\|_\infty)$ , where  $\|x\|_\infty = \sup_{0 \leq t \leq 1} |x(t)|$ , is not an inner product space, hence not a Hilbert space.

**Definition** A **strictly convex norm** on a normed linear space is a norm such that, for all  $x, y \in X$ ,  $\|x\| = \|y\| = 1$ ,  $y \neq x \Rightarrow \|x + y\| < 2$ .

- 2.3.P2. (a) Show that the norm on a Hilbert space is strictly convex.  
 (b) Show that the norm  $\|\cdot\|_\infty$  on  $C[0, 1]$  is not strictly convex.  
 (c) Show that the norm  $\|\cdot\|_1$  on  $C[0, 1]$  is not strictly convex.
- 2.3.P3. Let  $H$  be the collection of all absolutely continuous functions  $x: [0, 1] \rightarrow \mathbb{R}$  such that  $x(0) = 0$  and  $x' \in L^2[0, 1]$ . If  $(x, y) = \int_0^1 x'(t) \overline{y'(t)} dt$  for  $x, y \in H$ , show that  $H$  is a Hilbert space.
- 2.3.P4. Let  $H$  be a Hilbert space over  $\mathbb{R}$ . Show that there is a Hilbert space  $K$  over  $\mathbb{C}$  and a map  $U: H \rightarrow K$  such that (i)  $U$  is linear (ii)  $(Ux_1, Ux_2) = (x_1, x_2)$  for all  $x_1, x_2 \in H$ , (iii) for any  $z \in K$ , there are unique  $x_1, x_2 \in H$  such that  $z = Ux_1 + iUx_2$ .
- 2.3.P5. (a) Suppose  $x$  and  $y$  are vectors in a normed space such that  $\|x\| = \|y\|$ . If there exists  $t \in (0, 1)$  such that  $\|tx + (1 - t)y\| < \|x\|$ , then show that this strict inequality holds for all  $t \in (0, 1)$ .  
 (b) Let  $x$  and  $y$  belong to a real or complex strictly convex normed space. If  $\|x + y\| = \|x\| + \|y\|$  and  $x \neq 0 \neq y$ , show that there exists  $\alpha > 0$  such that  $y = \alpha x$ .
- 2.3.P6. The set of all vectors  $x = \{\eta_n\}_{n \geq 1}$  with  $|\eta_n| \leq \frac{1}{n}$ ,  $n = 1, 2, \dots$  in real  $\ell^2$  is called the *Hilbert cube*. Show that this set is compact in  $\ell^2$ .
- 2.3.P7. Let  $a = \{a_n\}_{n \geq 1}$  be a sequence of positive real numbers. Define  $\ell_a^2 = \{x = (x_1, x_2, \dots): x_i \in \mathbb{C} \text{ and } \sum_{i=1}^\infty a_i |x_i|^2 < \infty\}$ . Define an inner product on  $\ell_a^2$  by  $(x, y) = \sum_{i=1}^\infty a_i x_i \overline{y_i}$ . Show that  $\ell_a^2$  is a Hilbert space.
- 2.3.P8. For a real number  $s$ , we define on  $\mathbb{Z}$  a measure  $\mu_s$  by setting

$$\mu_s(\{n\}) = (1 + n^2)^{s/2}, \quad n \in \mathbb{Z}.$$

Put  $H^s = L^2(\mu_s)$ . Prove that for  $r < s$ , we have  $H^s \subseteq H^r$ .

- 2.3.P9. (a) Find a sequence  $a$  of positive real numbers such that  $(1, 1/2^2, 1/3^3, \dots) \notin \ell_a^2$  [see Problem 2.3.P7].  
 (b) Find a sequence  $a$  of positive real numbers such that all  $x = \{x_n\}_{n \geq 1}$  with  $|x_n| = n^n$  are in  $\ell_a^2$ .
- 2.3.P10. Let  $M$  be a closed subspace of a Hilbert space  $H$ , and let  $y \in H$ ,  $y \notin M$ . If  $M'$  is the subspace spanned by  $M$  and  $y$ , then  $M'$  is closed. In particular, a finite-dimensional subspace must be closed.

## 2.4 The Space $L^2(X, \mathfrak{M}, \mu)$

A class of spaces associated with  $\tilde{L}^p(X, \mathfrak{M}, \mu)$  for all  $p$ ,  $0 < p \leq \infty$  [see Definition 1.3.5] is important in analysis. Here we are concerned only with cases  $p = 1$  and  $p = 2$ . We shall see that there is a Hilbert space associated with  $\tilde{L}^2(X, \mathfrak{M}, \mu)$ . Henceforth, the symbols  $\mathfrak{M}$  and  $\mu$  will be omitted.

**Proposition 2.4.1**  $\tilde{L}^2(X)$  is a vector space.

*Proof* Suppose that  $f \in \tilde{L}^2(X, \mathfrak{M}, \mu)$  and  $\alpha \in \mathbb{C}$ . Then  $f$  is measurable and so  $\alpha f$  is measurable;  $|f|^2$  is integrable and so  $|\alpha f|^2 = |\alpha|^2 |f|^2$  is integrable. Thus,  $\alpha f \in \tilde{L}^2(X, \mathfrak{M}, \mu)$ .

Suppose that  $f, g \in \tilde{L}^2(X)$ . Then  $f$  and  $g$  are measurable and so  $f + g$  is measurable. For all complex numbers  $\alpha$  and  $\beta$ ,

$$|\alpha + \beta|^2 \leq 2|\alpha|^2 + 2|\beta|^2.$$

The above inequality may be seen to result on applying the Cauchy–Schwarz Inequality (Theorem 2.2.4) to the inner product  $((\alpha, \beta), (1, 1))$  in the inner product space  $\mathbb{C}^2$ . So, for all  $x \in X$ ,

$$0 \leq |f(x) + g(x)|^2 \leq 2(|f(x)|^2 + |g(x)|^2).$$

The function on the right is integrable and therefore so is the measurable function  $|f + g|^2$ . This proves that  $f + g \in \tilde{L}^2(X)$ .  $\square$

**Definition 2.4.2** If  $f$  and  $g$  are two complex-valued measurable functions defined on  $X$ , let us write ' $f \sim g$ ' if  $\{x \in X: f(x) \neq g(x)\}$  is a null set (a measurable set of measure zero). One says that the functions  $f$  and  $g$  are **equivalent** or that  $f - g$  is a **null function**.

Let  $\mathcal{N}$  denote the set of all null functions, that is,

$$\mathcal{N} = \{f : f \sim 0\}.$$

The relation ' $\sim$ ' is an equivalence relation on the set of all complex-valued measurable functions defined on  $X$ . As such, it partitions the set into disjoint equivalence classes, where a typical class, denoted by  $[f]$ , is given by

$$[f] = \{g : g \text{ measurable on } X \text{ and } g \sim f\}$$

and  $g \in [f]$  would be called a **representative** of this class.

Note that  $\mathcal{N}$  is a vector space of functions and  $\mathcal{N} \subseteq \tilde{L}^p(X)$  for all  $p > 0$ .

If  $f$  is a function in  $\tilde{L}^p(X)$ , then  $[f] = f + \mathcal{N}$  is the coset containing  $f$ , as defined towards the end of Sect. 1.1.

**Definition 2.4.3** The space  $L^2(X, \mathfrak{M}, \mu)$  is the set of all equivalence classes of functions in  $\tilde{L}^2(X)$ .

Thus, if  $f \in \tilde{L}^2(X)$ , then  $[f]$  is the corresponding member of  $L^2(X, \mathfrak{M}, \mu)$ . One says that  $f$  is a representative of the equivalence class  $[f]$ . The set just defined is what we intend to make into the promised Hilbert space associated with  $\tilde{L}^2(X, \mathfrak{M}, \mu)$ .

**Proposition 2.4.4**  $L^2(X)$  is a vector space.

*Proof*  $\mathcal{N}$  is a subspace of  $\tilde{L}^2(X)$ , and  $L^2(X)$  is actually the quotient space  $\tilde{L}^2(X, \mathfrak{M}, \mu) / \mathcal{N}$ . □

We next define an inner product on the space  $\tilde{L}^2(X)$ .

**Proposition 2.4.5** If  $f, g \in \tilde{L}^2(X)$  then  $fg \in \tilde{L}^1(X)$ .

*Proof* Suppose  $f, g \in \tilde{L}^2(X)$ . Then,  $f, g$  are measurable and so the product  $fg$  is measurable. The functions  $|f|^2$  and  $|g|^2$  are integrable and it follows from the inequality

$$|f(x)g(x)| \leq \frac{1}{2} \left( |f(x)|^2 + |g(x)|^2 \right), \quad x \in X$$

that  $fg$  is also integrable. □

Now let us define  $(f, g)$  for  $f, g \in \tilde{L}^2(X)$  by

$$(f, g) = \int_X f(x) \overline{g(x)} d\mu(x) \quad \text{or} \quad \int_X (f\overline{g}) d\mu.$$

Note that if  $g \in \tilde{L}^2(X)$  then  $\overline{g}$  (defined by  $\overline{g}(x) = \overline{g(x)}$ ) is also a member of  $\tilde{L}^2(X, \mathfrak{M}, \mu)$ , and so by Proposition 2.4.5,  $(f, g)$  is well defined. The reader can check that  $(\cdot, \cdot)$  has all the properties of an inner product except one. If  $(f, f) = 0$ , then

$$0 = (f, f) = \int_X |f(x)|^2 d\mu(x) = \int_X |f|^2 d\mu$$

and therefore,  $f \sim 0$ ; that is,  $f$  is a null function, but one cannot conclude that  $f = 0$ .

However, if  $f \sim f'$  and  $g \sim g'$ , then  $(f, g) = (f', g')$ . In fact,

$$\begin{aligned}
\left| \int_X (f\overline{g}) - \int_X f'\overline{g'} \right| &= \left| \int_X (f - f')\overline{g} + \int_X f'(\overline{g - g'}) \right| \\
&\leq \left| \int_X (f - f')\overline{g} \right| + \left| \int_X f'(\overline{g - g'}) \right| \\
&\leq (f - f', f - f')^{\frac{1}{2}}(g, g)^{\frac{1}{2}} + (f', f')^{\frac{1}{2}}(g - g', g - g')^{\frac{1}{2}} = 0,
\end{aligned}$$

using the Cauchy–Schwarz Inequality 2.2.4. In view of the inequality proved above, the integral

$$\int_X (f\overline{g})$$

depends only on the equivalence classes  $[f]$  and  $[g]$  of the functions.

We can now define  $(\cdot, \cdot)$  on  $L^2(X, \mathfrak{M}, \mu)$ .

**Definition 2.4.6** For  $[f], [g] \in L^2(X)$ , define

$$([f], [g]) = \int_X (f\overline{g})d\mu,$$

where  $f \in [f]$  and  $g \in [g]$ .

In view of the remarks preceding Definition 2.4.6,  $(\cdot, \cdot)$  is unambiguously defined on  $L^2(X, \mathfrak{M}, \mu)$ .

**Proposition 2.4.7** *The space  $L^2(X)$  with inner product as in Definition 2.4.6 is an inner product space.*

*Proof* For  $[f] \in L^2(X)$ , if  $([f], [f]) = 0$ , then  $\int_X |f|^2 d\mu = 0$  and hence  $[f] = 0$ , the zero of  $L^2(X)$ . The verification of the other axioms of an inner product is straightforward.  $\square$

**Remark 2.4.8** We shall adopt the usual practice and abandon the notation  $[f]$ . The symbol  $f$  will be used to denote both a function in  $\tilde{L}^2(X)$  and the corresponding equivalence class of functions in  $L^2(X)$ . Working mathematicians tend to ignore the distinction between a function and its equivalence class. The correspondence between statements about  $\tilde{L}^2(X)$  and  $L^2(X)$  is straightforward and gives rise to no confusion. In the subsequent discussions, it will always be clear whether calculations are in terms of functions or equivalence classes of functions.

Note that for  $f \in L^2(X)$ ,

$$\|f\| = (f, f)^{\frac{1}{2}} = \left( \int_X |f(x)|^2 d\mu(x) \right)^{\frac{1}{2}} = \left( \int_X |f|^2 d\mu \right)^{\frac{1}{2}}.$$

Now we can prove a result which is of central importance in analysis. It is often called the Riesz–Fischer Theorem.

**Theorem 2.4.9** (Riesz–Fischer Theorem)  $(L^2(X), (\cdot, \cdot))$  is a Hilbert space.

*Proof* Let  $\{f_n\}_{n \geq 1}$  be a Cauchy sequence in  $L^2(X)$ ; that is, for  $\varepsilon > 0$  there exists an integer  $n_0$  such that

$$\|f_n - f_m\| < \varepsilon \quad \text{whenever} \quad n, m \geq n_0.$$

Then there exists a subsequence  $\{f_{n_i}\}_{i \geq 1}$ ,  $n_1 < n_2 < \dots$ , such that

$$\|f_{n_{i+1}} - f_{n_i}\| = \left( \int_X |f_{n_{i+1}} - f_{n_i}|^2 d\mu \right)^{\frac{1}{2}} < \frac{1}{2^i}, \quad i = 1, 2, \dots$$

Indeed, if  $n_k$  has been selected, choose  $n_{k+1} > n_k$  such that  $n, m > n_{k+1}$  implies  $\|f_n - f_m\| < \frac{1}{2^{k+1}}$ .

Let

$$g_k = \sum_{i=1}^k |f_{n_{i+1}} - f_{n_i}|,$$

and

$$g = \sum_{i=1}^{\infty} |f_{n_{i+1}} - f_{n_i}|.$$

Then by the triangle inequality (Theorem 2.2.6),

$$\|g_k\| = \left\| \sum_{i=1}^k |f_{n_{i+1}} - f_{n_i}| \right\| \leq \sum_{i=1}^k \|f_{n_{i+1}} - f_{n_i}\| < \sum_{i=1}^k \frac{1}{2^i} < 1 \quad \text{for} \quad k = 1, 2, \dots$$

Hence, an application of Fatou's Lemma (Theorem 1.3.8) to  $\{g_k^2\}_{k \geq 1}$  gives

$$\begin{aligned}
\int_X g^2 &= \int_X (\liminf_k g_k^2) d\mu \\
&\leq \liminf_k \int_X g_k^2 d\mu \\
&\leq 1,
\end{aligned}$$

that is,  $\|g\| \leq 1$ . In particular,  $g(x) < \infty$  a.e. Indeed, if  $E = \{x: |g(x)| = \infty\}$  has positive measure, then  $\int_X g(x)^2 d\mu(x) = \infty$ . Therefore, the series

$$g = \sum_{i=1}^{\infty} (f_{n_{i+1}} - f_{n_i}). \quad (2.29)$$

converges absolutely for almost all  $x$ . Denote the sum of (2.29) by  $f(x)$  for those  $x$  at which (2.29) converges. Put  $f(x) = 0$  on the remaining set of measure zero. Since

$$f_{n_1} + \sum_{i=1}^{k-1} (f_{n_{i+1}} - f_{n_i}) = f_{n_k},$$

we see that

$$f(x) = \lim_i f_{n_i}(x) \text{ a.e.}$$

Having determined a function  $f$  which is the pointwise limit almost everywhere of  $\{f_{n_i}\}_{i \geq 1}$ , we have to show that  $f$  is the  $L^2$ -limit of  $\{f_n\}_{n \geq 1}$ , i.e.  $\lim_n \|f_n - f\| = 0$ , where  $\|\cdot\|$  denotes the  $L^2$ -norm. Recall that

$$\|f_n - f_m\| < \varepsilon \quad \text{whenever} \quad n, m \geq n_0.$$

For  $m > n_0$ , another application of Fatou's Lemma shows that

$$\int_X |f - f_m|^2 d\mu \leq \liminf_i \int_X |f_{n_i} - f_m|^2 d\mu \leq \varepsilon^2. \quad (2.30)$$

We conclude from (2.30) that  $f - f_m \in L^2(X)$  and hence  $f \in L^2(X)$  since  $f = (f - f_m) + f_m$ . Finally,

$$\|f - f_m\| \rightarrow 0 \text{ as } m \rightarrow \infty.$$

This completes the proof that  $L^2(X, \mathfrak{M}, \mu)$  is a Hilbert space.  $\square$

*Remark 2.4.10* In the course of the proof, we have shown that if  $\{f_n\}_{n \geq 1}$  is a Cauchy sequence in  $L^2(X)$  with limit  $f$ , then  $\{f_n\}_{n \geq 1}$  has a subsequence which converges pointwise almost everywhere to  $f$ .

The simple functions play an important role in  $L^2(X)$ .

**Theorem 2.4.11** *Let  $S$  be the collection of all measurable simple functions on  $X$  vanishing outside subsets of finite measure. Then  $S$  is dense in  $L^2(X)$ .*

*Proof* Clearly,  $S \subseteq L^2(X)$ . Let  $f \in L^2(X)$  and assume that  $f \geq 0$ . There exists a sequence  $\{s_n\}_{n \geq 1}$  of measurable simple functions such that

$$0 \leq s_1(x) \leq s_2(x) \leq \cdots \leq f(x) \quad \text{and} \quad s_n(x) \rightarrow f(x)$$

(see 1.3.3). Since  $0 \leq s_n(x) \leq f(x)$ , we have  $s_n \in L^2(X)$  and hence  $s_n \in S$ . Since  $|f - s_n|^2 \leq f^2$ , the Dominated Convergence Theorem 1.3.9 shows that  $\|f - s_n\| \rightarrow 0$  as  $n \rightarrow \infty$ . Then,  $f$  is in the  $L^2$ -closure of  $S$ . The general case when  $f$  is complex follows from the above considerations.  $\square$

**Theorem 2.4.12** *Let  $X = \mathbb{R}$ ,  $\mathfrak{M}$  be the  $\sigma$ -algebra of measurable subsets of  $\mathbb{R}$  and  $\mu$  be the usual Lebesgue measure on  $\mathbb{R}$ . Then the set of all continuous functions vanishing outside subsets of finite measure is dense in  $L^2(X)$ .*

*Proof* Let  $E \in \mathfrak{M}$  be such that  $\mu(E) < \infty$ . Then every bounded measurable (in particular continuous) function is square integrable on  $E$ . Consider now a nonempty closed subset  $F$  of  $E$  and its characteristic function  $\chi_F$ . For  $n = 1, 2, \dots$ , let

$$f_n(x) = \frac{1}{1 + n \operatorname{dist}(x, F)}, \quad x \in E,$$

where  $\operatorname{dist}(x, F) = \inf\{|x - y| : y \in F\}$ . Note that each  $f_n$  is continuous on  $E$ . Also,  $f_n(x) = 1$  for all  $x \in F$  and  $f_n(x) \rightarrow 0$  as  $n \rightarrow \infty$  for all  $x \notin F$ . Hence,  $(\chi_F - f_n)(x) \rightarrow 0$  for all  $x \in E$ . Since  $\mu(E) < \infty$  and  $|f_n(x)| \leq 1$  for all  $x \in E$ , the Dominated Convergence Theorem shows that

$$\int_E |\chi_F - f_n|^2 d\mu \rightarrow 0 \text{ as } n \rightarrow \infty.$$

In case  $F$  is empty,  $\chi_F$  is 0 everywhere and we may carry out the above argument with each  $f_n$  chosen to be 0 everywhere.

Now let  $E_0$  be any measurable subset of  $\mathbb{R}$  with  $\mu(E_0) < \infty$ . Let  $\varepsilon > 0$  be given. Then, there exists a closed set  $F \subseteq E_0$  such that  $\mu(E_0 \setminus F) < \varepsilon$ . This follows on using regularity of Lebesgue measure. Since

$$\begin{aligned} d(\chi_{E_0}, f_n) &= \|\chi_{E_0} - f_n\| \\ &\leq d(\chi_{E_0}, \chi_F) + d(\chi_F, f_n), \end{aligned}$$

and since

$$d(\chi_{E_0}, \chi_F) = \left( \int_X |\chi_{E_0} - \chi_F|^2 d\mu \right)^{\frac{1}{2}} = \mu(E_0 \setminus F)^{\frac{1}{2}} < \varepsilon^{\frac{1}{2}},$$

it follows that

$$d(\chi_{E_0}, f_n) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

We have thus proved that the characteristic functions of measurable subsets of finite measure can be approximated in  $L^2$ -norm by continuous functions which vanish outside sets of finite measure. The proof is now completed by using Theorem 2.4.11 and the triangle inequality.  $\square$

### Problem Set 2.4

- 2.4.P1. For which real  $\alpha$  does the function  $f_\alpha(t) = t^\alpha \exp(-t)$ ,  $t > 0$ , belong to  $L^2(0, \infty)$ ? What is  $\|f_\alpha\|$  when defined?
- 2.4.P2. (a) Show that the subspace  $M = \{x = \{x_n\}_{n \geq 1} \in \ell^2 : \sum_{n=1}^{\infty} \frac{1}{n} x_n = 0\}$  is closed in  $\ell^2$ .
- (b) Show that the subspace  $M = \{x(t) \in L^2[1, \infty) : \int_1^{\infty} \frac{1}{t} x(t) dt = 0\}$  is closed in  $L^2[1, \infty)$ .
- 2.4.P3.  $L^p[0, 1]$ ,  $1 \leq p < \infty$ ,  $p \neq 2$ , is not a Hilbert space.

## 2.5 A Subspace of $L^2(X, \mathfrak{M}, \mu)$

The following subspace of  $L^2(X, \mathfrak{M}, \mu)$  will play an important role in the discussion of applications of Hilbert space tools to problems in analysis. Here  $X = [a, b]$ ,  $\mathfrak{M}$  is the  $\sigma$ -algebra of Lebesgue measurable subsets of  $[a, b]$  and  $\mu$  is the Lebesgue measure.

*Example 2.5.1* Let  $[a, b]$  be a closed subinterval of  $\mathbb{R}$ . Let  $C[a, b]$  denote the space of complex-valued continuous functions defined on  $[a, b]$  with inner product given by

$$(f, g) = \int_a^b f(t) \overline{g(t)} dt, \quad f, g \in C[a, b]. \quad (2.31)$$

Then,  $C[a, b]$  is an inner product space (see Example 2.1.3(iv)) which is dense in  $L^2[a, b]$ . Extend  $f \in L^2[a, b]$  to  $\mathbb{R}$  by setting  $f = 0$  outside  $[a, b]$ . The extended function is defined on  $\mathbb{R}$  and is in  $L^2(\mathbb{R})$ . There exists a continuous function  $g$  vanishing outside a set of finite measure such that  $\|f - g\| < \varepsilon$  [see Theorem 2.4.12]. Consider the restriction of  $g$  to  $[a, b]$ , to be denoted by  $h$ . Then the given  $f$  is such that  $\|f - h\| < \varepsilon$ . Moreover,  $C[a, b] \neq L^2[a, b]$  as the following argument shows.

If two functions differ at a point and are both continuous there, then they differ on a neighbourhood of that point. Consequently, they cannot be equivalent. It follows that the function

$$f(x) = \begin{cases} 1 & \text{if } x \in [0, \frac{1}{2}) \\ -1 & \text{if } x \in (\frac{1}{2}, 1] \\ 0 & \text{if } x = \frac{1}{2} \end{cases}$$

is not equivalent to any continuous function. This proves the assertion that  $C[a, b] \neq L^2[a, b]$ .

Thus,  $C[a, b]$  is an inner product space which is not a Hilbert space.

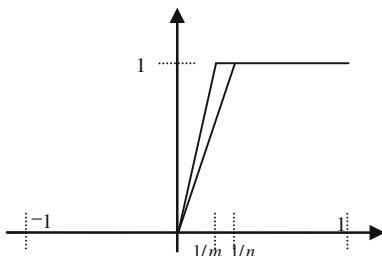
#### Remarks 2.5.2

- (i) The reader will note that if  $g_1$  and  $g_2$  are continuous functions on  $[a, b]$  and  $g_1 \sim g_2$  then  $g_1 = g_2$ . So, if  $f \in L^2[a, b]$  has a representative which is continuous, then that representative is unique. Thus,  $C[a, b] \rightarrow L^2[a, b]$  is an injection. Moreover, uniform convergence of continuous functions implies convergence in  $L^2$ -norm. The aforementioned implication stems from the finiteness of the Lebesgue measure of  $[a, b]$  and it fails if the bounded interval is replaced by an unbounded interval.
- (ii) That the inner product space  $C[a, b]$  with the inner product defined by (2.31) above is not complete can also be seen by exhibiting a Cauchy sequence in the space which converges to an element not lying in the space.

Let  $a$  and  $b$  be  $-1$  and  $1$ , respectively, and consider the sequence

$$f_n(t) = \begin{cases} 0 & -1 \leq t \leq 0 \\ nt & 0 < t \leq \frac{1}{n} \\ 1 & \frac{1}{n} < t \leq 1 \end{cases}.$$

Observe that ( $m > n$ )



$$\begin{aligned}
 \int_{-1}^1 |f_n(t) - f_m(t)|^2 dt &= \int_0^{1/m} (mt - nt)^2 dt + \int_{1/m}^{1/n} (1 - nt)^2 dt \\
 &= (m - n)^2 \frac{1}{3m^3} + \left( \frac{1}{n} - \frac{1}{m} \right) \\
 &\quad - n \left( \frac{1}{n^2} - \frac{1}{m^2} \right) + \frac{n^2}{3} \left( \frac{1}{n^3} - \frac{1}{m^3} \right) \\
 &= \frac{(m - n)^2}{3m^2 n}.
 \end{aligned}$$

The right-hand side of the above equality tends to zero as  $n, m \rightarrow \infty$ . Thus, the sequence  $\{f_n\}_{n \geq 1}$  is Cauchy. We next show that  $f_n \rightarrow f$  in the  $L^2$ -norm, where  $f(t) = 0$  for  $-1 \leq t \leq 0$  and  $f(t) = 1$  for  $0 < t \leq 1$ . In fact,

$$\int_{-1}^1 |f_n(t) - f(t)|^2 dt = \int_0^{1/n} (1 - nt)^2 dt = \frac{1}{n} - \frac{1}{n} + \frac{1}{3n} = \frac{1}{3n} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

The limit function is not equivalent to any continuous function for reasons similar to those in Example 2.5.1.

### Problem Set 2.5

2.5.P1. Prove that the system  $\{1, t^3, t^6, \dots\}$  has a dense linear span in the space  $L^2[0, 1]$  as well as in  $L^2[-1, 1]$ .

## 2.6 The Hilbert Space $A(\Omega)$

Suppose  $\Omega \subseteq \mathbb{C}$  is an arbitrary bounded domain whose boundary consists of smooth simple closed curves.

Consider the class of all holomorphic functions  $f$  in  $\Omega$  for which the integral

$$\iint_{\Omega} |f|^2 dm < \infty, \quad (2.32)$$

where  $dm$  is the two-dimensional Lebesgue measure, exists. In order to interpret the integral as a limit of the Riemann integral, we shall need the following:

**Proposition 2.6.1** *Let  $\Omega$  be a nonempty open subset of  $\mathbb{C}$ . Then there exists a sequence  $\{K_n\}_{n \geq 1}$  of closed and bounded subsets of  $\Omega$  such that  $\Omega$  is their union. Moreover, the sets  $K_n$  can be chosen to satisfy the following conditions:*

- (a)  $K_n \subseteq K_{n+1}^\circ$ ,  $n = 1, 2, \dots$ ;
- (b) every compact (closed and bounded) subset  $K$  of  $\Omega$  is contained in  $K_n$  for some  $n$ .

*Proof* For each positive integer  $n$ , let

$$K_n = \left\{ z : |z| \leq n \text{ and } \text{dist}(z, \mathbb{C} \setminus \Omega) \geq \frac{1}{n} \right\}.$$

Observe that  $K_n$  is both bounded and closed and hence compact or empty. Moreover,  $K_n \subseteq \Omega$ . Obviously,

$$K_n \subseteq \left\{ z : |z| < n+1 \text{ and } \text{dist}(z, \mathbb{C} \setminus \Omega) > \frac{1}{n+1} \right\} \subseteq K_{n+1}^\circ.$$

This proves (a).

We next show that  $\Omega = \bigcup_{n=1}^{\infty} K_n$ . For, if  $z \in \Omega$ , then there exist  $m_1$  and  $m_2$  such that  $|z| \leq m_1$  and  $\text{dist}(z, \mathbb{C} \setminus \Omega) \geq \frac{1}{m_2}$ . Thus,  $z \in K_n$  for some  $n$ . On the other hand,  $K_n \subseteq \Omega$ . This proves that  $\Omega = \bigcup_{n=1}^{\infty} K_n$ .

In view of (a), it follows that  $\Omega = \bigcup_{n=1}^{\infty} K_n^\circ$ . If  $K$  is a compact subset of  $\Omega$ , then the  $K_n^\circ$  form an open cover of  $K$ . By Definition 1.2.16, the compact set  $K$  is covered by finitely many  $K_n^\circ$ . Consequently,  $K$  is contained in  $K_n$  for some large  $n$ . This completes the proof.  $\square$

On letting

$$\varphi_n(z) = \begin{cases} |f(z)|^2 & z \in K_n \\ 0 & z \in \Omega \setminus K_n \end{cases},$$

we see that  $\varphi_n$  is monotonically increasing and  $\lim_n \varphi_n(z) = |f(z)|^2$ ,  $z \in \Omega$ . The Dominated Convergence Theorem now implies

$$\iint_{\Omega} \varphi_n dm \rightarrow \iint_{\Omega} |f|^2 dm \text{ as } n \rightarrow \infty,$$

that is,

$$\iint_{K_n} |f|^2 dm \rightarrow \iint_{\Omega} |f|^2 dm.$$

We now compute in terms of **Taylor coefficients** the integral (2.32) for  $\Omega = \{z: |z| < R\}$  and a holomorphic function  $f$  defined in  $\Omega$ .

The function  $f$  can be expanded in Taylor series

$$f(z) = \sum_{n=0}^{\infty} a_n z^n,$$

where  $a_n = f^{(n)}(0)/n!$ ,  $n = 0, 1, 2, \dots$

$$\begin{aligned} \iint_{\Omega} |f|^2 dm &= \iint_{\Omega} \left| \sum_{n=0}^{\infty} a_n z^n \right|^2 dm \\ &= \iint_{\Omega} \left( \sum_{n=0}^{\infty} a^n r^n e^{in\theta} \right) \left( \sum_{k=0}^{\infty} \overline{a_k} r^k e^{-ik\theta} \right) r dr d\theta \\ &= \int_0^R \int_0^{2\pi} \left( \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} a_n \overline{a_k} r^{n+k+1} e^{i(n-k)\theta} \right) dr d\theta \\ &= \int_0^R \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} a_n \overline{a_k} \int_0^{2\pi} e^{i(n-k)\theta} d\theta r^{n+k+1} dr, \end{aligned}$$

where term-by-term integration is valid because the power series representation converges uniformly on  $|z| \leq r$  for  $r < R$ . Then

$$\iint_{\Omega} |f|^2 dm = \int_0^R 2\pi \sum_{n=0}^{\infty} |a_n|^2 r^{2n+1} dr = \pi \sum_{n=0}^{\infty} |a_n|^2 \frac{R^{2n+2}}{n+1}. \quad (2.33)$$

**Definition 2.6.2** Let  $\Omega$  be a bounded domain in the complex plane. The class of all holomorphic functions in  $\Omega$  for which the integral  $\iint_{\Omega} |f|^2 dm$  is finite is denoted by  $A(\Omega)$  and is known as the **Bergman space**. Briefly,

$$A(\Omega) = \{f \text{ is holomorphic in } \Omega \text{ and } \iint_{\Omega} |f|^2 dm < \infty\}$$

The integral is to be understood in the sense of  $\lim_n \iint_{K_n} |f(z)|^2 dm(z)$ , where  $\{K_n\}$  is a nondecreasing sequence of compact subsets of  $\Omega$  whose union is  $\Omega$ .

The following inequality will be useful in proving that  $A(\Omega)$  is a Hilbert space.

**Proposition 2.6.3** *Suppose  $f \in A(\Omega)$  and  $d_z = \text{dist}(z, \partial\Omega)$ . Then*

$$|f(z)|^2 \leq \left( \iint_{\Omega} |f|^2 dm \right) / \pi d_z^2. \quad (2.34)$$

*Proof* Let  $D$  be the disc with centre  $z$  and radius  $d_z$ . Clearly,

$$\iint_{\Omega} |f|^2 dm \geq \iint_D |f|^2 dm.$$

It follows from the relation (2.33) above that

$$\iint_{\Omega} |f|^2 dm \geq \pi |a_0|^2 d_z^2 = \pi |f(z)|^2 d_z^2.$$

So,

$$|f(z)|^2 \leq \left( \iint_{\Omega} |f|^2 dm \right) / \pi d_z^2.$$

This completes the proof. □

The inequality  $|a + b|^2 \leq 2(|a|^2 + |b|^2)$  implies that

$$|af(z) + bg(z)|^2 \leq 2(|a|^2 |f(z)|^2 + |b|^2 |g(z)|^2) \quad (2.35)$$

for any two functions  $f, g \in A(\Omega)$ ; further, the identity

$$f\bar{g} = \frac{1}{2}|f + g|^2 + \frac{i}{2}|f + ig|^2 - \frac{1+i}{2}|f|^2 - \frac{1+i}{2}|g|^2 \quad (2.36)$$

holds.

**Definition 2.6.4** *For  $f, g \in A(\Omega)$ , we write*

$$(f, g) = \iint_{\Omega} f\bar{g} dm. \quad (2.37)$$

That the right side of (2.37) is finite follows from (2.35) and (2.36) above. It is easily verified that (2.37) defines an inner product on  $A(\Omega)$ .

**Theorem 2.6.5** *With  $(f, g)$  defined as in (2.37),  $A(\Omega)$  is a Hilbert space.*

*Proof* We show that  $A(\Omega)$  is a complete inner product space. In view of (2.35),  $A(\Omega)$  is closed under addition and scalar multiplication.

As usual,  $A(\Omega)$  with norm defined by

$$\|f\| = (f, f)^{\frac{1}{2}} = \left( \iint_{\Omega} |f|^2 dm \right)^{\frac{1}{2}},$$

becomes a normed space [see Definition 2.2.2 and Theorem 2.2.6]. It remains to show that  $A(\Omega)$  is complete in this norm. Suppose  $\{f_n\}_{n \geq 1}$  is a Cauchy sequence in this norm, that is,

$$\|f_n - f_p\|^2 = \left( \iint_{\Omega} |f_n - f_p|^2 dm \right) < \varepsilon$$

for  $n, p \geq N$ . For each compact set  $K \subset \Omega$ , it follows, on using Proposition 2.6.3, that

$$|f_n(z) - f_p(z)|^2 < \varepsilon / \pi d^2 \quad (z \in K),$$

where  $d = \inf\{d(z, \omega) : z \in K \text{ and } \omega \in \partial\Omega\}$ . This means by Weierstrass' Theorem that on each compact subset,  $K \subset \Omega$ , the sequence  $\{f_n\}_{n \geq 1}$  converges uniformly to a holomorphic function  $f$ :

$$f_n(z) \rightarrow f(z) \text{ as } n \rightarrow \infty \quad z \in K \subset \Omega.$$

Since

$$\iint_K |f_n - f_p|^2 dm \leq \iint_{\Omega} |f_n - f_p|^2 dm, \quad n, p \geq N,$$

it follows, on letting  $p \rightarrow \infty$ , that

$$\iint_K |f_n - f|^2 dm \leq \varepsilon, \quad n \geq N,$$

for each compact set  $K$ ; hence,

$$\iint_K |f_n - f|^2 dm \leq \varepsilon, \quad n \geq N,$$

The last inequality implies  $f \in A(\Omega)$  and  $\|f_n - f\| \rightarrow 0$  as  $n \rightarrow \infty$ . This completes the proof.  $\square$

**Problem Set 2.6**

2.6.P1. Let  $\Omega$  be an arbitrary domain in  $\mathbb{C}$  whose boundary consists of a finite number of smooth simple closed curves and let  $A(\Omega)$  be the collection of all holomorphic functions  $f: \Omega \rightarrow \mathbb{C}$  for which

$$\iint_{\Omega} |f(z)|^2 dx dy \left[ \text{same as } \iint_{\Omega} |f|^2 dm \text{ of Sect. 2.6} \right] < \infty$$

holds.

- (a) Show that every  $f \in A(\Omega)$ , where  $\Omega = \{z \in \mathbb{C}: 0 < |z| < 1\}$  has a removable singularity at  $z = 0$ ;
- (b) Show that if  $\alpha \in \Omega$ , then  $\{f \in A(\Omega): f(\alpha) = 0\}$  is closed in  $A(\Omega)$ .

**2.7 Direct Sum of Hilbert Spaces**

The definition of the external direct sum of vector spaces [Definition 1.1.7] is extended to any arbitrary family of vector spaces (each vector space is over the field  $\mathbb{R}$  of real numbers or the field  $\mathbb{C}$  complex numbers) beginning with any finite such family. This procedure lays bare the intricacies involved and aids understanding.

The direct sum

$$H = H_1 \oplus H_2 \oplus \cdots \oplus H_n$$

of the vector spaces  $H_1, H_2, \dots, H_n$  is the set  $H = H_1 \times H_2 \times \cdots \times H_n$ , in which the addition and scalar multiplication are defined by the formula

$$\begin{aligned} (x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) &= (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n), \\ \lambda(x_1, x_2, \dots, x_n) &= (\lambda x_1, \lambda x_2, \dots, \lambda x_n), \end{aligned}$$

where  $(x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n)$  are in  $H_1 \times H_2 \times \dots \times H_n$  and  $\lambda \in \mathbb{R}$  or  $\mathbb{C}$ .

It is quite clear that  $H$  contains a subspace  $Y_i$  of  $H$ , where

$$Y_i = \{(x_1, x_2, \dots, x_n) : x_j = 0 \text{ for } j \neq i\},$$

which is isomorphic to  $H_i$ . It is sometimes convenient to refer to the space  $H_i$  itself as a subspace of  $H$ , and when such reference is made, it is the isomorphic space  $Y_i$  that is to be understood. The map of  $H$  into  $H_i$  given by

$$(x_1, x_2, \dots, x_n) \rightarrow (0, 0, \dots, x_i, 0, \dots)$$

is a projection and is sometimes called projection of  $H$  onto  $H_i$ .

If  $H_1, H_2, \dots, H_n$  are Hilbert spaces, then  $H$  is the uniquely determined Hilbert space with inner product

$$((x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n)) = \sum_{i=1}^n (x_i, y_i)_i, \quad (2.38)$$

where  $(\cdot, \cdot)_i$  is the inner product in  $H_i$ . Then the norm in a direct sum of Hilbert spaces is given by

$$\|(x_1, x_2, \dots, x_n)\| = |((x_1, x_2, \dots, x_n), (x_1, x_2, \dots, x_n))|^{\frac{1}{2}}. \quad (2.39)$$

To summarise, we state the following definition:

**Definition 2.7.1** For each  $i = 1, 2, \dots, n$ , let  $H_i$  be a Hilbert space with inner product  $(\cdot, \cdot)_i$ . The **direct sum** of Hilbert spaces  $H_1, H_2, \dots, H_n$  is the vector space  $H = H_1 \oplus H_2 \oplus \dots \oplus H_n$  in which the inner product and the norm are defined by (2.38) and (2.39).

Henceforth, the subscripts  $i$  in the notation for the inner products and the norms will be omitted because the context will make it clear which one is intended.

**Proposition 2.7.2** With notations as above,  $H = H_1 \oplus H_2 \oplus \dots \oplus H_n$  is a Hilbert space.

*Proof* It must be shown that (2.38) defines an inner product on  $H$  and  $H$  is complete with respect to the norm defined by (2.39). We shall check the second assertion only.

Let  $\{(x_1^{(m)}, x_2^{(m)}, \dots, x_n^{(m)})\}_{m \geq 1}$  be a Cauchy sequence in  $H$ , that is,  $\sum_{i=1}^n \|x_i^{(m)} - x_i^{(\ell)}\|^2 \rightarrow 0$  as  $m, \ell \rightarrow \infty$ . For each  $k$ ,  $\|x_k^{(m)} - x_k^{(\ell)}\|^2 \leq \sum_{i=1}^n \|x_i^{(m)} - x_i^{(\ell)}\|^2$  shows that  $\{x_k^{(m)}\}_{m \geq 1}$  is Cauchy in  $H_k$ . Since  $H_k$  is a Hilbert space, there exists  $x_k$  in  $H_k$  such that  $x_k^{(m)} \rightarrow x_k$  as  $m \rightarrow \infty$ . Clearly, the vector  $(x_1, x_2, \dots, x_n)$  is in  $H$ . It will be shown that  $(x_1^{(m)}, x_2^{(m)}, \dots, x_n^{(m)}) \rightarrow (x_1, x_2, \dots, x_n)$  in  $H$ . Let  $\varepsilon > 0$  be given. For each  $k$ , there exists an integer  $m_k$  such that

$$\|x_k^{(m)} - x_k\| < \varepsilon / \sqrt{n} \quad \text{for } m \geq m_k.$$

Consequently,

$$\sum_{k=1}^n \|x_k^{(m)} - x_k\|^2 < \varepsilon^2 \quad \text{for } m \geq \max\{m_1, m_2, \dots, m_n\}.$$

This completes the proof of the assertion made. □

We next define  $H_1 \oplus H_2 \oplus \cdots$ , also written as  $\bigoplus_{i=1}^{\infty} H_i$ , for a sequence of Hilbert spaces  $H_1, H_2, \dots$ . Let

$$H = \left\{ \{x_n\}_{n \geq 1} : x_n \in H_n, \quad n = 1, 2, \dots \quad \text{and} \quad \sum_{n=1}^{\infty} \|x_n\|^2 < \infty \right\}.$$

For  $x = \{x_n\}_{n \geq 1}$  and  $y = \{y_n\}_{n \geq 1}$  in  $H$ , define

$$(x, y) = \sum_{n=1}^{\infty} (x_n, y_n). \quad (2.40)$$

The sum on the right is seen to be finite by using the Cauchy–Schwarz Inequality for each  $H_i$  and then for  $\ell^2$ . It can then be verified that  $(\cdot, \cdot)$  is an inner product on  $H$  and the norm relative to the inner product is  $\|x\| = (\sum_{n=1}^{\infty} \|x_n\|^2)^{\frac{1}{2}}$ .

With this inner product,  $H$  can be shown to be a Hilbert space.

**Proposition 2.7.3** *With notations as above,  $H = H_1 \oplus H_2 \oplus \cdots = \bigoplus_{i=1}^{\infty} H_i$  is a Hilbert space.*

*Proof* It must be shown that (2.40) defines an inner product on  $H$  and  $H$  is complete with respect to the norm defined by

$$\|x\| = \left( \sum_{n=1}^{\infty} \|x_n\|^2 \right)^{\frac{1}{2}}.$$

For  $x = \{x_n\}_{n \geq 1}$  and  $y = \{y_n\}_{n \geq 1}$  in  $H$ ,

$$\sum_{n=1}^{\infty} |(x_n, y_n)| \leq \sum_{n=1}^{\infty} \|x_n\| \|y_n\| \leq \left( \sum_{n=1}^{\infty} \|x_n\|^2 \right)^{\frac{1}{2}} \left( \sum_{n=1}^{\infty} \|y_n\|^2 \right)^{\frac{1}{2}},$$

using the Cauchy–Schwarz Inequality twice. Hence, the series on the right of (2.40) converges absolutely. Consequently,  $(\cdot, \cdot)$  is well defined. It is a routine exercise to show that  $(\cdot, \cdot)$  is an inner product on  $H$ . It remains to show that  $H$  is a complete space. Suppose  $\{x^{(m)}\}_{m \geq 1} = \{(x_1^{(m)}, x_2^{(m)}, \dots)\}_{m \geq 1}$  is a Cauchy sequence in  $H$ , that is,  $\|x^{(m)} - x^{(n)}\| \rightarrow 0$  as  $m, n \rightarrow \infty$ . For each  $k$ ,  $\|x_k^{(m)} - x_k^{(n)}\|^2 \leq \sum_{j=1}^{\infty} \|x_j^{(m)} - x_j^{(n)}\|^2 = \|x^{(m)} - x^{(n)}\|^2$ , which shows that the sequence  $\{x_k^{(1)}, x_k^{(2)}, \dots\}$  of  $k$ th components is Cauchy. Since  $H_k$  is a Hilbert space,  $x_k^{(n)} \rightarrow x_k$  as  $n \rightarrow \infty$  for suitable  $x_k$  in  $H_k$ . It will be shown that  $\sum_{n=1}^{\infty} \|x_k^{(n)}\|^2 < \infty$  and  $x^{(n)} \rightarrow x$ , where  $x = \{x_k\}_{k \geq 1}$ .

Given  $\varepsilon > 0$ . Let  $p$  be an integer such that  $\|x^{(m)} - x^{(n)}\| < \varepsilon$  whenever  $m, n \geq p$ . For any positive integer  $r$ , one has

$$\sum_{k=1}^r \left\| x_k^{(m)} - x_k^{(n)} \right\|^2 \leq \left\| x^{(m)} - x^{(n)} \right\|^2 \leq \varepsilon^2,$$

provided  $m, n \geq p$ . Letting  $m \rightarrow \infty$ ,

$$\sum_{k=1}^r \left\| x_k - x_k^{(n)} \right\|^2 \leq \varepsilon^2$$

provided  $n \geq p$ . Since  $r$  is arbitrary,

$$\sum_{k=1}^{\infty} \left\| x_k - x_k^{(n)} \right\|^2 \leq \varepsilon^2 \quad (2.41)$$

provided  $n \geq p$ . In particular,

$$\sum_{k=1}^{\infty} \left\| x_k - x_k^{(p)} \right\|^2 \leq \varepsilon^2;$$

hence, the sequence  $\{x_k - x_k^{(p)}\}_{k \geq 1}$  belongs to  $H$ . Consequently, the sequence

$$\{x_k\}_{k \geq 1} = \left\{ x_k - x_k^{(p)} + x_k^{(p)} \right\}_{k \geq 1}$$

belongs to  $H$ . It follows from (2.41) that  $\|x - x^{(n)}\| \leq \varepsilon$  whenever  $n \geq p$ . Thus  $x^{(n)} \rightarrow x$ .  $\square$

**Definition 2.7.4** If  $H_1, H_2, \dots$  are Hilbert spaces, the space  $H$  in Proposition 2.7.3 is called the **direct sum** of  $H_1, H_2, \dots$

For our next definition, a summation over an arbitrary (possibly uncountable) indexing set is to be understood in the following sense:

Suppose  $S = \{x_\alpha: \alpha \in \Lambda\}$ , where  $\Lambda$  is an indexing set, is a collection of elements from a normed linear space  $X$ .  $\{x_\alpha: \alpha \in \Lambda\}$  is said to be **summable to  $x \in X$** , written

$$\sum_{\alpha \in \Lambda} x_\alpha = x \quad \text{or} \quad \sum_{\alpha} x_\alpha = x,$$

if for all  $\varepsilon > 0$ , there exists some finite set of indices  $J_0 \subseteq \Lambda$ , such that for any finite set of indices  $J \supseteq J_0$ ,

$$\left\| \sum_{\alpha \in J} x_\alpha - x \right\| < \varepsilon.$$

**Definition 2.7.5** For each  $\alpha$  in the index set  $\Lambda$ , let  $H_\alpha$  be a Hilbert space. The **direct sum**  $\oplus_\alpha H_\alpha$  of Hilbert spaces  $H_\alpha$  is defined to be the family of all functions  $\{x_\alpha\}$  on  $\Lambda$  such that for each  $\alpha$ ,  $x_\alpha \in H_\alpha$  and  $\sum_{\alpha \in \Lambda} \|x_\alpha\|^2 < \infty$ .

If  $x, y \in H$ ,  $(x, y) = \sum_\alpha (x_\alpha, y_\alpha)$  is an inner product on  $H$ , then  $H$  is a Hilbert space with respect to the norm  $\|x_\alpha\| = \left(\sum_{\alpha \in \Lambda} \|x_\alpha\|^2\right)^{\frac{1}{2}}$ . The proof is not included.

Permuting the index set  $\Lambda$  results in an isomorphic Hilbert space.

Let  $X = C^n[a, b]$ , the linear space of all scalar valued  $n$  times continuously differentiable functions on  $[a, b]$ . For  $x, y$  in  $X$ , define

$$(x, y) = \sum_{j=0}^n \int_a^b x^{(j)}(t) \overline{y^{(j)}(t)} dt.$$

Clearly,  $(\cdot, \cdot)$  is an inner product on  $X$  and

$$\|x\|^2 = (x, x) = \sum_{j=0}^n \int_a^b |x^{(j)}(t)|^2 dt, \quad x \in X.$$

Let

$H = \{x \in C[a, b]: x^{(n-1)} \text{ is defined and absolutely continuous, } x^{(n)} \in L^2[a, b]\}.$

For  $x, y \in H$ , let

$$(x, y) = \sum_{j=0}^n \int_a^b x^{(j)}(t) \overline{y^{(j)}(t)} dt.$$

It can be seen that  $(\cdot, \cdot)$  is an inner product on  $H$  and

$$\|x\|^2 = (x, x) = \sum_{j=0}^n \int_a^b |x^{(j)}(t)|^2 dt, \quad x \in H.$$

**Theorem 2.7.6** With notations as in the paragraph above,  $H$  is a Hilbert space and  $X$  is dense in  $H$ .

*Proof* Consider the direct sum of  $(n + 1)$  copies of  $L^2[a, b]$ , i.e.

$$\oplus_{(n+1)\text{copies}} L^2[a, b].$$

Then  $\oplus_{(n+1)\text{copies}} L^2[a, b]$  is seen to be a Hilbert space by Proposition 2.7.2. Let  $T: H \rightarrow \oplus_{n+1} L^2[a, b]$  be defined by

$$Tx = \left( x, x^{(1)}, x^{(2)}, \dots, x^{(n)} \right).$$

Observe that  $T$  is both linear and injective; moreover, it preserves inner products. We shall next show that  $T(H)$  is a closed subspace of  $\oplus_{n+1} L^2[a, b]$  and is consequently a Hilbert space. This will imply that  $H$  is a Hilbert space.

Let  $\{x_k\}_{k \geq 1}$  be a Cauchy sequence in  $H$  and let

$$y = (y_0, y_1, \dots, y_n) = \lim_{k \rightarrow \infty} Tx_k = \lim_{k \rightarrow \infty} \left( x_k, x_k^{(1)}, x_k^{(2)}, \dots, x_k^{(n)} \right),$$

that is,  $\{x_k^{(j)}\}_{k \geq 1}$  converges in  $L^2[a, b]$  to  $y_j$ ,  $j = 0, 1, 2, \dots, n$ , where  $x_k^{(0)}$  means  $x_k$ . Now, for  $j = 1, 2, \dots, n$  and each  $t \in [a, b]$ ,

$$x_k^{(j-1)}(t) = x_k^{(j-1)}(a) + \int_a^t x_k^{(j)}(\tau) d\tau. \quad (2.42)$$

Observe that

$$\left| \int_a^t y_j(\tau) d\tau - \int_a^t x_k^{(j)}(\tau) d\tau \right| \leq (b-a)^{\frac{1}{2}} \|y_j - x_k^{(j)}\| \quad (2.43)$$

for  $j = 1, 2, \dots, n$  and all  $t \in [a, b]$ . Therefore, the sequence of continuous functions  $\left\{ \int_a^t x_k^{(j)}(\tau) d\tau \right\}_{k \geq 1}$  is uniformly convergent to the continuous function  $\int_a^t y_j(\tau) d\tau$ . It is also convergent as a sequence in  $L^2[a, b]$ . But  $\{x_k^{(j-1)}\}_{k \geq 1}$  is convergent in  $L^2[a, b]$  to  $y_{j-1}$  and so by (2.42), the sequence  $\{x_k^{(j-1)}(a)\}_{k \geq 1}$  of constant functions is convergent in  $L^2[a, b]$ . Therefore, the sequence  $\{x_k^{(j-1)}(a)\}_{k \geq 1}$  is convergent in  $\mathbb{C}$ , and the function on the right of (2.42) is uniformly convergent to a continuous function. Thus, it follows that

$$y_{j-1}(t) = y_{j-1}(a) + \int_a^t y_j(\tau) d\tau.$$

and hence,  $y_{j-1}$  is an absolutely continuous function. We have shown that  $y_0 \in H$ .

Consequently,  $y = Tx$ , where  $x = y_0$ .

We next show that  $X$  is dense in  $H$ .

Let  $x \in H$ . Then  $x^{(n)} \in L^2[a, b]$  so that we can find a sequence  $\{z_m\}_{m \geq 1}$  in  $C[a, b]$  such that  $\|z_m - x^{(n)}\|_2 \rightarrow 0$  as  $m \rightarrow \infty$ . Define recursively  $u_m^{(1)}, u_m^{(2)}, \dots, u_m^{(n)}$  by the formula

$$u_m^{(j)}(t) = \int_a^t u_m^{(j-1)}(\tau) d\tau + x^{(n-j)}(a), \quad j = 1, 2, \dots, n,$$

where  $u_m^{(0)} = z_m$ . Observe that  $u_m^{(n)} \in X$ . We claim that

$$\|u_m^{(j)} - x^{(n-j)}\|_2 \rightarrow 0 \quad \text{as } m \rightarrow \infty.$$

The result is true for  $j = 0$ . Assume that it is true for  $j \geq 0$ . Then

$$\left| u_m^{(j+1)}(t) - x^{n-(j+1)}(t) \right| \leq \|u_m^{(j)} - x^{(n-j)}\|_2 (b-a)^{\frac{1}{2}}, \quad \text{using (2.42) and (2.43)}$$

above.

Hence,  $\{u_m^{(j+1)}(t)\}_{m \geq 1}$  converges uniformly to  $x^{n-(j+1)}$  for  $t \in [a, b]$ , so that  $\|u_m^{(j+1)} - x^{n-(j+1)}\|_2 \rightarrow 0$  as  $m \rightarrow \infty$ . This completes the argument for  $j = 1, 2, \dots, n$ . Consequently,  $u_m^{(n)} \rightarrow x$  in  $H$ . The proof is now complete.  $\square$

## 2.8 Orthogonal Complements

In the familiar Euclidean space, we assign a length to each vector and to each pair of vectors an angle between them. The first notion has been made abstract in the definition of a norm. An appropriate notion of angle and the associated notion of orthogonality are introduced below. The introduction of the concept of orthogonality depends on the definition of inner product in a pre-Hilbert space.

Recall from Definition 2.1.1 that a real vector space  $H$  equipped with an inner product is called a real pre-Hilbert space. The angle  $\theta$  between two nonzero vectors in a real pre-Hilbert space may be defined in a manner consistent with the properties of an inner product by means of the relation  $(x, y) = \|x\| \|y\| \cos \theta$ . Observe that the Cauchy–Schwarz Inequality then says that  $|\cos \theta| \leq 1$ . This definition is not satisfactory in a complex pre-Hilbert space, for  $(x, y)$  is in general a complex number. Nevertheless, if the condition  $(x, y) = 0$  is taken as the definition of orthogonality (perpendicularity), then the concept is just as useful here as in the real case.

**Definition 2.8.1** Let  $H$  be a pre-Hilbert space. Two vectors  $x$  and  $y$  in  $H$  are said to be **orthogonal** if  $(x, y) = 0$ ; we write  $x \perp y$ .

Since  $(x, y) = 0$  implies  $(y, x) = 0$ , we have  $x \perp y$  if, and only if,  $y \perp x$ . It is also clear that  $x \perp 0$  for every  $x$ . Also, the relation  $(x, x) = \|x\|^2$  shows that  $0$  is the only vector orthogonal to itself.

**Definition 2.8.2** A set  $M$  of nonzero vectors in a pre-Hilbert space  $H$  is said to be an **orthogonal set** if  $x \perp y$  whenever  $x$  and  $y$  are distinct vectors of  $M$ . A set  $M$  of vectors in a pre-Hilbert space  $H$  is said to be **orthonormal** if

- (i)  $M$  is orthogonal and (ii)  $\|x\| = 1$  for every  $x \in M$ .

An orthonormal set  $M$  of vectors in a pre-Hilbert space is called **complete** (or **maximal**) **orthonormal system** provided it is not a proper subset of some other orthonormal set.

*Remarks 2.8.3*

- (i) If  $x$  is orthogonal to  $y_1, y_2, \dots, y_n$ , then  $x$  is orthogonal to every linear combination of the  $y_k$ . In fact, if  $x \perp y_k$  for all  $k$  and  $y = \sum_{k=1}^n \lambda_k y_k$ , then

$$(x, y) = \left( x, \sum_{k=1}^n \lambda_k y_k \right) = \sum_{k=1}^n \overline{\lambda_k} (x, y_k) = 0.$$

- (ii) If  $x \perp y$ , then  $\|x + y\|^2 = \|x\|^2 + \|y\|^2$  since  $\|x + y\|^2 = (x + y, x + y) = \|x\|^2 + (x, y) + (y, x) + \|y\|^2 = \|x\|^2 + \|y\|^2$ , using  $(x, y) = 0 = (y, x)$ .
- (iii) An orthogonal subset  $M$  of  $H$  not containing the zero vector is linearly independent. Indeed, if  $\sum_{k=1}^n \lambda_k y_k = 0$ , where  $y_1, y_2, \dots, y_n$  are orthogonal, then on taking the inner product of the sum on the left-hand side with  $y_m$ , we find that  $\lambda_m = 0$ .

*Examples 2.8.4*

- (i) The sequence  $\{x_j\}_{j \geq 1}$ , where  $x_j = (0, 0, \dots, 0, \lambda_j, 0, \dots)$  and the scalar  $\lambda_j$  occurs at the  $j$ th place, in the space  $\ell_0$  of finitely nonzero sequences, is an orthogonal sequence (a sequence whose range is an orthogonal set). The sequence  $\{e_j\}_{j \geq 1}$ , where  $e_j = (0, 0, \dots, 0, 1, 0, \dots)$  and 1 occurs at the  $j$ th place is an orthonormal sequence, in the space  $\ell_0$  of finitely nonzero sequences.
- (ii) Let  $H = C[-\pi, \pi]$  and let  $f_n(x) = \sin nx$ ,  $n = 1, 2, \dots$  and  $g_n(x) = \cos nx$ ,  $n = 1, 2, \dots$ . Since

$$\int_{-\pi}^{\pi} \sin mx \sin nx \, dx = 0 = \int_{-\pi}^{\pi} \cos mx \cos nx \, dx, \quad n \neq m,$$

it follows that  $\{f_n\}_{n \geq 1}$  and  $\{g_n\}_{n \geq 1}$  are orthogonal sequences in  $C[-\pi, \pi]$ . In the space, the vectors

$$u_n(x) = \frac{1}{\sqrt{\pi}} \sin nx, \quad n = 1, 2, \dots$$

form an orthonormal sequence and so do the vectors

$$v_0(x) = \frac{1}{\sqrt{2\pi}}, \quad v_n(x) = \frac{1}{\sqrt{\pi}} \cos nx, \quad n = 1, 2, \dots$$

Also, note that the vectors

$$v_0, v_1, u_1, v_2, u_2, \dots$$

form an orthonormal sequence in  $C[-\pi, \pi]$ .

Recall that  $(f, u_k)$ ,  $k = 1, 2, \dots$  and  $(f, v_k)$ ,  $k = 1, 2, \dots$  are called Fourier coefficients of the function  $f \in C[-\pi, \pi]$ .

- (iii) The sequence  $\varphi_n(z) = \sqrt{\frac{n}{\pi}} z^{n-1}$ ,  $n = 1, 2, \dots$  is orthonormal in  $A(D)$ , where  $D = \{z \in \mathbb{C}: |z| < 1\}$ . In fact,

$$\begin{aligned} (\varphi_n, \varphi_m) &= \iint_D \varphi_n \overline{\varphi_m} \, dx \, dy \\ &= \frac{\sqrt{nm}}{\pi} \int_0^1 \int_0^{2\pi} r^{n+m-1} e^{i(n-m)\theta} \, dr \, d\theta \\ &= \frac{\sqrt{nm}}{\pi(n+m)} \int_0^{2\pi} e^{i(n-m)\theta} \, d\theta \\ &= \begin{cases} 0 & \text{if } n \neq m \\ 1 & \text{if } n = m. \end{cases} \end{aligned}$$

The following definition generalises the notion of Fourier coefficients to any arbitrary infinite-dimensional pre-Hilbert space.

**Definition 2.8.5** If  $\{x_n\}_{n \geq 1}$  is an orthonormal sequence in a pre-Hilbert space  $H$ , then for any  $x \in H$ ,  $(x, x_n)$  is called the **Fourier coefficient** of  $x$  with respect to  $\{x_n\}_{n \geq 1}$ . The **Fourier series** of  $x$  with respect to  $\{x_n\}_{n \geq 1}$  is the series  $\sum_{n=1}^{\infty} (x, x_n) x_n$ .

In the Hilbert space  $\ell^2$ , let

$$e_1 = (1, 0, 0, \dots), e_2 = (0, 1, 0, \dots), e_3 = (0, 0, 1, \dots), \dots$$

Then  $\{e_n\}_{n \geq 1}$  is an orthonormal sequence in  $\ell^2$ . If  $x = \{\lambda_j\}_{j \geq 1} \in \ell^2$ , then  $(x, e_j) = \lambda_j$  and  $x = \sum_{j=1}^{\infty} (x, e_j) e_j$  is its Fourier series with respect to the orthonormal sequence  $\{e_n\}_{n \geq 1}$ . Observe that  $\sum_{j=1}^{\infty} |(x, e_j)|^2 = \sum_{j=1}^{\infty} |\lambda_j|^2 < \infty$ . That this result holds for any orthonormal sequence is a consequence of the following:

**Theorem 2.8.6** (Bessel's Inequality) *Let  $x_1, x_2, \dots, x_n$  be orthonormal vectors in a pre-Hilbert space  $H$ . For every  $x \in H$ ,*

$$\left\| x - \sum_{k=1}^n (x, x_k) x_k \right\|^2 = \|x\|^2 - \sum_{k=1}^n |(x, x_k)|^2,$$

hence

$$\sum_{k=1}^n |(x, x_k)|^2 \leq \|x\|^2.$$

*Proof* For  $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$ ,

$$\left\| \sum_{k=1}^n \lambda_k x_k \right\|^2 = \left( \sum_{k=1}^n \lambda_k x_k, \sum_{k=1}^n \lambda_k x_k \right) = \sum_{k=1}^n |\lambda_k|^2.$$

So,

$$\begin{aligned} \left\| x - \sum_{k=1}^n \lambda_k x_k \right\|^2 &= \left( x - \sum_{k=1}^n \lambda_k x_k, x - \sum_{k=1}^n \lambda_k x_k \right) \\ &= \|x\|^2 - \sum_{k=1}^n \lambda_k (x_k, x) - \sum_{k=1}^n \overline{\lambda_k} (x, x_k) + \sum_{k=1}^n |\lambda_k|^2 \\ &= \|x\|^2 - \sum_{k=1}^n |(x_k, x)|^2 + \sum_{k=1}^n |(x, x_k) - \lambda_k|^2. \end{aligned}$$

In particular, if  $\lambda_k = (x, x_k)$ , then

$$\left\| x - \sum_{k=1}^n (x, x_k) x_k \right\|^2 = \|x\|^2 - \sum_{k=1}^n |(x_k, x)|^2.$$

Since the left-hand side of the above equality is nonnegative, we get

$$\sum_{k=1}^n |(x, x_k)|^2 \leq \|x\|^2.$$

□

Since Bessel's Inequality holds for each  $n$  orthonormal vectors, it yields the following corollary:

**Corollary 2.8.7** *If  $x_1, x_2, \dots$  is any orthonormal sequence of vectors, then for any  $x$  in the pre-Hilbert space  $H$ ,*

$$\sum_{n=1}^{\infty} |(x, x_n)|^2 \leq \|x\|^2.$$

In particular,  $(x, x_n) \rightarrow 0$  as  $n \rightarrow \infty$ .

*Remarks 2.8.8*

- (i) As a special case of Corollary 2.8.7, we obtain the following inequality in the pre-Hilbert space  $C[-\pi, \pi]$ : For  $f \in C[-\pi, \pi]$ ,

$$\sum_{n=1}^{\infty} |(f, u_n)|^2 + \sum_{n=0}^{\infty} |(f, v_n)|^2 \leq \int_{-\pi}^{\pi} |f(x)|^2 dx,$$

where  $u_n(x) = \frac{1}{\sqrt{\pi}} \sin nx$ ,  $v_0(x) = \frac{1}{\sqrt{2\pi}}$ ,  $v_n(x) = \frac{1}{\sqrt{\pi}} \cos nx$ ,  $n = 1, 2, \dots$  [see Example 2.8.4(ii)].

- (ii) Let  $M$  denote the linear manifold spanned by orthonormal vectors  $x_1, x_2, \dots, x_n$ . Then the proof of Theorem 2.8.6 shows that the distance  $\|x - \sum_{k=1}^n \lambda_k x_k\|$  is minimised if we set  $\lambda_k = (x, x_k)$ ,  $k = 1, 2, \dots, n$ ; i.e.  $\|x - \sum_{k=1}^n (x, x_k) x_k\| \leq \|x - \sum_{k=1}^n \lambda_k x_k\|$ , where  $\lambda_1, \lambda_2, \dots, \lambda_n$  are arbitrary scalars.

Thus,  $y = \sum_{k=1}^n (x, x_k) x_k$  is the vector in  $M$  which provides the ‘best approximation’ to the vector  $x$  in the pre-Hilbert space  $H$ . Also note that if  $n > m$ , then in the best approximation by the linear span of  $x_1, x_2, \dots, x_n$ , the first  $m$  coefficients are precisely the same as required for the best approximation in the linear span of  $x_1, x_2, \dots, x_m$ .

- (iii) We set  $z = x - y$ , where  $y = \sum_{k=1}^n (x, x_k) x_k$  provides the best approximation amongst the vectors in  $M$ , then  $(z, x_k) = (x, x_k) - (y, x_k) = 0$  for  $k = 1, 2, \dots, n$ . Hence  $(z, y) = 0$ . Thus,  $x = y + z$ , where  $y$  is a linear combination of  $x_1, x_2, \dots, x_n$  providing the best approximation to  $x$  and  $z \perp x_k$ ,  $k = 1, 2, \dots, n$ , is a decomposition of  $x$ . The decomposition is unique. Indeed, the vector in  $M$  providing the best approximation to  $x \in H$  is unique. If  $x = y_1 + z_1$  is the another decomposition of  $x$ , where  $y_1$  provides the best approximation amongst the vectors in  $M$  and  $z_1 \perp x_k$ ,  $k = 1, 2, \dots, n$ , then  $y + z = y_1 + z_1$  implies  $y - y_1 = z_1 - z$ , which in turn says  $y = y_1$  and  $z = z_1$ , since  $y, y_1$  are in  $M$  and  $z, z_1$  are orthogonal to  $M$ .

It follows from Remark 2.8.3(iii) that every orthonormal sequence in  $H$  is linearly independent. Conversely, given any countable linearly independent sequence in  $H$ , we can construct an orthonormal sequence, keeping the span of the elements at each step of construction [see Theorem 2.8.9 below] in tact.

**Theorem 2.8.9** (Gram–Schmidt orthonormalisation) *Let  $x_1, x_2, \dots$  be a linearly independent sequence in an inner product space  $H$ . Define  $y_1 = x_1$ ,  $u_1 = \frac{x_1}{\|x_1\|}$  and for  $n = 2, 3, \dots$ ,*

$$y_n = x_n - (x_n, u_1)u_1 - (x_n, u_2)u_2 - \cdots - (x_n, u_{n-1})u_{n-1}$$

and

$$u_n = \frac{y_n}{\|y_n\|}.$$

Then  $\{u_1, u_2, \dots\}$  is an orthonormal sequence in  $H$  and for  $n = 1, 2, \dots$ ,

$$\text{span}\{u_1, u_2, \dots, u_n\} = \text{span}\{x_1, x_2, \dots, x_n\}.$$

*Proof* As  $\{x_1\}$  is a linearly independent set,  $y_1 = x_1 \neq 0$  and  $u_1 = \frac{x_1}{\|x_1\|}$  is such that  $\|u_1\| = 1$  and  $\text{span}\{u_1\} = \text{span}\{x_1\}$ .

For  $n \geq 1$ , assume that we have defined  $y_1, y_2, \dots, y_n$  and  $u_1, u_2, \dots, u_n$  as stated above and proved that  $\{u_1, u_2, \dots, u_n\}$  is an orthonormal sequence satisfying  $\text{span}\{u_1, u_2, \dots, u_n\} = \text{span}\{x_1, x_2, \dots, x_n\}$ . Define

$$y_{n+1} = x_{n+1} - (x_{n+1}, u_1)u_1 - (x_{n+1}, u_2)u_2 - \cdots - (x_{n+1}, u_n)u_n.$$

Since the set  $\{x_1, x_2, \dots, x_{n+1}\}$  is linearly independent,  $x_{n+1}$  does not belong to  $\text{span}\{x_1, x_2, \dots, x_n\} = \text{span}\{u_1, u_2, \dots, u_n\}$ . Hence,  $y_{n+1} \neq 0$  and let  $u_{n+1} = \frac{y_{n+1}}{\|y_{n+1}\|}$ . Then  $\|u_{n+1}\| = 1$  and for  $j \leq n$ ,

$$\begin{aligned} (y_{n+1}, u_j) &= (x_{n+1}, u_j) - \sum_{k=1}^n (x_{n+1}, u_k)(u_k, u_j) \\ &= (x_{n+1}, u_j) - (x_{n+1}, u_j) \\ &= 0, \end{aligned}$$

since  $(u_k, u_j) = 0$  for all  $k \neq j$ ,  $k = 1, 2, \dots, n$ . Thus

$$(u_{n+1}, u_j) = \frac{(y_{n+1}, u_j)}{\|y_{n+1}\|} = 0 \quad \text{for } j = 1, 2, \dots, n.$$

Hence,  $\{u_1, u_2, \dots, u_{n+1}\}$  is an orthonormal sequence. Moreover,

$$\begin{aligned} \text{span}\{u_1, u_2, \dots, u_{n+1}\} &= \text{span}\{x_1, x_2, \dots, x_n, u_{n+1}\} \\ &= \text{span}\{x_1, x_2, \dots, x_{n+1}\}. \end{aligned}$$

The argument is now complete in view of mathematical induction.  $\square$

*Remarks 2.8.10*

- (i) The Gram–Schmidt orthonormalisation process as described above yields an orthonormal sequence which is unique.

Let  $e_1, \dots, e_n$  and  $f_1, \dots, f_n$  be  $n$ -term ( $n > 1$ ) orthogonal sequences of nonzero vectors in an inner product space. Suppose they have the same linear span and so do  $e_1, \dots, e_{n-1}$  and  $f_1, \dots, f_{n-1}$ . Then the vectors  $e_n$  and  $f_n$  are scalar multiples of each other, as the following argument shows.

It is sufficient to argue that  $f_n$  is a scalar multiple of  $e_n$ .

Since  $f_n$  lies in the linear span of  $e_1, \dots, e_n$ , there exist scalars  $\lambda_1, \dots, \lambda_n$  such that  $f_n = \sum_{1 \leq k \leq n} \lambda_k e_k$ . However, the vectors  $e_1, \dots, e_{n-1}$  lie in the linear span of  $f_1, \dots, f_{n-1}$ . Therefore, the sum of the first  $n-1$  terms in the preceding sum can be written as  $\sum_{1 \leq k \leq n-1} \lambda_k e_k = \sum_{1 \leq k \leq n-1} c_k f_k$  for some scalars  $c_1, \dots, c_{n-1}$ . Thus,

$$f_n = \sum_{1 \leq k \leq n-1} c_k f_k + \lambda_n e_n. \quad (2.44)$$

By the orthogonality of  $f_1, \dots, f_n$ , for  $1 \leq j \leq n-1$ , we have

$$0 = (f_n, f_j) = \left( \sum_{1 \leq k \leq n-1} c_k f_k + \lambda_n e_n, f_j \right) = c_j (f_j, f_j) + \lambda_n (e_n, f_j).$$

But  $f_j$  lies in the linear span of  $e_1, \dots, e_{n-1}$  and  $e_n$  is orthogonal to this linear span. Therefore,  $(e_n, f_j) = 0$  and the above equality becomes  $c_j (f_j, f_j) = 0$ . As each  $f_j$  is nonzero, it now follows that each  $c_j = 0$  ( $1 \leq j \leq n-1$ ). Using this in (2.44), we get  $f_n = \lambda_n e_n$ .

If the vectors  $e_n$  and  $f_n$  have the same norm, then it further follows that the scalar  $\lambda_n$  has absolute value 1.

- (ii) If  $e_1, e_2, \dots$  and  $f_1, f_2, \dots$  are the orthogonal sequences of nonzero vectors in an inner product space and

$$\text{span}\{e_1, \dots, e_n\} = \text{span}\{f_1, \dots, f_n\} \quad \text{for } n = 1, 2, \dots,$$

then  $e_n$  and  $f_n$  are scalar multiples of each other. If the vectors  $e_n$  and  $f_n$  have the same norm, then it further follows that the scalar factor has absolute value 1.

- (iii) Let  $Q_0, Q_1, \dots$  be the sequence of polynomials obtained from the sequence of polynomials  $1, t, t^2, \dots$  (on the domain  $[-1, 1]$ ) by orthonormalisation, and let  $P_0, P_1, \dots$  be the sequence of Legendre polynomials defined in 2.8.13 below. The first  $k$  functions in either sequence span the space of polynomials of degree at most  $k-1$ . It follows from what has been proved above that each  $Q_n$  is a scalar multiple of  $P_n$  and vice versa. The value of the scalar can be obtained by comparing (a) the leading coefficients or (b) the constant terms or (c) the integrals over  $[-1, 1]$ .
- (iv) The Gram–Schmidt procedure when applied to a finite sequence  $\{x_1, x_2, \dots, x_n\}$  of independent vectors leads to orthonormal vectors  $\{u_1, u_2, \dots, u_n\}$  such that

$$\text{span}\{u_1, u_2, \dots, u_k\} = \text{span}\{x_1, x_2, \dots, x_k\} \quad \text{for } k = 1, 2, \dots, n.$$

As an immediate consequence, we record the following:

**Corollary 2.8.11** *If  $H$  is a pre-Hilbert space of dimension  $n$ , then it has a basis of orthonormal vectors.*

**Theorem 2.8.12** *Every finite-dimensional pre-Hilbert space is complete and is, therefore, a Hilbert space.*

*Proof* By Corollary 2.8.11, there is a basis  $u_1, u_2, \dots, u_n$  of orthonormal vectors. If  $x = \sum_{k=1}^n \lambda_k u_k$ , then  $\|x\|^2 = \sum_{k=1}^n |\lambda_k|^2$ , using Remark 2.8.3(ii). The completeness follows as in Example 2.3.4(i).  $\square$

The following examples illustrate the orthogonalisation procedure.

*Examples 2.8.13*

- (i) Let  $H = \ell^2$ . For  $n = 1, 2, \dots$ , let  $x_n = (1, 1, \dots, 1, 0, 0, \dots)$ , where 1 occurs only in the first  $n$  places. The Gram–Schmidt orthonormalisation process yields

$$y_n = (0, 0, \dots, 0, 1, 0, \dots), \quad n = 1, 2, \dots,$$

where 1 occurs only in the  $n$ th place.

The vector  $y_1 = x_1 = (1, 0, 0, \dots)$ . The vector  $y_2 = \frac{x_2 - (x_2, x_1)x_1}{\|x_2 - (x_2, x_1)x_1\|} = (0, 1, 0, \dots)$ . By induction, it can be shown that

$$y_n = (0, 0, \dots, 0, 1, 0, \dots), \quad n = 1, 2, \dots,$$

where 1 occurs only in the  $n$ th place. The sequence of vectors  $\{y_n\}_{n \geq 1}$  is an orthonormal sequence in  $\ell^2$ .

The set of finite linear combinations of the sequence  $\{y_n\}_{n \geq 1}$  is dense in  $\ell^2$ . Let  $x = \{\lambda_i\}_{i \geq 1} \in \ell^2$ . Given  $\varepsilon > 0$ , there exists  $n_0$  such that  $n > n_0$  implies  $\sum_{n_0+1 \leq i < \infty} |\lambda_i|^2 < \varepsilon$ . Then the vector

$$y = \lambda_1 y_1 + \lambda_2 y_2 + \dots + \lambda_{n_0} y_{n_0}$$

is such that

$$\|x - y\|_2^2 = \sum_{n_0+1 \leq i < \infty} |\lambda_i|^2 < \varepsilon.$$

- (ii) **Legendre polynomials.** Consider the sequence  $\{1, t, t^2, \dots\}$  of vectors in  $L^2[-1, 1]$ . Since any nontrivial finite linear combination  $\sum_{i=1}^n a_i t^{k_i}$  is a polynomial of degree  $m = \max_i k_i$ , it has at most  $m$  zeros. This shows that the

vectors  $\{t^k\}_{k \geq 0}$  are linearly independent. We next calculate the first three orthonormal vectors by the Gram–Schmidt procedure.

Let  $y_0(t) = x_0(t) = 1$ , so that  $\|y_0\|^2 = \int_{-1}^1 ds = 2$  and  $u_0 = \frac{y_0(t)}{\|y_0\|} = \frac{1}{\sqrt{2}}$ . Next,

$$y_1(t) = x_1(t) - (x_1, u_0)u_0(t) = t - \left( \int_{-1}^1 \frac{s}{\sqrt{2}} ds \right) \frac{1}{\sqrt{2}} = t,$$

so that  $\|y_1\|^2 = \int_{-1}^1 s^2 ds = \frac{2}{3}$  and  $u_1(t) = \frac{y_1(t)}{\|y_1\|} = \sqrt{\frac{3}{2}}t$ .

Further,

$$\begin{aligned} y_2(t) &= x_2(t) - (x_2, u_0)u_0(t) - (x_2, u_1)u_1(t) \\ &= t^2 - \left( \int_{-1}^1 \frac{s^2}{\sqrt{2}} ds \right) \frac{1}{\sqrt{2}} - \left( \int_{-1}^1 \sqrt{\frac{3}{2}} s^3 ds \right) \left( \sqrt{\frac{3}{2}} t \right) = t^2 - \frac{1}{3}, \end{aligned}$$

so that  $\|y_2\|^2 = \int_{-1}^1 (s^2 - \frac{1}{3})^2 ds = \frac{8}{45}$  and

$$u_2(t) = \frac{y_2(t)}{\|y_2\|} = \sqrt{10} (3t^2 - 1)/4.$$

We shall next prove that the general form of these orthonormal polynomials is  $\sqrt{n + \frac{1}{2}} P_n(t)$ , where

$$P_n(t) = \frac{1}{2^n n!} \frac{d^n}{dt^n} ((t^2 - 1)^n), \quad n = 0, 1, 2, \dots \quad (2.45)$$

The reader can check by using (2.45) that  $P_0(t) = 1$ ,  $P_1(t) = t$  and  $P_2(t) = \frac{1}{2}(3t^2 - 1)$  and consequently the first three normalised polynomials are  $\frac{1}{\sqrt{2}}$ ,  $\sqrt{\frac{3}{2}}t$  and  $\frac{\sqrt{10}}{4}(3t^2 - 1)$ . That the general form of these polynomials is  $\left\{ \sqrt{n + \frac{1}{2}} P_n(t) \right\}$  will be verified below. We begin by showing that

$$\int_{-1}^1 P_n(t) P_m(t) dt = \begin{cases} 0 & \text{if } n \neq m \\ \frac{2}{2n+1} & \text{if } n = m. \end{cases}$$

For  $m \neq n$ ,

$$\begin{aligned}
2^{n+m} n! m! \int_{-1}^1 P_n(t) P_m(t) dt &= \int_{-1}^1 \frac{d^n}{dt^n} ((t^2 - 1)^n) \frac{d^m}{dt^m} ((t^2 - 1)^m) dt \\
&= \frac{d^{n-1}}{dt^{n-1}} ((t^2 - 1)^n) \frac{d^m}{dt^m} ((t^2 - 1)^m) \Big|_{-1}^1 \\
&\quad - \int_{-1}^1 \frac{d^{m+1}}{dt^{m+1}} ((t^2 - 1)^m) \frac{d^{n-1}}{dt^{n-1}} ((t^2 - 1)^n) dt \\
&= - \int_{-1}^1 \frac{d^{m+1}}{dt^{m+1}} ((t^2 - 1)^m) \frac{d^{n-1}}{dt^{n-1}} ((t^2 - 1)^n) dt
\end{aligned}$$

since  $\frac{d^{n-k}}{dt^{n-k}} ((t^2 - 1)^n)$ ,  $k = 1, 2, \dots, n$ , is zero at  $t = \pm 1$ . Hence, if  $n > m$  and we continue the process of integration by parts, we obtain

$$\pm \int_{-1}^1 \frac{d^{n-m-1}}{dt^{n-m-1}} ((t^2 - 1)^n) \frac{d^{2m+1}}{dt^{2m+1}} ((t^2 - 1)^m) dt,$$

which is equal to zero (the second factor in the integrand is identically zero).

For  $m = n$ , we have

$$\begin{aligned}
\int_{-1}^1 P_n(t)^2 dt &= \frac{(-1)^n}{2^{2n} (n!)^2} \int_{-1}^1 (t^2 - 1)^n \frac{d^{2n}}{dt^{2n}} ((t^2 - 1)^n) dt \\
&= \frac{(-1)^n}{2^{2n} (n!)^2} (2n)! \int_{-1}^1 (t^2 - 1)^n dt
\end{aligned} \tag{2.46}$$

since

$$\frac{d^{2n}}{dt^{2n}} ((t^2 - 1)^n) = (2n)!.$$

Setting  $t = \cos \theta$  in (2.46) and using Wallis' formula,

$$\int_0^{\pi/2} \sin^{2n+1} \theta d\theta = \frac{2^n n!}{1 \cdot 3 \cdot \dots \cdot (2n+1)},$$

[which can be derived by integrating by parts repeatedly], it follows that

$$\int_{-1}^1 P_n(t)^2 dt = \frac{2}{2n+1}.$$

Thus  $\left\{ \sqrt{\frac{2n+1}{2}} P_n(t) \right\}_{n \geq 0}$  is an orthonormal sequence in  $L^2[-1, 1]$ .

It now follows readily that the functions  $\left\{ \sqrt{\frac{2n+1}{2}} P_n(t) \right\}_{n \geq 0}$  are obtained from  $\{1, t, t^2, \dots\}$  by the Gram–Schmidt orthonormalisation procedure since  $P_n(t)$  is a polynomial of degree  $n$ . The essential uniqueness pointed out in Remark 2.8.10(ii), and the fact that the leading coefficients in  $P_n(t)$  and in the  $n$ th polynomial obtained via orthonormalisation are both positive lead to the result.

(iii) **Hermite functions.** Consider the sequence of functions  $\{f_n\}_{n \geq 0}$  on  $\mathbb{R}$ , where  $f_n(t) = t^n \exp(-\frac{t^2}{2})$ . Since

$$\int_0^1 t^{2n} \exp(-t^2) dt < \int_0^1 t^{2n} dt \quad (\exp(-t^2) < 1 \text{ for } 0 < t < 1) \quad (2.47)$$

and

$$\int_1^\infty t^{2n} \exp(-t^2) dt < (n+1)! \int_1^\infty \frac{1}{t^2} dt \quad (e^{t^2} > \frac{t^{2n+2}}{(n+1)!}, t > 1), \quad (2.48)$$

it follows on using (2.47) and (2.48) that

$$\int_{-\infty}^\infty |t^n \exp(-\frac{t^2}{2})|^2 dt = \int_{-\infty}^\infty t^{2n} \exp(-t^2) dt = 2 \int_0^\infty t^{2n} \exp(-t^2) dt$$

is finite. Thus, each  $f_n \in L^2(\mathbb{R})$ . Moreover,  $f_n$ 's are linearly independent, because any nontrivial finite linear combination of functions  $f_n$  is a polynomial multiplied by  $\exp(-\frac{t^2}{2})$ , which is zero for no  $t \in \mathbb{R}$  and any nonzero polynomial has at most finitely many zeros.

We next orthonormalise the functions  $\{f_n\}_{n \geq 0}$ , where  $f_n(t) = t^n \exp(-\frac{t^2}{2})$ , and obtain the first three orthonormal vectors.

To begin with,  $(f_0, f_0) = \int_{-\infty}^\infty \exp(-t^2) dt = \sqrt{\pi}$ , using a well-known formula from advanced calculus. Thus

$$u_0(t) = \frac{\exp(-\frac{t^2}{2})}{\pi^{1/4}}.$$

The next orthonormal vector is

$$\begin{aligned} u_1(t) &= \frac{t \exp(-\frac{t^2}{2}) - \left( t \exp(-\frac{t^2}{2}), \frac{\exp(-\frac{t^2}{2})}{\pi^{1/4}} \right) \frac{\exp(-\frac{t^2}{2})}{\pi^{1/4}}}{\left\| t \exp(-\frac{t^2}{2}) - \left( t \exp(-\frac{t^2}{2}), \frac{\exp(-\frac{t^2}{2})}{\pi^{1/4}} \right) \frac{\exp(-\frac{t^2}{2})}{\pi^{1/4}} \right\|}} \\ &= \frac{\sqrt{2} t \exp(-\frac{t^2}{2})}{\pi^{1/4}} = \frac{2t \exp(-\frac{t^2}{2})}{(2\pi^{1/2})^{1/2}}. \\ u_2(t) &= \frac{t^2 \exp(-\frac{t^2}{2}) - \left( t^2 \exp(-\frac{t^2}{2}), u_0 \right) u_0 - \left( t^2 \exp(-\frac{t^2}{2}), u_1 \right) u_1}{\left\| t^2 \exp(-\frac{t^2}{2}) - \left( t^2 \exp(-\frac{t^2}{2}), u_0 \right) u_0 - \left( t^2 \exp(-\frac{t^2}{2}), u_1 \right) u_1 \right\|}} \\ &= \frac{(4t^2 - 2) \exp(-\frac{t^2}{2})}{(\pi^{1/2} 2^2 2!)^{1/2}}. \end{aligned}$$

We shall next prove that the general form of these orthonormal functions is

$$v_n(t) = \frac{H_n(t) \exp(-\frac{t^2}{2})}{(2^n \cdot n! \pi^{1/2})^{1/2}},$$

where

$$H_n(t) = (-1)^n \exp(t^2) \exp^{(n)}(-t^2), \quad (2.49)$$

and the superscript ‘ $(n)$ ’ indicates the  $n$ th derivative of the function  $t \rightarrow \exp(-t^2)$ . The functions  $\{H_n\}_{n \geq 0}$  are easily seen to be polynomials and are called Hermite polynomials. The degree of  $H_n$  is  $n$ , as shown in (2.50) below. The functions  $v_n$  are called Hermite functions.

For  $n = 0, 1$  and  $2$ , it can be verified that

$$H_0(t) = 1, \quad H_1(t) = 2t \quad \text{and} \quad H_2(t) = 4t^2 - 2.$$

We shall establish below that

$$H'_n(t) = 2nH_{n-1}(t), \quad n = 1, 2, \dots \quad (2.50)$$

In order to do so, we first prove by induction that

$$\exp^{(n+1)}(-t^2) = -2t \exp^{(n)}(-t^2) - 2n \exp^{(n-1)}(-t^2). \quad (2.51)$$

This is true for  $n = 0$ . Assume that it is true for  $n = k - 1$ . Then

$$\begin{aligned} \exp^{(k+1)}(-t^2) &= \frac{d}{dt} \exp^{(k)}(-t^2) \\ &= \frac{d}{dt} (-2t \exp^{(k-1)}(-t^2) - 2(k-1) \exp^{(k-2)}(-t^2)) \\ &= -2t \exp^{(k)}(-t^2) - 2 \exp^{(k-1)}(-t^2) - 2(k-1) \exp^{(k-1)}(-t^2) \\ &= -2t \exp^{(k)}(-t^2) - 2k \exp^{(k-1)}(-t^2). \end{aligned}$$

This proves (2.51) for all  $n = 1, 2, \dots$ . Now, differentiating (2.49) and using (2.51), we obtain

$$\begin{aligned} H'_n(t) &= (-1)^n \left\{ 2t \exp(t^2) \exp^{(n)}(-t^2) + \exp(t^2) \left( -2t \exp^{(n)}(-t^2) - 2n \exp^{(n-1)}(-t^2) \right) \right\} \\ &= (-1)^{n-1} 2n \exp(t^2) \exp^{(n-1)}(-t^2) \\ &= 2n H_{n-1}(t). \end{aligned}$$

This establishes (2.50).

The orthogonality of the Hermite functions may be obtained from

$$\int_{-\infty}^{\infty} H_m(t) H_n(t) e^{-t^2} dt = (-1)^n \int_{-\infty}^{\infty} H_m(t) \exp^{(n)}(-t^2) dt.$$

For  $n > m$ , repeated integration by parts, using (2.50) and the fact that  $\exp(-t^2)$  and all its derivatives vanish for  $t = \pm\infty$ , we obtain

$$\begin{aligned} \int_{-\infty}^{\infty} H_m(t) H_n(t) e^{-t^2} dt &= (-1)^{n-1} 2m \int_{-\infty}^{\infty} H_{m-1}(t) \exp^{(n-1)}(-t^2) dt \\ &= (-1)^{n-m} 2^m m! \int_{-\infty}^{\infty} H_0(t) \exp^{(n-m)}(-t^2) dt = 0. \end{aligned}$$

For  $n = m$ ,

$$\begin{aligned}\int_{-\infty}^{\infty} H_n(t)^2 \exp(-t^2) dt &= 2^n n! \int_{-\infty}^{\infty} H_0(t) \exp(-t^2) dt \\ &= 2^n n! \sqrt{\pi}.\end{aligned}$$

Thus, the functions

$$v_n(t) = \frac{H_n(t) \exp(-\frac{t^2}{2})}{(2^n \cdot n! \pi^{1/2})^{1/2}}, \quad n = 0, 1, 2, \dots \quad (2.52)$$

form an orthonormal sequence.

The reader can check using (2.52) that

$$v_j(t) = u_j(t), \quad j = 0, 1, 2.$$

The vector  $H_n(t) \exp(-\frac{t^2}{2})$  is a linear combination of  $f_0, \dots, f_n$ . Since the sets  $\{H_k(t) \exp(-\frac{t^2}{2}) : k = 0, \dots, n\}$  and  $\{f_k : k = 0, \dots, n\}$  are linearly independent, it follows by Remark 2.8.10(ii) that  $v_j = \pm u_j$  for all  $j$ . The ambiguity of sign can be removed by using the following observation. The leading coefficient of each  $H_n$  is positive in view of (2.50) and so is that of  $f_n \exp(\frac{t^2}{2})$ .

- (iv) **Laguerre functions.** Consider the sequence of functions  $\{f_n\}_{n \geq 0}$  on  $(0, \infty)$ , where  $f_n(t) = t^n \exp(-\frac{t}{2})$ . Since  $\int_0^\infty t^{2n} \exp(-t) dt = \Gamma(2n+1)$ , where  $\Gamma(\cdot)$  is the gamma function and  $\int_0^\infty \exp(-t) dt = 1$ , it follows that each  $f_n \in L^2(0, \infty)$ . Moreover,  $f_n$ 's are linearly independent because any nontrivial finite linear combination of functions  $f_n$  is a polynomial multiplied by  $\exp(-\frac{t}{2})$ , which is zero for no  $t \in (0, \infty)$  and any nonzero polynomial has at most finitely many zeros.

We next orthonormalise the functions  $\{f_n\}_{n \geq 0}$ , where  $f_n(t) = t^n \exp(-\frac{t}{2})$ , and obtain the first three orthonormal vectors.

$$(f_0, f_0) = \int_0^\infty \exp(-t) dt = 1.$$

Thus,  $u_0(t) = \exp(-\frac{t}{2})$ .

The next two orthonormal vectors are given as

$$\begin{aligned}
u_1(t) &= \frac{t \exp(-\frac{t}{2}) - (t \exp(-\frac{t}{2}), u_0) u_0}{\|t \exp(-\frac{t}{2}) - (t \exp(-\frac{t}{2}), u_0) u_0\|} = (t - 1) \exp(-\frac{t}{2}), \text{ since} \\
(t, u_0) &= 1 \text{ and } \|t \exp(-\frac{t}{2}) - (t \exp(-\frac{t}{2}), u_0) u_0\| = 1. \\
u_2(t) &= \frac{t^2 \exp(-\frac{t}{2}) - (t^2 \exp(-\frac{t}{2}), u_1) u_1 - (t^2 \exp(-\frac{t}{2}), u_0) u_0}{\|t^2 \exp(-\frac{t}{2}) - (t^2 \exp(-\frac{t}{2}), u_1) u_1 - (t^2 \exp(-\frac{t}{2}), u_0) u_0\|} \\
&= \left(\frac{1}{2}t^2 - 2t + 1\right) \exp(-\frac{t}{2}), \text{ since} \\
(t^2 \exp(-\frac{t}{2}), u_1(t)) &= 4, \quad (t^2 \exp(-\frac{t}{2}), \exp(-\frac{t}{2})) = 1 \text{ and} \\
\|t^2 \exp(-\frac{t}{2}) - (t^2 \exp(-\frac{t}{2}), u_1) u_1 - (t^2 \exp(-\frac{t}{2}), u_0) u_0\| &= 2.
\end{aligned}$$

We shall next prove that the general form of these orthonormal functions is

$$v_n(t) = \frac{1}{n!} \exp(-\frac{t}{2}) L_n(t), \quad (2.53)$$

where

$$L_n(t) = (-1)^n \exp(t) \frac{d^n}{dt^n} (t^n \exp(-t)), \quad n = 0, 1, 2, \dots$$

Using Leibniz's formula for higher derivatives of a product, we have

$$L_n(t) = (-1)^n \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} n(n-1) \cdots (n-k+1) t^{n-k}. \quad (2.54)$$

The reader can check using (2.53) that  $v_0(t) = \exp(-\frac{t}{2}) = u_0(t)$ ,  $v_1(t) = (t - 1) \exp(-\frac{t}{2}) = u_1(t)$  and  $v_2(t) = (\frac{1}{2}t^2 - 2t + 1) \exp(-\frac{t}{2}) = u_2(t)$ , where  $u_0$ ,  $u_1$  and  $u_2$  are the orthonormalised vectors computed using the Gram-Schmidt orthonormalisation process.

We begin by showing that

$$\int_0^\infty \exp(-t) L_n(t) L_m(t) dt = 0 \quad \text{for } n > m.$$

For  $m < n$ ,

$$\begin{aligned}
\int_0^{\infty} \exp(-t) t^m L_n(t) dt &= (-1)^n \int_0^{\infty} t^m \frac{d^n}{dt^n} (t^n \exp(-t)) dt \\
&= (-1)^{n+m} m! \int_0^{\infty} \frac{d^{n-m}}{dt^{n-m}} (t^n \exp(-t)) dt = 0,
\end{aligned}$$

by repeated integration by parts. Also,

$$\begin{aligned}
\int_0^{\infty} \exp(-t) L_n^2(t) dt &= (-1)^n \int_0^{\infty} \left( \frac{d^n}{dt^n} (t^n \exp(-t)) \right) L_n(t) dt \\
&= (-1)^n \int_0^{\infty} \frac{d^n}{dt^n} (t^n \exp(-t)) \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} n(n-1) \dots (n-k+1) t^{n-k} dt \\
&= \int_0^{\infty} (-1)^{2n} t^n \frac{d^n}{dt^n} (t^n \exp(-t)) dt \\
&= n! \int_0^{\infty} t^n \exp(-t) dt = (n!)^2,
\end{aligned}$$

which shows that  $\{v_n\}_{n \geq 0}$  is orthonormal.

The vector  $L_n(t) \exp(-\frac{t}{2})$  is a linear combination of  $f_0, \dots, f_n$ . Since the sets  $\{L_k(t) \exp(-\frac{t}{2}) : k = 0, \dots, n\}$  and  $\{f_k : k = 0, \dots, n\}$  are linearly independent, it follows by Remark 2.8.10(ii) that  $v_j = \pm u_j$  for all  $j$ . The ambiguity of sign can be removed by using the following observation. The leading coefficient of each  $L_n$  is positive in view of (2.54) and so is that of  $f_n \exp(-\frac{t}{2})$ .

(v) **Rademacher functions.** Consider the sequence  $\{r_n\}$  of functions defined on the interval  $[0, 1]$  by

$$r_0(t) = 1, \quad r_k(t) = \operatorname{sgn}(\sin 2^k \pi t), \quad k = 1, 2, \dots, \quad t \in [0, 1].$$

This sequence was introduced by Rademacher and the  $r_k$  are known as Rademacher functions. If the interval  $[0, 1]$  is divided into  $2^k$  ( $k \geq 1$ ) equal parts, then  $r_k(t)$  assumes on the interiors of those segments the values  $+1$  and  $-1$  alternately while at the endpoints,  $r_k(t) = 0$ .

The reader will note that  $\|r_n\| = (\int_0^1 |r_n(t)|^2 dt)^{\frac{1}{2}} = 1$ , i.e.  $r_n \in L^2[0, 1]$  and  $\|r_n\| = 1$ ,  $n = 0, 1, 2, \dots$ . To prove orthogonality, let  $n > m \geq 0$ . Let  $I$  be the open segment which lies between some two consecutive points of subdivision of the interval  $[0, 1]$  corresponding to the function  $r_m$ . Then,  $r_m$  has constant value  $+1$  or  $-1$  on  $I$ . Furthermore,  $I$  is composed of an even number, precisely  $2^{n-m}$ , of intervals

of equal length. On half of these intervals,  $r_n(t)$  has value +1, whereas on the other half,  $r_n(t)$  has value -1. Consequently,

$$\int_I r_m(t)r_n(t)dt = \pm \int_I r_n(t)dt = 0.$$

Summing up over all such segments  $I$ , we have

$$\int_0^1 r_m(t)r_n(t)dt = 0.$$

Since the vectors of an orthonormal system in a pre-Hilbert space cannot be linearly dependent, it follows that the Rademacher sequence  $\{r_n\}_{n \geq 0}$  of orthonormal functions in  $L^2[0, 1]$  is a linearly independent sequence. Moreover, the function  $f(t) = \cos 2\pi t$  is such that

$$\int_0^1 r_n(t)f(t)dt = \sum_{k=1}^{2^n} (-1)^{k-1} \int_{\frac{k-1}{2^n}}^{\frac{k}{2^n}} f(t)dt = 0,$$

since the  $k$ th term and the  $(2^n - (k - 1))$ th term are equal in magnitude and opposite in sign because  $\cos 2\pi t = \cos 2\pi(1 - t)$ , and consequently add up to 0.

*Remark* The sequence  $\{r_n\}_{n \geq 0}$  converges for  $t = 0, 1, \frac{k}{2^n}, 1 \leq k < 2^n, n = 1, 2, \dots$ , the points of subdivision of the interval  $[0, 1]$ , and converges for no  $t$  other than these points of subdivision, for if  $t \neq 0, 1$  or any of the points of subdivision,  $\{r_n(t)\}_{n \geq 0}$  assumes the values +1 and -1 infinitely often. (For any  $t$  that is not of the form  $\frac{k}{2^n}$ , there exists an integer  $j$  such that  $\frac{j}{2^n} < t < \frac{j+1}{2^n}$ , that is,  $j\pi < 2^n\pi t < (j+1)\pi$ : so,  $r_n(t)$  is 1 if  $j$  is even and -1 if  $j$  is odd. As  $n$  increases, the parity of  $j$  keeps changing between even and odd [see Problem 2.8.P10.]). Thus, the sequence converges only on the set  $\{0, 1, \frac{k}{2^n}; 1 \leq k < 2^n, n = 1, 2, \dots\}$  of measure zero. However, the arithmetic averages of  $\{r_n\}_{n \geq 0}$  converge to the zero function almost everywhere.

**Lemma 2.8.14** For distinct nonnegative integers  $k_1, k_2, \dots, k_n$ ,

$$\int_0^1 r_{k_1} r_{k_2} \dots r_{k_n} = 0.$$

*Proof* This is left as Problem 2.8.P11. □

**Theorem 2.8.15** *Let  $\{r_n\}_{n \geq 1}$  be the Rademacher functions. Then the sequence  $\{(\sum_{k=1}^n r_k)/n\}_{n \geq 1}$  of arithmetic means converges to zero almost everywhere with respect to the Lebesgue measure on  $[0, 1]$ .*

*Proof* Set  $f_n = [(\sum_{k=1}^n r_k)/n]^4$ ,  $n = 1, 2, \dots$ . Observe that each  $f_n$  belongs to  $L^1[0, 1]$ . Indeed,  $|f_n| \leq [(\sum_{k=1}^n |r_k|)/n]^4 = 1$  and  $\int_0^1 f_n(t) dt \leq 1$ . Next, on using  $r_k^2 = 1$ ,  $k = 1, 2, \dots$  (except at the finitely many points of subdivision), we have

$$\begin{aligned}
 n^4 f_n &= \left[ \left( \sum_{k=1}^n r_k \right)^2 \right]^2 \\
 &= \left( \sum_{k=1}^n r_k^2 + 2 \sum_{\substack{k,m=1 \\ k < m}}^n r_k r_m \right)^2 \\
 &= \left( n + 2 \sum_{\substack{k,m=1 \\ k < m}}^n r_k r_m \right)^2 \\
 &= n^2 + 4n \sum_{\substack{k,m=1 \\ k < m}}^n r_k r_m + 4 \left( \sum_{\substack{i,j=1 \\ i < j}}^n r_i r_j \right) \left( \sum_{\substack{k,m=1 \\ k < m}}^n r_k r_m \right) \\
 &= n^2 + 4n \sum_{\substack{k,m=1 \\ k < m}}^n r_k r_m + 4 \sum_{\substack{k,m=1 \\ k < m}}^n r_k^2 r_m^2 + 4 \left( \sum_{j=1}^n 2r_j^2 \sum_{\substack{k,m=1 \\ k < m \\ k,m \neq j}}^n r_k r_m + 2 \sum_{\substack{k,m=1 \\ i,j=1 \\ k < m \\ i < j \\ i,j,k,m \text{ dist}}}^n r_i r_j r_k r_m \right) \\
 &= n^2 + 4n \sum_{\substack{k,m=1 \\ k < m}}^n r_k r_m + 4((n-1) + (n-2) + \dots + 2 + 1) + 8 \sum_{j=1}^n \sum_{\substack{k,m=1 \\ k < m \\ k,m \neq j}}^n r_k r_m \\
 &\quad + 8 \sum_{\substack{k,m=1 \\ i,j=1 \\ k < m \\ i < j \\ i,j,k,m \text{ dist}}}^n r_i r_j r_k r_m \\
 &= n^2 + 2n(n-1) + 4n \sum_{\substack{k,m=1 \\ k < m}}^n r_k r_m + 8 \sum_{j=1}^n \sum_{\substack{k,m=1 \\ k < m \\ k,m \neq j}}^n r_k r_m + 8 \sum_{\substack{k,m=1 \\ i,j=1 \\ k < m \\ i < j \\ i,j,k,m \text{ dist}}}^n r_i r_j r_k r_m.
 \end{aligned}
 \tag{2.55}$$

Dividing both sides of (2.55) by  $n^4$ , integrating and using Lemma 2.8.14, we obtain

$$\int_0^1 f_n dt = \frac{1}{n^2} + \frac{2n(n-1)}{n^4} < \frac{3}{n^2}.$$

Consequently,

$$\sum_{n=1}^{\infty} \int_0^1 f_n dt < \infty.$$

By Corollary 1.3.7 and Remark 1.3.13, it follows that the sequence  $\{f_n\}_{n \geq 1}$  converges to zero almost everywhere; that is, the sequence  $\{(\sum_{k=1}^n r_k)/n\}_{n \geq 1}$  of arithmetic averages converges to zero almost everywhere with respect to Lebesgue measure. This completes the proof.  $\square$

### Problem Set 2.8

- 2.8.P1. Using Bessel's Inequality, obtain the Cauchy–Schwarz Inequality.  
 2.8.P2. Give an example to show that strict inequality can hold in the Corollary 2.8.7 to Bessel's Inequality.  
 2.8.P3. Let  $\{e_k\}_{k \geq 1}$  be any orthonormal sequence in an inner product space  $X$ . Show that for any  $x, y \in X$ ,

$$\sum_{n=1}^{\infty} |(x, e_k)(y, e_k)| \leq \|x\| \|y\|.$$

- 2.8.P4. Let  $\{e_k\}_{k \geq 1}$  be any orthonormal sequence in a Hilbert space  $H$  and let  $M = \text{span}\{e_k\}$ . Show that for any  $x \in H$ , we have  $x \in \overline{M}$  if, and only if,  $x$  can be represented by

$$x = \sum_{k=1}^{\infty} (x, e_k) e_k.$$

- 2.8.P5. Let  $f(x)$  be a differentiable  $2\pi$ -periodic function on  $[-\pi, \pi]$  with derivative  $f'(x) \in L^2[-\pi, \pi]$ . Let  $f_n$ ,  $n \in \mathbb{Z}$ , be the Fourier coefficients of  $f(x)$  in the system  $\{e^{inx}/\sqrt{2\pi}\}_{n \in \mathbb{Z}}$ . Prove that  $\sum_{n=-\infty}^{\infty} |f_n| < \infty$ .  
 2.8.P6. Show that the system  $\{1, t^2, t^4, \dots\}$  is complete in the space  $L^2[0, 1]$ . It is not complete in  $L^2[-1, 1]$ .  
 2.8.P7. Find a nonzero vector in  $\mathbb{C}^3$  orthogonal to  $(1, 1, 1)$  and  $(1, \omega, \omega^2)$ , where  $\omega = \exp(2\pi i/3)$ .

- 2.8.P8. Let  $\alpha \in \mathbb{C}$  be such that  $|\alpha| \neq 1$ . Find the Fourier coefficients of  $f \in \text{RL}^2$ , where  $f(z) = (z - \alpha)^{-1}$ , with respect to the orthonormal sequence  $\{e_j\}_{j=-\infty}^{\infty}$ ,  $e_j(z) = z^j$ .
- 2.8.P9. If the series  $\frac{|a_0|^2}{2} + \sum_{n=1}^{\infty} (|a_n|^2 + |b_n|^2)$  converges, show that there exists a function  $f \in L^2[0, 2\pi]$  having the  $a_k, b_k$  as its Fourier coefficients, i.e. the equations

$$a_0 = \frac{1}{\pi} \int_0^{2\pi} f(t) dt, \quad a_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos kt dt, \quad b_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin kt dt, \\ k = 1, 2, \dots$$

are valid. This function is uniquely defined up to a set of measure zero; i.e. if there are two such functions, they differ only on a set of measure zero.

- 2.8.P10. Show that for any  $t \in [0, 1]$  that is not of the form  $\frac{k}{2^n}$  (i.e.  $t$  is not a ‘dyadic rational’) for any integers  $k$  and  $n$ , the parity of the (obviously unique) integer  $j$  such that  $\frac{j-1}{2^n} < t < \frac{j}{2^n}$  keeps changing between even and odd as  $n$  increases.
- 2.8.P11. Prove Lemma 2.8.14: for distinct nonnegative integers  $k_1, k_2, \dots, k_n$ ,  $\int_0^1 r_{k_1} r_{k_2} \dots r_{k_n} = 0$ .
- 2.8.P12. Show that completeness of the orthonormal set of Hermite functions in  $L^2(-\infty, \infty)$  is equivalent to that of the orthonormal set of Laguerre functions in  $L^2(0, \infty)$ .
- 2.8.P13. Let  $X$  be a complex inner product space of dimension  $n$ . Show that  $X$  is isometrically isomorphic to  $\mathbb{C}^n$  and is hence complete.

## 2.9 Complete Orthonormal Sets

Recall that a set  $M$  of vectors in a pre-Hilbert space is said to be orthogonal if  $x \perp y$  whenever  $x$  and  $y$  are distinct vectors of  $M$  [Definition 2.8.1]. The orthogonal set  $M$  is said to be orthonormal if, in addition,  $\|x\| = 1$  for every vector  $x$  in  $M$ .

An orthonormal set is said to be complete if it is a maximal orthonormal set [Definition 2.8.2]. We shall show that there are complete orthonormal sets in any nontrivial inner product space and discuss a few of the many important examples. One also speaks of complete orthogonal sets, which are defined analogously. The classical result of Riesz–Fischer and Parseval will be proved. These will lead to the identification of all infinite-dimensional Hilbert spaces.

We begin by showing that a nontrivial inner product space  $H$  ( $H \neq \{0\}$ ) contains a complete orthonormal set.

**Theorem 2.9.1** *Let  $H$  be an inner product space over  $\mathbb{F}$  and let  $H \neq \{0\}$ . Then  $H$  contains a complete orthonormal set.*

*Proof* Let  $S$  denote the collection of all orthonormal sets in  $X$ . Since, for any nonzero vector  $x$ , the set  $\{\frac{x}{\|x\|}\}$  is an orthonormal set; it follows that  $S \neq \emptyset$ . The collection  $S$  is partially ordered by inclusion. We wish to show that every totally ordered subset of  $S$  has an upper bound in  $S$ . It will then follow by Zorn's Lemma that  $S$  has a maximal element, namely a complete orthonormal set.

Let  $T = \{A_\alpha\}_{\alpha \in \Lambda}$ , where  $\Lambda$  is an indexing set, be any totally ordered subset of  $S$ . Then the set  $\bigcup_\alpha A_\alpha$  is an upper bound for  $T$ ; indeed,  $A_\beta \subseteq \bigcup_\alpha A_\alpha$  for each  $\beta$ . We next show that  $\bigcup_\alpha A_\alpha$  is orthonormal. Let  $x$  and  $y$  be any two distinct elements of  $\bigcup_\alpha A_\alpha$  so that  $x \in A_\beta$  and  $y \in A_\gamma$  for some  $\beta$  and  $\gamma$  in the indexing set  $\Lambda$ . Since  $T$  is totally ordered, either  $A_\beta \subseteq A_\gamma$  or  $A_\gamma \subseteq A_\beta$ . Supposing  $A_\beta \subseteq A_\gamma$ , it follows that  $x, y \in A_\gamma$ . So  $x \perp y$  and  $\|x\| = \|y\| = 1$ . Thus,  $\bigcup_\alpha A_\alpha$  is seen to be orthonormal.

By Zorn's Lemma [Sect. 1.3],  $S$  has a maximal element. This completes the proof.  $\square$

A slight modification of the proof of Theorem 2.9.1 yields the following corollary.

**Corollary 2.9.2** *Let  $H$  be an inner product space over  $\mathbb{F}$ . If  $E \subseteq H$  is an orthonormal set, then there exists a complete orthonormal set  $S$  such that  $E \subseteq S$ .*

The next result contains an alternate description of complete orthonormal sets.

**Theorem 2.9.3** *Let  $H$  be an inner product space over  $\mathbb{F}$ . Suppose that  $S \subseteq H$  is an orthonormal set. Then the following are equivalent:*

- (a)  $S$  is a complete orthonormal set;
- (b) If  $x \in H$  is such that  $x \perp S$ , then  $x = 0$ .

*Proof* Suppose  $S$  is a complete orthonormal set. If  $x \in H$  is such that  $x \perp S$ , and  $x \neq 0$ , then  $S \cup \{\frac{x}{\|x\|}\}$  is an orthonormal set that properly contains  $S$ , contradicting the fact that  $S$  is a complete orthonormal set.

On the other hand, suppose  $x \perp S$  implies  $x = 0$ . If  $S$  were not a complete orthonormal set, there would exist some orthonormal set  $T \subseteq H$  such that  $T$  properly contains  $S$ . Hence, if  $x \in T \setminus S$  then  $\|x\| = 1$  and  $x \perp S$ . This contradicts the assumption that  $x \perp S$  implies  $x = 0$ .

Therefore, the orthonormal set  $S$  is complete.  $\square$

So far we have considered examples of countable orthonormal sets in pre-Hilbert spaces. If a Hilbert space contains a countable complete orthonormal set, then it is said to be **separable**. This definition of separability is equivalent to Definition 1.2.10 as the next theorem shows.

Let  $S$  be a countable dense set in a Hilbert space  $H \neq \{0\}$ . By progressively reducing  $S$ , if necessary, it can be turned into a linearly independent set. The Gram-Schmidt orthonormalisation process applied to the linearly independent set renders

it into an orthonormal set. This orthonormal set is in fact complete. More precisely, we have the following theorem.

**Theorem 2.9.4** *Let  $H \neq \{0\}$  be a Hilbert space that contains a countable dense subset  $S$ . Then  $H$  contains a countable complete orthonormal set that is obtained from  $S$  by the Gram–Schmidt orthonormalisation process. Thus  $H$  is separable.*

*Let  $H \neq \{0\}$  contain a countable, complete orthonormal set  $T$ , then  $H$  contains a countable dense set, namely the finite rational linear combinations of vectors in  $T$ .*

*Proof* We assume, as we may, that  $0 \notin S$ . Enumerate the vectors in  $S$  as a sequence  $\{x_n\}_{n \geq 1}$  and let  $y_1 = x_{n_1}$ , where  $n_1 = 1$ . If all the  $x_n$  for  $n > n_1$  are scalar multiples of  $x_{n_1}$ , then the set  $\{x_{n_1}\}$  is the linearly independent set obtained from  $S$ . Otherwise, let  $y_2 = x_{n_2}$  be the first  $x_n$  which is not a scalar multiple of  $x_{n_1}$ . Then for  $n < n_2$ ,  $x_n$  is a scalar multiple of  $x_{n_1}$ . If all the  $x_n$  for  $n > n_2$  are expressible as linear combinations of  $x_{n_1}$  and  $x_{n_2}$ , then the set  $\{x_{n_1}, x_{n_2}\}$  is the linearly independent set obtained from  $S$ . Otherwise, let  $y_3 = x_{n_3}$  be the first  $x_n$  which is independent of  $x_{n_1}$  and  $x_{n_2}$ . Then for  $n < n_3$ ,  $x_n$  is a linear combination of  $x_{n_1}$  and  $x_{n_2}$ . The proof continues inductively, and we thus obtain a finite or countably infinite linearly independent set  $\{y_1, y_2, \dots\} \subseteq S$ . Let  $X$  be the smallest linear subspace of  $H$  containing  $\{y_1, y_2, \dots\}$ . It is clear that  $S \subseteq X$  since if  $x_j \in S$  then  $x_j$  is a linear combination of  $y_1, y_2, \dots, y_k$ , where  $k$  is chosen so that  $n_k \leq j < n_{k+1}$ . This says that  $X$  is dense in  $H$ . Orthonormalise  $\{y_1, y_2, \dots\}$  by the Gram–Schmidt procedure to obtain the orthonormal set  $\{u_1, u_2, \dots\}$ . It remains to show that the orthonormal set  $\{u_1, u_2, \dots\}$  is complete.

Let  $x \in H$  be such that  $(x, u_k) = 0$  for  $k = 1, 2, \dots$ . Then  $(x, \sum_{k=1}^n \alpha_k u_k) = 0$  for all finite linear combinations of the  $u_n$  and so  $(x, y) = 0$  for all  $y \in X$ . Let  $\{z_n\}_{n \geq 1}$  be a sequence in  $X$  such that  $\|x - z_n\| \rightarrow 0$  as  $n \rightarrow \infty$ . Then  $\|x\|^2 = (x, x) = (x, z_n) = (x, x - z_n) \leq \|x\| \|x - z_n\| \rightarrow 0$  as  $n \rightarrow \infty$ .

Clearly, the closure of the rational linear combinations of the vectors of  $T = \{x_k\}$  contains all possible linear combinations of  $T$ , i.e. contains  $[T]$  and is hence the same as  $[\overline{T}]$ . Let  $x \in H$ . Now,  $\sum_{k=1}^n (x, x_k) x_k \in [T]$ . Using Bessel's Inequality [Theorem 2.8.6], it follows that  $\sum_{n=1}^{\infty} (x, x_k) x_k$  converges to some  $y \in H$ . In fact,  $y \in [\overline{T}]$ . Suppose  $y \neq x$ . Then

$$(x - y, x_k) = (x, x_k) - (y, x_k) = (x, x_k) - (x, x_k) = 0.$$

Using the completeness of  $T$ , it follows that  $x - y = 0$ . Thus  $x \in [\overline{T}]$ . This completes the proof.  $\square$

However, there are Hilbert spaces which contain non-denumerable orthonormal sets and are, therefore, **nonseparable**. We give below examples of such Hilbert spaces.

### Examples 2.9.5

- (i) Consider the collection  $X$  of functions on  $\mathbb{R}$  representable in the form

$$x(t) = \sum_{k=1}^n a_k e^{i\lambda_k t}$$

for arbitrary  $n$ , real numbers  $\lambda_1, \lambda_2, \dots, \lambda_n$  and complex coefficients  $a_1, a_2, \dots, a_n$ .  $X$  is a vector space, and an inner product in  $X$  is defined by

$$(x, y) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t) \overline{y(t)} dt.$$

If  $y(t) = \sum_{k=1}^n b_k e^{i\mu_k t}$ , then

$$\begin{aligned} (x, y) &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \sum_{j=1}^n \sum_{k=1}^m a_j \overline{b_k} e^{i(\lambda_j - \mu_k)t} dt \\ &= \sum_{j=1}^n \sum_{k=1}^m a_j \overline{b_k} \end{aligned} \quad (2.56)$$

since

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T e^{i\lambda t} dt = \begin{cases} 1 & \text{if } \lambda = 0 \\ 0 & \text{if } \lambda \neq 0. \end{cases}$$

The reader will note that the summation in (2.56) is taken over all  $j$  and  $k$  for which  $\lambda_j = \mu_k$ .  $X$  together with the inner product defined in (2.56) is an inner product space. This is known as the space of trigonometric polynomials on  $\mathbb{R}$ . Its completion  $H$  is a Hilbert space. The set  $\{u_r(t) = e^{irt}: r \in \mathbb{R}\}$  is an uncountable orthonormal set in the Hilbert space  $H$ , where  $H = \overline{X}$ , the closure of  $X$ .

- (ii) Let  $X$  be a nonempty set. Consider the Hilbert space  $L^2(X, \mathfrak{S}, \mu)$ , where  $\mathfrak{S}$  denotes the collection of all subsets of  $X$  and  $\mu$  is the counting measure on  $X$ ; that is, if  $E \in \mathfrak{S}$ ,  $\mu(E)$  is equal to the number of points in  $E$  when  $E$  is finite and is infinite if  $E$  is infinite. The space  $L^2(X, \mathfrak{S}, \mu)$  is denoted by  $\ell^2(X)$ .

Consider the subset of  $\ell^2(X)$  consisting of all characteristic functions of one point sets in  $X$ , i.e.  $\{\chi_{\{x\}}: x \in X\}$ . Observe that  $(\chi_{\{x\}}, \chi_{\{y\}}) = 0$  for  $x \neq y$  and  $\|\chi_{\{x\}}\| = 1$ . Suppose now  $x \neq y$  and consider the distance between  $\chi_{\{x\}}$  and  $\chi_{\{y\}}$ :

$$\|\chi_{\{x\}} - \chi_{\{y\}}\|_2^2 = \sum_{z \in X} |\chi_{\{x\}} - \chi_{\{y\}}|^2 = 2.$$

Thus

$$\|\chi_{\{x\}} - \chi_{\{y\}}\|_2 = \sqrt{2}.$$

The open balls  $S(\chi_{\{x\}}, 1/\sqrt{2})$  with centres  $\chi_{\{x\}}$  and radii  $1/\sqrt{2}$  are nonoverlapping, since no ball  $S(\chi_{\{x\}}, 1/\sqrt{2})$  contains a point of the set  $\{\chi_{\{x\}}: x \in X\}$  other than its centre. Now suppose that  $X$  is an uncountably infinite set. We claim that the space  $\ell^2(X)$  is nonseparable. Suppose not and let  $\{z_k\}$  be a countable dense set in  $\ell^2(X)$ . Each of the balls  $S(\chi_{\{x\}}, 1/\sqrt{2})$  will contain a point  $z_k$  of the countable dense set. Since the balls are nonoverlapping, the points contained in different balls cannot be identical. We thus have an injective map from  $X$  into the countable dense set, which is not possible as  $X$  is uncountable.

In view of the examples above, we consider orthonormal sets which are not necessarily countable. We begin with the following definition that formalises the remarks above Definition 2.7.5.

**Definition 2.9.6** Suppose  $\{x_\alpha: \alpha \in \Lambda\}$ , where  $\Lambda$  is an indexing set, is a collection of elements from a normed linear space  $X$ .  $\{x_\alpha: \alpha \in \Lambda\}$  is said to be **summable to**  $x \in X$ , written

$$\sum_{\alpha \in \Lambda} x_\alpha = x \quad \text{or} \quad \sum_{\alpha} x_\alpha = x,$$

if for all  $\varepsilon > 0$ , there exists some finite set of indices  $J_0 \subseteq \Lambda$ , such that for any finite set of indices  $J \supseteq J_0$ ,

$$\left\| \sum_{\alpha \in J} x_\alpha - x \right\| < \varepsilon.$$

This notion of summability can be easily reconciled with the usual notion of summability of a series when  $\Lambda$  consists of the natural numbers.

*Remarks 2.9.7*

- (i) Suppose  $S = \{x_\alpha: \alpha \in \Lambda\}$ , where  $\Lambda$  is an indexing set, is a collection of elements from the normed linear space  $\mathbb{R}$ . If  $0 \leq x_\alpha < \infty$  for each  $\alpha \in \Lambda$ , then  $\sum_{\alpha} x_\alpha$  is the supremum of the set of all finite sums  $x_{\alpha_1} + x_{\alpha_2} + \cdots + x_{\alpha_n}$ , where  $\alpha_1, \alpha_2, \dots, \alpha_n$  are distinct members of  $\Lambda$ . In this situation, the sum can be infinity, which is outside the space  $\mathbb{R}$ . However, if it is within the space, then the two notions of summation are identical. If  $x_\alpha = \infty$  for some  $\alpha \in \Lambda$ , then the sum  $\sum_{\alpha} x_\alpha$  is equal to infinity.
- (ii) It is easy to check that if  $\sum_{\alpha} x_\alpha = x$  and  $\sum_{\alpha} y_\alpha = y$ , then

$$\sum_{\alpha} (x_\alpha + y_\alpha) = x + y \quad \text{and} \quad \sum_{\alpha} \lambda x_\alpha = \lambda x, \quad \lambda \in \mathbb{F}.$$

The following proposition says though we are summing over an arbitrary indexing set, it is the sum over only a countable set of indices that matters.

**Proposition 2.9.8** *Let  $X$  be a Banach space over  $\mathbb{F}$  and suppose  $\{x_j; j \in \Lambda\} \subseteq X$ . The family  $\{x_j; j \in \Lambda\}$  is summable if, and only if, for every  $\varepsilon > 0$ , there exists a finite set  $J_0$  of indices such that  $\left\| \sum_{j \in J} x_j \right\| < \varepsilon$  whenever  $J$  is a finite set of indices disjoint from  $J_0$ .*

*If  $\{x_j\}$  is summable, then the set of those indices for which  $x_j \neq 0$  is countable.*

*Proof* If  $\{x_j\}$  is a summable family with sum  $x$ , then for every  $\varepsilon > 0$ , there exists a finite set  $J_0$  such that  $\left\| x - \sum_{j \in J_1} x_j \right\| < \varepsilon/2$  whenever  $J_1 \supseteq J_0$  and is finite. It follows that if  $J \cap J_0 = \emptyset$ , then

$$\left\| \sum_{j \in J} x_j \right\| = \left\| \sum_{j \in J \cup J_0} x_j - \sum_{j \in J_0} x_j \right\| \leq \left\| \sum_{j \in J \cup J_0} x_j - x \right\| + \left\| x - \sum_{j \in J_0} x_j \right\| < \varepsilon.$$

The reader will note that we have not used the completeness of  $X$  in the above argument.

If, conversely, the condition is satisfied, then for every positive integer  $n$ , there exists a finite set  $J_n$  such that  $\left\| \sum_{j \in J} x_j \right\| < 1/n$  whenever  $J$  is a finite set of indices and  $J \cap J_n = \emptyset$ . By replacing  $J_n$  by  $J_1 \cup J_2 \cup \dots \cup J_n$ ,  $n = 1, 2, \dots$ , we see that there is a sequence  $\{J_n\}$  of finite sets of indices which is increasing. If  $n < m$ , then

$$\left\| \sum_{j \in J_m} x_j - \sum_{j \in J_n} x_j \right\| = \left\| \sum_{j \in J_m \setminus J_n} x_j \right\| < 1/n$$

since  $(J_m \setminus J_n) \cap J_n = \emptyset$ . By the completeness of  $X$ , it follows that there exists  $x$  such that  $\left\| \sum_{j \in J_n} x_j - x \right\| \rightarrow 0$ . For  $\varepsilon > 0$ , there exists  $n_0 > 2/\varepsilon$  such that  $\left\| \sum_{j \in J_{n_0}} x_j - x \right\| < \varepsilon/2$ . If  $J$  is any finite set of indices containing  $J_{n_0}$ , then

$$\left\| \sum_{j \in J} x_j - x \right\| \leq \left\| \sum_{j \in J_{n_0}} x_j - x \right\| + \left\| \sum_{j \in J \setminus J_{n_0}} x_j \right\| < \varepsilon/2 + 1/n_0 < \varepsilon.$$

Consequently, the family  $\{x_j\}$  is summable with sum  $x$ .

Finally we show that  $x_j = 0$  for all but countably many  $j$ . If  $j$  is an index which does not belong to  $J_1 \cup J_2 \cup \dots$ , then  $\|x_j\| < 1/n$  for every  $n$ . The reader will note that we have not used the completeness of  $X$  in this argument. This completes the proof.  $\square$

If the sequence  $\{x_n\}_{n \geq 1}$  in a Hilbert space is orthogonal, then  $\sum_{n=1}^{\infty} x_n$  converges if, and only if,  $\sum_{n=1}^{\infty} \|x_n\|^2 < \infty$ . More generally, the following theorem holds.

**Theorem 2.9.9** *Let  $H$  be a Hilbert space and let  $\{x_j; j \in \Lambda\}$  be an orthogonal family in  $H$ , i.e.  $x_j \perp x_k$  for  $j \neq k$ . Then  $\sum_{j \in \Lambda} x_j$  converges if, and only if,  $\sum_{j \in \Lambda} \|x_j\|^2 < \infty$ . Moreover, if  $\sum_{j \in \Lambda} x_j = x$ , then  $\|x\|^2 = \sum_{j \in \Lambda} \|x_j\|^2$ .*

*Proof* If  $\{x_j\}$  is summable, then for every positive number  $\varepsilon$  there exists a finite set  $J_0$  such that  $\left\| \sum_{j \in J} x_j \right\| < \varepsilon$  whenever  $J \cap J_0 = \emptyset$ , and consequently

$$\sum_{j \in J} \|x_j\|^2 = \left\| \sum_{j \in J} x_j \right\|^2 < \varepsilon^2$$

whenever  $J \cap J_0 = \emptyset$ .

If, conversely,  $\sum_{j \in \Lambda} \|x_j\|^2 < \infty$ , then for every positive  $\varepsilon$  there exists a finite set  $J_0$  such that  $\sum_{j \in J} \|x_j\|^2 < \varepsilon^2$  (consequently  $\left\| \sum_{j \in J} x_j \right\|^2 < \varepsilon^2$ ) whenever  $J \cap J_0 = \emptyset$ . Summability now follows from the previous Theorem 2.9.8.

Observe that

$$\begin{aligned} \|x\|^2 &= (x, x) = \left( \sum_{j \in \Lambda} x_j, x \right) = \sum_{j \in \Lambda} \left( x_j, \sum_{k \in \Lambda} x_k \right) = \sum_{j \in \Lambda} \sum_{k \in \Lambda} (x_j, x_k) = \sum_{j \in \Lambda} (x_j, x_j) \\ &= \sum_{j \in \Lambda} \|x_j\|^2. \end{aligned}$$

□

The following general form of Bessel's Inequality holds.

**Theorem 2.9.10** (Bessel's Inequality) *Let  $S = \{x_\alpha; \alpha \in \Lambda\}$  be an orthonormal set in an inner product space  $H$  and let  $x \in H$ . Then we have*

$$\sum_{\alpha \in \Lambda} |(x, x_\alpha)|^2 \leq \|x\|^2.$$

*Proof* The inequality in Theorem 2.8.6 implies that for each finite set  $J \subseteq \Lambda$  of indices, we have

$$\sum_{\alpha \in J} |(x, x_\alpha)|^2 \leq \|x\|^2.$$

It now follows using Remark 2.9.7 that

$$\sum_{\alpha \in \Lambda} |(x, x_\alpha)|^2 = \sup \left\{ \sum_{\alpha \in J} |(x, x_\alpha)|^2 : J \subseteq \Lambda, J \text{ finite} \right\} \leq \|x\|^2.$$

□

*Remark 2.9.11* The set  $A = \{\alpha \in \Lambda : (x, x_\alpha) \neq 0\}$  is countable. By Bessel's Inequality,  $\sum_{\alpha \in \Lambda} |(x, x_\alpha)|^2 \leq \|x\|^2$ . The countability now follows from Theorem 2.9.8.

**Theorem 2.9.12** *Let  $\{x_\alpha : \alpha \in \Lambda\}$  be an orthonormal set in a Hilbert space  $H$ . For every  $x \in H$ , the vector  $y = \sum_{\alpha \in \Lambda} (x, x_\alpha)x_\alpha$  exists in  $H$  and  $x - y \perp x_\alpha$  for every  $\alpha \in \Lambda$ .*

*Proof* By Bessel's Inequality 2.9.10, there is a countable set of  $x_\alpha$  for which  $(x, x_\alpha) \neq 0$ . Arrange them as a sequence  $x_1, x_2, \dots$ . Let  $\varepsilon > 0$  be given. Then

$$\begin{aligned} \left\| \sum_{i=n}^{n+k} (x, x_i)x_i \right\|^2 &= \left( \sum_{i=n}^{n+k} (x, x_i)x_i, \sum_{j=n}^{n+k} (x, x_j)x_j \right) \\ &= \sum_{i=n}^{n+k} \sum_{j=n}^{n+k} (x, x_i) \overline{(x, x_j)} (x_i, x_j) \\ &= \sum_{i=n}^{n+k} |(x, x_i)|^2 < \varepsilon \end{aligned}$$

for large  $n$  and any positive integer  $k$ , again using Bessel's Inequality. It follows that the sequence of partial sums  $\{\sum_{i=1}^n (x, x_i)x_i\}_{n \geq 1}$  is Cauchy in  $H$ , and  $H$  being a Hilbert space,  $y = \sum_{i=1}^\infty (x, x_i)x_i$  exists in  $H$  and equals  $\sum_{\alpha \in \Lambda} (x, x_\alpha)x_\alpha$ . Note that the foregoing argument is valid whether  $\mathbb{F} = \mathbb{C}$  or  $\mathbb{R}$ , as in the latter case  $\overline{(x, x_j)} = (x, x_j)$ .

It remains to show that  $x - y \perp x_\alpha$  for every  $\alpha \in \Lambda$ . For each  $n$ , let  $y_n = \sum_{k=1}^n (x, x_k)x_k$ . We first prove that  $(x - y_n, x_\alpha) = 0$  for those  $\alpha$  for which  $(x, x_\alpha) = 0$  and any  $n$ . Note that  $x_\alpha$  cannot appear in the representation of  $y_n$  for any  $n$ . Therefore,  $(x_k, x_\alpha) = 0$  for all  $k$  and hence

$$(x - y_n, x_\alpha) = (x, x_\alpha) - \sum_{k=1}^n (x, x_k)(x_k, x_\alpha) = 0.$$

Next we prove  $(x - y_n, x_\alpha) = 0$  for those  $\alpha$  for which  $(x, x_\alpha) \neq 0$  and sufficiently large  $n$ . Note that  $x_\alpha$  must appear in the representation of  $y_n$  for sufficiently large  $n$ . Therefore,

$$(x - y_n, x_\alpha) = (x, x_\alpha) - \sum_{k=1}^n (x, x_k)(x_k, x_\alpha) = (x, x_\alpha) - (x, x_\alpha) = 0$$

for sufficiently large  $n$ .

Now, for sufficiently large  $n$ ,

$$\begin{aligned} |(x - y, x_\alpha)| &\leq |(x - y_n, x_\alpha)| + |(y_n - y, x_\alpha)| \\ &\leq 0 + \|y_n - y\| \|x_\alpha\| \\ &= \|y_n - y\| \text{ (using orthonormality of } \{x_\alpha : \alpha \in \Lambda\}) \end{aligned}$$

Since  $\|y_n - y\| \rightarrow 0$  as  $n \rightarrow \infty$ , it follows that  $x - y \perp x_\alpha$  for every  $\alpha \in \Lambda$ .  $\square$

We next investigate the problem of writing an arbitrary element  $x$  in a Hilbert space  $H$  as a limit of linear combinations of elements of an orthonormal set. We begin with a definition.

**Definition 2.9.13** Let  $H$  be a Hilbert space and  $S = \{x_\alpha : \alpha \in \Lambda\}$  be an orthonormal set in  $H$ . We say that  $S$  is a **basis (orthonormal basis)** in  $H$  if for every  $x \in H$ , the following holds:

$$x = \sum_{\alpha \in \Lambda} (x, x_\alpha) x_\alpha.$$

The following theorem provides a characterisation of a basis in a Hilbert space.

**Theorem 2.9.14** *If  $H$  is a Hilbert space, then  $S = \{x_\alpha : \alpha \in \Lambda\}$  consisting of orthonormal vectors in  $H$  is a basis if, and only if,  $S$  is a complete orthonormal system of vectors.*

*Proof* Suppose  $S$  is a basis in  $H$ . Then if  $x \in H$  satisfies  $(x, x_\alpha) = 0$ ,  $\alpha \in \Lambda$ , then the definition of the basis gives

$$x = \sum_{\alpha \in \Lambda} (x, x_\alpha) x_\alpha = 0.$$

Thus,  $S$  is complete by Theorem 2.9.3.

On the other hand, suppose that  $S$  is complete in  $H$ . Let  $\beta \in \Lambda$ . Then for any  $x \in H$ , the sum  $\sum_{\alpha \in \Lambda} (x, x_\alpha) x_\alpha$  exists by Theorem 2.9.12 and

$$\left( x - \sum_{\alpha \in \Lambda} (x, x_\alpha) x_\alpha, x_\beta \right) = (x, x_\beta) - \sum_{\alpha \in \Lambda} (x, x_\alpha) (x_\alpha, x_\beta) = (x, x_\beta) - (x, x_\beta) = 0,$$

using the fact that  $S = \{x_\alpha : \alpha \in \Lambda\}$  consists of orthonormal vectors. Thus, the vector  $x - \sum_{\alpha \in \Lambda} (x, x_\alpha) x_\alpha$  is orthogonal to  $x_\beta$  for every  $\beta \in \Lambda$ . The hypothesis, together with Theorem 2.9.3, now implies

$$x = \sum_{\alpha \in \Lambda} (x, x_\alpha) x_\alpha,$$

i.e.  $S$  is a basis in  $H$ . □

### Examples 2.9.15

- (i) The set  $e_1 = (1, 0, 0, \dots)$ ,  $e_2 = (0, 1, 0, 0, \dots)$ , ... is a complete orthonormal set (basis) in  $\ell^2$ . Indeed, if  $x = (x_1, x_2, \dots) \in \ell^2$  and  $(x, e_j) = 0$ ,  $j = 1, 2, \dots$ , then  $x_j = 0$ ,  $j = 1, 2, \dots$  and so  $x = 0$ . Moreover, if  $x = (x_1, x_2, \dots) \in \ell^2$  then  $x = \sum_{i=1}^{\infty} (x, e_i) e_i$ , where the partial sums converge in the  $\ell^2$ -norm:  $\|\sum_{i=1}^n (x, e_i) e_i - x\|^2 = \sum_{i=n+1}^{\infty} |x_i|^2$  is small for large  $n$ .
- (ii) [Cf. Examples 2.9.5(ii)] Let  $X$  be a non-denumerable set. The set  $\ell^2(X) = L^2(X, \mathfrak{S}, \mu)$ , where  $\mathfrak{S}$  denotes the collection of all subsets of  $X$  and  $\mu$  is the counting measure on  $X$ , is a nonseparable Hilbert space. The set  $\{\chi_{\{x\}} : x \in X\}$  of characteristic functions is an uncountable orthonormal set in  $\ell^2(X)$ . In fact, it is a complete orthonormal set. If  $f \in \ell^2(X)$  and for  $x \in X$ ,  $(f, \chi_{\{x\}}) = 0$ , then  $\sum_{y \in X} f(y) \chi_{\{x\}}(y) = 0$ , which implies  $f(x) = 0$ . So  $f$  is the identically zero function [see Theorem 2.9.3].
- (iii) The Rademacher system is not complete. The function  $f(x) = \cos 2\pi x$  is orthogonal to all the Rademacher functions [see Example 2.8.13(v)].

The following theorem provides various characterisations of complete orthonormal sets and helps decide which orthonormal sets are complete. Some of the characterisations have already been described.

**Theorem 2.9.16** *Let  $S = \{x_\alpha : \alpha \in \Lambda\}$  be an orthonormal set in Hilbert space  $H$ . Each of the following conditions implies the other five:*

- (a)  $S$  is a complete orthonormal set in  $H$ ;
- (b)  $x \perp S$  implies  $x = 0$ ;
- (c)  $x \in H$  implies  $x = \sum_{\alpha \in \Lambda} (x, x_\alpha) x_\alpha$ ; that is,  $S$  is a basis in  $H$ ;
- (d)  $\|x\|^2 = \sum_{\alpha \in \Lambda} |(x, x_\alpha)|^2$  for each  $x \in H$ ; (**Parseval's Identity**)
- (e) for  $x, y \in H$ ,  $(x, y) = \sum_{\alpha \in \Lambda} (x, x_\alpha) \overline{(y, x_\alpha)}$ ;
- (f)  $\overline{[S]} = H$ ; that is, the smallest subspace of  $H$  containing  $S$  is dense in  $H$ .

The equality in (c) means that the right-hand side has only a countable number of nonzero terms, and every rearrangement of this series converges to  $x$  [Definition 2.9.6]. The equations in (d) and (e) are to be interpreted analogously.

Of course, (d) is a special case of (e).

*Proof* The equivalence of (a) and (b) has been proved [Theorem 2.9.3]. So also the equivalence of (a) and (c) [Theorem 2.9.14]. We shall prove that (b)  $\Rightarrow$  (f)  $\Rightarrow$  (d)  $\Rightarrow$  (e)  $\Rightarrow$  (b).

(b) implies (f). Let  $M = \overline{[S]}$ . Since  $[S]$  is a subspace, so is  $M$ . (For  $x, y \in M$ , there exist sequences  $\{x_n\}_{n \geq 1}$  and  $\{y_n\}_{n \geq 1}$  such that  $x_n \rightarrow x$  and  $y_n \rightarrow y$ ; then

$x_n + y_n \rightarrow x + y$  and  $\lambda x_n \rightarrow \lambda x$ ,  $\lambda \in \mathbb{F}$ .) Suppose  $[S]$  is not dense in  $H$ . Then,  $M \neq H$  so that there exists a nonzero vector  $x$  in  $H$  which is not in  $M$ . The vector  $y = \sum_{\alpha \in \Lambda} (x, x_\alpha) x_\alpha$  exists in  $H$  and  $x - y \perp x_\alpha$  for every  $\alpha \in \Lambda$  [Theorem 2.9.12]. Moreover,  $x \neq y$  since  $y \in M$  and  $x \notin M$  and hence  $x - y \neq 0$ . This contradicts (b).

(f) implies (d). Suppose (f) holds. For  $x \in H$  and  $\varepsilon > 0$ , there exists a finite set  $\{x_{\alpha_1}, x_{\alpha_2}, \dots, x_{\alpha_n}\}$  such that some linear combinations of these vectors have distance less than  $\varepsilon$  from  $x$ . By Remark 2.8.8(ii), the vector  $z = \sum_{i=1}^n (x, x_{\alpha_i}) x_{\alpha_i}$  provides the best approximation to the vector  $x$  in the linear span of  $\{x_{\alpha_1}, x_{\alpha_2}, \dots, x_{\alpha_n}\}$ ; so  $\|x - z\| < \varepsilon$ , and hence  $\|x\| < \|z\| + \varepsilon$  which implies  $(\|x\| - \varepsilon)^2 < \|z\|^2 = \sum_{i=1}^n |(x, x_{\alpha_i})|^2 \leq \sum_{i=1}^\infty |(x, x_{\alpha_i})|^2$ . Since  $\varepsilon > 0$  is arbitrary, we obtain  $\|x\|^2 \leq \sum_{i=1}^\infty |(x, x_{\alpha_i})|^2$ . The result now follows using Bessel's Inequality 2.9.10.

(d) implies (e). Note that (d) can be written as

$$(x, x) = \left( \sum_{\alpha \in \Lambda} (x, x_\alpha) x_\alpha, \sum_{\alpha \in \Lambda} (x, x_\alpha) x_\alpha \right).$$

Fix  $x, y \in H$ . If (d) holds, then

$$(x + \lambda y, x + \lambda y) = \left( \sum_{\alpha \in \Lambda} (x + \lambda y, x_\alpha) x_\alpha, \sum_{\alpha \in \Lambda} (x + \lambda y, x_\alpha) x_\alpha \right)$$

for any scalar  $\lambda$ . Hence,

$$\bar{\lambda}(x, y) + \lambda(y, x) = \bar{\lambda} \left( \sum_{\alpha \in \Lambda} (x, x_\alpha) x_\alpha, \sum_{\alpha \in \Lambda} (y, x_\alpha) x_\alpha \right) + \lambda \left( \sum_{\alpha \in \Lambda} (y, x_\alpha) x_\alpha, \sum_{\alpha \in \Lambda} (x, x_\alpha) x_\alpha \right) \quad (2.57)$$

Taking  $\lambda = 1$  and  $\lambda = i$ , (2.57) shows that the real and imaginary parts of  $(x, y)$  and  $(\sum_{\alpha \in \Lambda} (x, x_\alpha) x_\alpha, \sum_{\alpha \in \Lambda} (y, x_\alpha) x_\alpha)$  are equal. Hence

$$\begin{aligned} (x, y) &= \left( \sum_{\alpha \in \Lambda} (x, x_\alpha) x_\alpha, \sum_{\alpha \in \Lambda} (y, x_\alpha) x_\alpha \right) \\ &= \sum_{\alpha \in \Lambda} (x, x_\alpha) \overline{(x_\alpha, y)}. \end{aligned}$$

(e) implies (b). Finally, if (b) fails to be true, there exists a vector  $z \neq 0$  so that  $(z, x_\alpha) = 0$  for all  $\alpha \in \Lambda$ . If  $x = y = z$ , the  $\|z\|^2 = (x, y) \neq 0$  but  $(z, x_\alpha) \overline{(x_\alpha, z)} = 0$ . Hence (e) fails to hold. Thus, (e) implies (b) and the proof is complete.  $\square$

To deal with completeness of orthonormal sets in the next few examples, we will use their equivalent descriptions provided in Theorem 2.9.16.

*Examples 2.9.17*

- (i) In the completion  $H$  of the inner product space of trigonometric polynomials [see Example 2.9.5(i)], the uncountable orthonormal set  $\{u_r(t) = \exp(irt) : r \in \mathbb{R}\}$  is complete since  $\overline{\{u_r\}} = H$  [equivalence of (a) and (d) in Theorem 2.9.16].
- (ii) Let  $H = L^2[-1, 1]$  and for  $n = 0, 1, 2, \dots$ , let  $P_n$  denote the Legendre polynomial of degree  $n$ . Note that  $P_n$  is obtained by applying the Gram–Schmidt orthonormalisation process to the linearly independent vectors  $\{1, t, t^2, \dots, t^n\}$ . Moreover,

$$\text{span}\{1, t, t^2, \dots, t^n\} = \text{span}\{P_0, P_1, \dots, P_n\}. \quad (2.58)$$

This is true for each  $n$ . Let  $x \in H$  and  $\varepsilon > 0$ . By Example 2.5.1, there exists  $y \in C[-1, 1]$  such that  $\|x - y\| < \varepsilon$ . By Weierstrass' Theorem, there is a polynomial  $Q(t)$  such that  $|y(t) - Q(t)| < \varepsilon$  for all  $t \in [-1, 1]$ . Then

$$\|y - Q\|_2^2 = \int_{-1}^1 |y(t) - Q(t)|^2 dt < 2\varepsilon^2.$$

Thus

$$\|x - Q\|_2 \leq \|x - y\|_2 + \|y - Q\|_2 < (1 + \sqrt{2})\varepsilon.$$

This shows that the set of all polynomials on  $[-1, 1]$  is dense in  $H$ .

In view of (2.58), the set  $\{P_0, P_1, \dots\}$  constitutes a complete orthonormal basis [Theorem 2.9.16(f)].

- (iii) Let  $H = L^2([-\pi, \pi], \frac{dt}{2\pi})$  and for  $n = 0, \pm 1, \pm 2, \dots$ ,

$$u_n(t) = e^{int}, \quad t \in [-\pi, \pi].$$

Then  $\{u_n : n = 0, \pm 1, \pm 2, \dots\}$  is an orthonormal set in  $H$ . Indeed,

$$(u_n, u_m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-m)t} dt = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{if } n \neq m \end{cases}$$

The orthonormal set  $\{u_n: n = 0, \pm 1, \pm 2, \dots\}$  is usually called *trigonometric system*. For  $x \in H$  and  $n = 0, \pm 1, \pm 2, \dots$ ,

$$(x, u_n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} x(t) e^{-int} dt = \hat{x}(n),$$

where  $\hat{x}(n)$  is the  $n$ th Fourier coefficient of  $x$ .

We shall show that if  $\hat{x}(n) = 0, n = 0, \pm 1, \pm 2, \dots$ , then  $x = 0$  a.e. This will prove that the trigonometric system is complete in  $L^2([-\pi, \pi], \frac{dt}{2\pi})$ .

Set

$$y(t) = \frac{1}{2\pi} \int_{-\pi}^t x(s) ds.$$

Since  $L^2([-\pi, \pi], \frac{dt}{2\pi}) \subseteq L^1([-\pi, \pi], \frac{dt}{2\pi})$ , it is evident that  $y$  is a well-defined absolutely continuous function on  $[-\pi, \pi]$  [see 1–5]. In particular,  $y \in L^2([-\pi, \pi], \frac{dt}{2\pi})$ . Moreover,  $y(-\pi) = 0$  and  $y(\pi) = 0$ , using the fact that  $\hat{x}(0) = 0$  by hypothesis. Let  $a$  be any constant. On integrating by parts, we obtain

$$\int_{-\pi}^{\pi} [y(t) - a] e^{int} dt = 0, \quad n = \pm 1, \pm 2, \dots \quad (2.59)$$

Choose  $a$  so that (2.59) holds for  $n = 0$  as well. Since  $y(t) - a$  is a continuous periodic function, for  $\varepsilon > 0$ , there is a trigonometric polynomial

$$T(t) = \sum_{k=-n}^n c_k e^{ikt}$$

such that

$$\sup\{|y(t) - a - T(t)|: t \in [-\pi, \pi]\} < \varepsilon,$$

using Weierstrass' Theorem.

Now using (2.59) and the choice of  $a$ , we obtain

$$\begin{aligned}
\int_{-\pi}^{\pi} |y(t) - a|^2 dt &= \int_{-\pi}^{\pi} (y(t) - a) \left( \overline{y(t)} - \bar{a} - T(t) \right) dt \\
&\leq \varepsilon \int_{-\pi}^{\pi} |y(t) - a| dt \\
&\leq \varepsilon \left[ \int_{-\pi}^{\pi} |y(t) - a|^2 dt \right]^{\frac{1}{2}} \left[ \int_{-\pi}^{\pi} dt \right]^{\frac{1}{2}},
\end{aligned}$$

which implies

$$\int_{-\pi}^{\pi} |y(t) - a|^2 dt \leq 2\pi\varepsilon^2.$$

Thus,  $y(t)$  is constant and  $x(t) = 0$  almost everywhere. This completes the proof.

*Remarks* (a) In the above proof, we have used the fact that  $x \in L^1\left([-\pi, \pi], \frac{dt}{2\pi}\right)$  and have proved that if  $\hat{x}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} x(t) e^{-int} dt = 0$  for  $n = 0, \pm 1, \pm 2, \dots$ , then  $x = 0$  a.e.

(b) We put some of the results proved for abstract Hilbert spaces in the present setting of  $L^2\left([-\pi, \pi], \frac{dt}{2\pi}\right)$ . For  $x \in L^2\left([-\pi, \pi], \frac{dt}{2\pi}\right)$ , associate a function  $\hat{x}$  defined on  $\mathbb{Z}$ , the set of integers. The Fourier series of  $x$  is

$$\sum_{n=-\infty}^{\infty} \hat{x}(n) e^{int} \quad (2.60)$$

and its partial sums are

$$S_N = \sum_{n=-N}^N \hat{x}(n) e^{int}, \quad N = 0, 1, 2, \dots,$$

The Parseval identity asserts

$$\sum_{n=-\infty}^{\infty} \hat{x}(n) \overline{\hat{y}(n)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} x(t) \overline{y(t)} dt, \quad x, y \in L^2\left([-\pi, \pi], \frac{dt}{2\pi}\right). \quad (2.61)$$

The Fourier series (2.60) converges to  $x$  in the  $L^2$ -norm:

$$\lim_N \|x - S_N\|_2 = 0. \quad (2.62)$$

(c) **(The Riemann–Lebesgue Lemma)** If  $x \in L^2([-\pi, \pi], \frac{dt}{2\pi})$ , then

$$\int_{-\pi}^{\pi} x(t) e^{-int} dt \rightarrow 0 \quad \text{as } n \rightarrow \pm\infty.$$

Indeed, the Parseval's identity (2.61) gives  $\sum_{n=-\infty}^{\infty} |\hat{x}(n)|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |x(t)|^2 dt < \infty$ . (d) The relation (2.62) leads to the question whether the Fourier series of  $x \in L^2([-\pi, \pi], \frac{dt}{2\pi})$  tends to  $x$  pointwise. This is not true even for a continuous function, as was demonstrated by du Bois-Reymond in 1876. However, Fejér proved in 1900 that the Fourier series of a continuous function is Cesàro summable and the sum is the function itself. For a function  $x \in L^2([-\pi, \pi], \frac{dt}{2\pi})$ , Lusin's conjecture that  $\{S_N\}_{N \geq 0}$ , where  $S_N = \sum_{n=-N}^N \hat{x}(n) e^{int}$  converges to  $x$  pointwise a.e. was proved by Carleson in 1966.

(iv) A complete orthonormal system for the space  $H = L^2(0, \infty)$  is given by the Laguerre functions

$$v_n(t) = \frac{1}{n!} \exp\left(-\frac{t}{2}\right) L_n(t), \quad (2.63)$$

where

$$L_n(t) = (-1)^n \exp(t) \frac{d^n}{dt^n} (t^n \exp(-t)), \quad n = 0, 1, 2, \dots$$

In fact,  $\{v_n(t)\}_{n \geq 0}$  constitute an orthonormal set in  $H$  [Example 2.8.13(iv)]. In order to show that the system (2.63) is complete in  $H$ , it will be enough to show that if  $f \in H$  and  $\int_0^\infty f(t) \exp(-\frac{t}{2}) L_n(t) dt = 0$ ,  $n = 0, 1, 2, \dots$ , then  $f = 0$  a.e. Let

$$g(t) = f(t) \exp\left(-\frac{t}{2}\right), \quad 0 < t < \infty.$$

Since  $f \in L^2(0, \infty)$  and  $\exp(-\frac{t}{2}) \in L^2(0, \infty)$ , it follows that  $g \in L^1(0, \infty)$ . Indeed, by the Cauchy–Schwarz Inequality,

$$\begin{aligned}
\int_0^{\infty} |g(t)| dt &= \int_0^{\infty} \left| f(t) \exp\left(-\frac{t}{2}\right) \right| dt \\
&\leq \left[ \int_0^{\infty} |f(t)|^2 dt \right]^{\frac{1}{2}} \left[ \int_0^{\infty} \exp(-t) dt \right]^{\frac{1}{2}} \\
&\leq \|f\|_2.
\end{aligned}$$

Each  $L_n$  is a polynomial of degree  $n$  [see (2.64) of Examples 2.8.13(iv)]. Therefore, each  $t^n$  is a linear combination of  $L_0, \dots, L_n$ . Thus, we need only to show that

$$\int_0^{\infty} g(t) t^n dt = 0, \quad n = 0, 1, 2, \dots \text{ implies } g(t) = 0 \quad \text{a.e.} \quad (2.64)$$

Now consider

$$\begin{aligned}
\Phi(z) &= \int_0^{\infty} \exp(-tz) g(t) dt \\
&= \int_0^{\infty} \exp(-tx) \exp(-ity) g(t) dt, \quad \Re z > 0.
\end{aligned} \quad (2.65)$$

We shall show that  $\Phi(z)$  is holomorphic in  $\Re z > 0$ .

Since  $g \in L^1(0, \infty)$  and  $|\exp(-ity)| = 1$ , the integral in (2.65) exists as a Lebesgue integral. Moreover,  $\Phi(z)$  is continuous in  $\Re z > 0$ . Indeed, if  $z_n \rightarrow z$  in  $\Re z > 0$ , then  $g(t) \exp(-tz_n) \rightarrow g(t) \exp(-tz)$ . Both the sequence of functions and the limit function are integrable and are dominated by the integrable function  $g(t)$ , an application of the Lebesgue Dominated Convergence Theorem 1.3.9 proves the assertion. If  $\Delta$  denotes the boundary of any closed triangle in  $\Re z > 0$ , then

$$\begin{aligned}
\oint_{\Delta} \Phi(z) dz &= \oint_{\Delta} \left( \int_0^{\infty} g(t) \exp(-tz) dt \right) dz \\
&= \int_0^{\infty} g(t) \left( \oint_{\Delta} \exp(-tz) dz \right) dt \quad [\text{Fubini's Theorem}] \\
&= \int_0^{\infty} g(t) \cdot 0 dt \quad [\text{Cauchy's Theorem}] = 0.
\end{aligned}$$

It now follows by using Morera's Theorem that  $\Phi(z)$  is holomorphic in  $\Re z > 0$ .

On using integration by parts and induction on  $n$ , we see that

$$\int_0^{\infty} \exp(-t) t^{2n} dt = (2n)! \leq (2^n n!)^2. \quad (2.66)$$

We next consider the series

$$\sum_{n=0}^{\infty} s^n \frac{1}{n!} \int_0^{\infty} g(t) t^n dt, \quad s > 0, \quad (2.67)$$

and show that the series converges for  $0 \leq s < \frac{1}{2}$  to the function  $\Phi(s)$ .

$$\begin{aligned} \left| \sum_{n=0}^{\infty} s^n \frac{1}{n!} \int_0^{\infty} g(t) t^n dt \right| &\leq \sum_{n=0}^{\infty} s^n \frac{1}{n!} \int_0^{\infty} |f(t)| \exp\left(-\frac{t}{2}\right) t^n dt \\ &\leq \sum_{n=0}^{\infty} s^n \frac{1}{n!} \|f\|_2 \cdot [(2n)!]^{1/2}, \end{aligned}$$

using the Cauchy–Schwarz Inequality. From (2.66), we have

$$\leq \|f\|_2 \cdot \sum_{n=0}^{\infty} (2s)^n.$$

And the series on the right converges for  $0 \leq s < \frac{1}{2}$ .

Note that  $\exp(-st)g(t) = \sum_{n=0}^{\infty} (-1)^n s^n \frac{1}{n!} g(t) t^n$ , and

$$\begin{aligned} \sum_{n=0}^{\infty} \int_0^{\infty} \left| (-1)^n s^n \frac{1}{n!} g(t) t^n \right| dt &\leq \sum_{n=0}^{\infty} \int_0^{\infty} s^n |g(t)| t^n dt \\ &= \sum_{n=0}^{\infty} s^n \frac{1}{n!} \int_0^{\infty} |f(t)| \exp\left(-\frac{t}{2}\right) t^n dt \\ &\leq \sum_{n=0}^{\infty} s^n \frac{1}{n!} \left[ \int_0^{\infty} |f(t)|^2 dt \int_0^{\infty} t^{2n} \exp(-t) dt \right]^{\frac{1}{2}} \quad (\text{Cauchy–Schwarz}) \\ &\leq \sum_{n=0}^{\infty} s^n \frac{1}{n!} \|f\|_2 \cdot 2^n \cdot n! \quad (\text{using (2.66)}) \\ &= \sum_{n=0}^{\infty} \|f\|_2 (2s)^n < \infty \quad \text{for } 0 \leq s < \frac{1}{2}. \end{aligned}$$

Then, on using Corollary 1.3.10,

$$\sum_{n=0}^{\infty} (-1)^n s^n \frac{1}{n!} g(t) t^n \in L^1(0, \infty)$$

and

$$\Phi(s) = \int_0^{\infty} \exp(-st) g(t) dt = \sum_{n=0}^{\infty} \int_0^{\infty} (-1)^n s^n g(t) t^n dt.$$

On using the hypothesis, we obtain  $\Phi(s) = 0$  for all  $s \geq 0$ . Now,

$$\Phi(s) = \int_0^1 t^{s-1} g(-\ln t) dt, \quad s \geq 0.$$

It may be observed that  $t^{-1}g(-\ln t)$  is in  $L^1[0, 1]$ , using the substitution  $u = -\ln t$ . Using Proposition 1.3.11, it now follows that  $g(-\ln t) = 0$  a.e. on  $(0, 1)$ , which implies [see Problem 2.9.P8] that  $g(t) = 0$  a.e. on  $(0, \infty)$ . This completes the proof.

- (v) A complete orthonormal system for the space  $H = L^2(-\infty, \infty)$  is given by the Hermite functions

$$v_n(t) = \frac{H_n(t) \exp\left(-\frac{t^2}{2}\right)}{(2^n \cdot n! \pi^{1/2})^{1/2}}, \quad n = 0, 1, 2, \dots, \quad (2.68)$$

where

$$H_n(t) = (-1)^n \exp(t^2) \exp^{(n)}(-t^2).$$

In fact,  $\{v_n(t)\}_{n \geq 0}$  constitute an orthonormal set in  $L^2(-\infty, \infty)$  [Example 2.8.13 (iii)]. In order to show that the system (2.68) is complete in  $H$ , it will be enough to show that if  $f \in L^2(-\infty, \infty)$  and  $\int_{-\infty}^{\infty} f(t) \exp\left(-\frac{t^2}{2}\right) H_n(t) dt = 0$ , or equivalently,  $\int_{-\infty}^{\infty} f(t) \exp\left(-\frac{t^2}{2}\right) t^n dt = 0$ , for  $n = 0, 1, 2, \dots$  implies  $f = 0$  a.e. on  $(-\infty, \infty)$ . The equivalence follows exactly as in (iv) above, since  $H_n$  is a polynomial of degree  $n$ .

We consider the function

$$F(x) = \int_{-\infty}^{\infty} f(t) e^{-itx} \exp\left(-\frac{t^2}{2}\right) dt, \quad -\infty < x < \infty.$$

This integral exists, since  $f \in L^2(-\infty, \infty)$ ,  $\exp(-\frac{t^2}{2}) \in L^2(-\infty, \infty)$  and  $|e^{-itx}| = 1$ . In fact,

$$\begin{aligned} \int_{-\infty}^{\infty} \left| f(t) e^{-itx} \exp\left(-\frac{t^2}{2}\right) \right| dt &= \int_{-\infty}^{\infty} \left| f(t) \exp\left(-\frac{t^2}{2}\right) \right| dt \\ &\leq \left[ \int_{-\infty}^{\infty} |f(t)|^2 dt \right]^{\frac{1}{2}} \left[ \int_{-\infty}^{\infty} \exp(-t^2) dt \right]^{\frac{1}{2}} < \infty, \end{aligned}$$

using the Cauchy–Schwarz Inequality.

We write

$$f(t) e^{-itx} = \sum_{n=0}^{\infty} (-i)^n \frac{x^n}{n!} f(t) t^n$$

and observe that

$$\begin{aligned} &\int_{-\infty}^{\infty} \sum_{n=0}^{\infty} \frac{|x|^n}{n!} |f(t)| |t^n| \exp\left(-\frac{t^2}{2}\right) dt \\ &= \int_{-\infty}^{\infty} |f(t)| \exp(|xt|) \exp\left(-\frac{t^2}{2}\right) dt \\ &= \int_{-\infty}^{\infty} |f(t)| \exp\left(-\frac{t^2}{4}\right) \exp(|xt|) \exp\left(-\frac{t^2}{4}\right) dt \\ &\leq \left[ \int_{-\infty}^{\infty} |f(t)|^2 \exp\left(-\frac{t^2}{2}\right) dt \right]^{\frac{1}{2}} \left[ \int_{-\infty}^{\infty} \exp(2|xt|) \exp\left(-\frac{t^2}{2}\right) dt \right]^{\frac{1}{2}} < \infty, \end{aligned}$$

since

$$\int_{-\infty}^{\infty} |f(t)|^2 \exp\left(-\frac{t^2}{2}\right) dt \leq \int_{-\infty}^{\infty} |f(t)|^2 dt = \|f\|_2^2,$$

and

$$\begin{aligned}
 \int_{-\infty}^{\infty} \exp(2|x|t) \exp\left(-\frac{t^2}{2}\right) dt &= 2 \int_0^{\infty} \exp(2|x|t) \exp\left(-\frac{t^2}{2}\right) dt \\
 &= 2 \int_0^{\infty} \exp\left(-\frac{t^2}{2} + 2|x|t - 2x^2\right) \exp(2x^2) dt \\
 &= 2 \exp(2x^2) \int_0^{\infty} \exp\left(-\frac{1}{2}(t - 2|x|)^2\right) dt < \infty.
 \end{aligned}$$

Using Corollary 1.3.10, it follows that

$$\begin{aligned}
 F(x) &= \int_{-\infty}^{\infty} f(t) e^{-itx} \exp\left(-\frac{t^2}{2}\right) dt \\
 &= \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} (-i)^n \frac{x^n}{n!} f(t) \exp\left(-\frac{t^2}{2}\right) t^n dt \\
 &= \sum_{n=0}^{\infty} (-i)^n \frac{x^n}{n!} \int_{-\infty}^{\infty} f(t) \exp\left(-\frac{t^2}{2}\right) t^n dt = 0,
 \end{aligned}$$

i.e.

$$\int_{-\infty}^{\infty} f(t) e^{-itx} \exp\left(-\frac{t^2}{2}\right) dt = 0$$

for all real  $x$ . It follows on using Proposition 1.3.12 that  $f = 0$  a.e., which was to be proved.

- (vi) The set  $\varphi_n(z) = \sqrt{\frac{n}{\pi}} z^{n-1}$ ,  $n = 1, 2, \dots$  is an orthonormal set in  $H = A(D)$ , where  $D = \{z \in \mathbb{C}: |z| < 1\}$  [see 2.8.4(iii)]. We shall show that the Parseval formula

$$\sum_{n=1}^{\infty} |(f, \varphi_n)|^2 = \iint_{|z| < 1} |f(z)|^2 dz, \quad f \in L^2(D)$$

holds, thereby establishing the completeness of  $\{\varphi_n(z)\}_{n \geq 1}$  in  $A(D)$  [Theorem 2.9.16].

For  $f \in A(D)$ , the Fourier coefficients are given by

$$\begin{aligned}
\gamma_n = (f, \varphi_n) &= \sqrt{\frac{n}{\pi}} \iint_D f(z) \bar{z}^{n-1} dx dy \\
&= \lim_{r \rightarrow 1} \sqrt{\frac{n}{\pi}} \iint_{|z| < r} f(z) \bar{z}^{n-1} dx dy.
\end{aligned}$$

On applying the complex Green's formula [29, p. 124], we obtain

$$\gamma_j = \lim_{r \rightarrow 1} \frac{1}{2i} \sqrt{\frac{n}{\pi}} \int_{|z|=r} f(z) \frac{\bar{z}^n}{n} dz.$$

Since  $|z|^2 = r^2$ , we have  $\bar{z} = r^2 z^{-1}$  and so

$$\gamma_n = \lim_{r \rightarrow 1} \frac{1}{\sqrt{n\pi}} \frac{r^{2n}}{2i} \int_{|z|=r} f(z) \frac{dz}{z^n}. \quad (2.69)$$

Now if

$$f(z) = a_0 + a_1 z + \dots, \quad |z| < 1$$

is the power series expansion of  $f$ , then

$$a_{n-1} = \frac{1}{2\pi i} \int_{|z|=1} \frac{f(z)}{z^n} dz. \quad (2.70)$$

From (2.69) to (2.70), we obtain

$$\begin{aligned}
\gamma_n &= \lim_{r \rightarrow 1} \frac{1}{\sqrt{n\pi}} r^{2n} \pi a_{n-1} \\
&= \sqrt{\frac{\pi}{n}} a_{n-1}, \quad n = 1, 2, \dots
\end{aligned} \quad (2.71)$$

Also,

$$\iint_{|z| < 1} |f(z)|^2 dx dy = \pi \sum_{n=1}^{\infty} \frac{|a_{n-1}|^2}{n} \quad (2.72)$$

[see (2.70) above Definition 2.6.2].

From (2.71) to (2.72), it follows that Parseval's formula holds. The argument is therefore complete.

**Theorem 2.9.18** *Any two complete orthonormal sets in a given Hilbert space  $H \neq \{0\}$  have the same cardinal number.*

*Proof* Let  $H$  be a Hilbert space of dimension  $n$  and  $A$  be any complete orthonormal set in  $H$ . It consists of linearly independent vectors and therefore can have at most  $n$  vectors in it. We shall argue that it contains precisely  $n$  vectors: by Theorem 2.9.14,  $A$  is a basis. Since it is finite, it is a Hamel basis and must therefore contain precisely  $n$  vectors.

Let  $A = \{x_\alpha: \alpha \in \Lambda\}$  and  $B = \{y_\beta: \beta \in \Gamma\}$  be complete orthonormal sets in  $H$ . For any  $x_\alpha \in A$ , the set

$$B_{x_\alpha} = \{y_\beta \in B: (x_\alpha, y_\beta) \neq 0\}$$

must be countable [see Remark 2.9.11]. Clearly,  $\bigcup_\alpha B_{x_\alpha} \subseteq B$ . We next show that  $B \subseteq \bigcup_\alpha B_{x_\alpha}$ . Let  $y_\beta \in B$ . Suppose  $y_\beta \notin B_{x_\alpha}$  for no  $\alpha$ . Then  $(x_\alpha, y_\beta) = 0$  for all  $\alpha \in \Lambda$ . In other words,  $y_\beta \perp A$ . Since  $A$  is complete, it follows that  $y_\beta = 0$ , which is impossible since  $\|y_\beta\| = 1$ . Hence,  $y_\beta \in B_{x_\alpha}$  for some  $x_\alpha$ . Thus

$$B = \bigcup_\alpha B_{x_\alpha}.$$

It follows that  $|B|$ , the cardinality of  $B$ , satisfies  $|B| \leq \aleph_0 |A| = |A|$ . Interchanging the roles of  $A$  and  $B$ , we also have  $|A| \leq |B|$ . This completes the proof.  $\square$

**Definition 2.9.19** Let  $H$  be a Hilbert space. If  $H \neq \{0\}$ , we define the **orthogonal dimension** of  $H$  to be the unique cardinal number of a complete orthonormal set in  $H$ . If  $H = \{0\}$ , we say that  $H$  has orthogonal dimension 0.

If  $H$  is finite dimensional, then the orthogonal dimension of  $H$  is the cardinal of a Hamel basis.

**Theorem 2.9.20** (Riesz–Fischer) *Let  $\{x_\alpha\}_{\alpha \in A}$  be a complete orthonormal system in a Hilbert space  $H$  and  $\ell^2(A) = L^2(A, \mathfrak{S}, \mu)$ , where  $\mathfrak{S}$  denotes the collection of all subsets of  $A$  and  $\mu$  is counting measure on  $A$ . Then  $H$  is isometrically isomorphic to  $\ell^2(A)$ .*

*Proof* For  $x \in H$ , let  $T(x)$  be that function on  $A$  such that

$$[T(x)](\alpha) = (x, x_\alpha), \quad \alpha \in A.$$

Then  $T$  maps  $H$  into  $\ell^2(A)$ , for  $\sum_{\alpha \in A} |(x, x_\alpha)|^2 < \infty$  by Bessel's Inequality. Also, for  $x, y \in H$ , we have

$$\begin{aligned} [T(x+y)](\alpha) &= (x+y, x_\alpha) = (x, x_\alpha) + (y, x_\alpha) \\ &= [T(x)](\alpha) + [T(y)](\alpha), \quad \alpha \in A, \end{aligned}$$

i.e.  $T(x+y) = T(x) + T(y)$ . It is equally easy to show that  $T(ax) = aT(x)$  for scalar  $a$ . Thus,  $T$  is linear. Using Theorem 2.9.16(e), we have

$$\begin{aligned}
(T(x), T(y)) &= \sum_{\alpha \in A} [T(x)](\alpha) \overline{[T(y)](\alpha)} \\
&= \sum_{\alpha \in A} (x, x_\alpha) \overline{(y, x_\alpha)} \\
&= (x, y)
\end{aligned}$$

and so  $T$  preserves inner products. It remains to show that the mapping  $T: H \rightarrow \ell^2(A)$  is onto.

Let  $f \in \ell^2(A)$ . Then  $\sum_{\alpha \in A} |f(\alpha)|^2 < \infty$ . Let  $\alpha_1, \alpha_2, \dots$  be those  $\alpha$ 's for which  $f(\alpha) \neq 0$ . The condition  $\sum_{\alpha \in A} |f(\alpha)|^2 < \infty$  becomes  $\sum_{i=1}^{\infty} |f(\alpha_i)|^2 < \infty$ . It follows from Theorem 2.9.9 that  $x = \sum_{i=1}^{\infty} f(\alpha_i) \alpha_i$  is in  $H$ . Since  $(x, x_{\alpha_j}) = f(\alpha_j)$ , so  $(x, x_{\alpha_j}) \neq 0$ . For a fixed  $p$  and any  $m \geq p$ , we have

$$\begin{aligned}
|(x, x_{\alpha_p}) - f(\alpha_p)| &= \left| (x, x_{\alpha_p}) - \sum_{i=1}^m f(\alpha_i) (x_{\alpha_i}, x_{\alpha_p}) \right| \\
&\leq \left\| x - \sum_{i=1}^m f(\alpha_i) x_{\alpha_i} \right\| \|x_{\alpha_p}\| \rightarrow 0
\end{aligned}$$

as  $m \rightarrow \infty$ . Therefore,  $f(\alpha_j) = (x, x_{\alpha_j}) = (T(x))(\alpha_j)$  for all  $j = 1, 2, \dots$ . The equality also holds for those  $\alpha$ 's for which  $f(\alpha) = 0$ . This completes the proof.  $\square$

### Remarks 2.9.21

- (i) The following form of the above Theorem was originally proved by Riesz and Fischer in 1907:

Let  $\{\alpha_n\}_{n \in \mathbb{Z}}$  be in  $\ell^2(\mathbb{Z})$ , that is,  $\sum_{n=-\infty}^{\infty} |\alpha_n|^2 < \infty$ . Then, there exists a function  $f$  in  $L^2([-\pi, \pi], \frac{dt}{2\pi})$ , such that  $\hat{f}(n) = \alpha_n, n \in \mathbb{Z}$ , where  $\hat{f}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt$  is the  $n$ th Fourier coefficient of  $f$  with respect to the orthonormal basis  $\{e^{-int}; n \in \mathbb{Z}\}$ .

- (ii) A Hilbert space is completely determined up to an isometric isomorphism by its orthogonal dimension, i.e. by the cardinality of its complete orthonormal basis. The space  $L^2([-\pi, \pi], \frac{dt}{2\pi})$  is isometrically isomorphic to  $\ell^2(\mathbb{Z})$  and hence also to  $\ell^2(\mathbb{N})$ .

### Problem Set 2.9

- 2.9.P1. Let  $\{e_n\}_{n \geq 1}$  and  $\{\tilde{e}_n\}_{n \geq 1}$  be orthonormal sequences in a Hilbert space  $H$  and let  $M_1 = \text{span}(e_n)$  and  $M_2 = \text{span}(\tilde{e}_n)$ . Show that  $\overline{M_1} = \overline{M_2}$  if, and only if,

$$e_n = \sum_{m=1}^{\infty} \alpha_{nm} \tilde{e}_m, \quad \tilde{e}_n = \sum_{m=1}^{\infty} \bar{\alpha}_{nm} e_m, \quad \alpha_{nm} = (e_n, \tilde{e}_m).$$

2.9.P2. Let  $H$  be a Hilbert space. Then show that the following hold:

- (a) If  $H$  is separable, every orthonormal set in  $H$  is countable.
- (b) If  $H$  contains an orthonormal sequence which is complete in  $H$ , then  $H$  is separable.

2.9.P3. Let  $A \subseteq [-\pi, \pi]$  and  $A$  be measurable. Prove that

$$\lim_{n \rightarrow \infty} \int_A \cos nt \, dt = \lim_{n \rightarrow \infty} \int_A \sin nt \, dt = 0.$$

2.9.P4. Let  $n_1 < n_2 < n_3 < \dots$  be positive integers and

$$E = \{x \in [-\pi, \pi] : \lim_k \sin n_k x \text{ exists}\}.$$

Prove that  $m(E) = 0$ , where  $m(E)$  denotes the Lebesgue measure of  $E$ .

2.9.P5. Let  $e_j(z) = z^j$ ,  $j \in \mathbb{Z}$ . Show that  $\{e_j\}_{j=-\infty}^{\infty}$  is an orthonormal sequence in  $RL^2$  (notation as in Example 2.1.3(vi)).

2.9.P6. Let  $\alpha_1, \alpha_2 \in D(0,1) = \{z: |z| < 1\}$  and  $\alpha_1 \neq \alpha_2$ . Show that the vectors

$$e_1(z) = \frac{(1 - |\alpha_1|^2)^{\frac{1}{2}}}{1 - \bar{\alpha}_1 z} \quad \text{and} \quad e_2(z) = \frac{z - \alpha_1}{1 - \bar{\alpha}_1 z} \frac{(1 - |\alpha_2|^2)^{\frac{1}{2}}}{1 - \bar{\alpha}_2 z}$$

constitute an orthonormal system in  $RH^2$  (notation as in Example 2.1.3(vi)).

2.9.P7. Let  $\{e_n\}_{n \geq 1}$  be an orthonormal basis in  $H$ . Show that for any orthonormal set  $\{f_n\}_{n \geq 1}$ , if

$$\sum_{n=1}^{\infty} \|e_n - f_n\|^2 < \infty,$$

then  $\{f_n\}_{n \geq 1}$  is an orthonormal basis.

2.9.P8. A real-valued function on an interval having a continuous nonvanishing derivative on the interior of its domain maps a set of (Lebesgue) measure zero into a set of measure zero. In case the domain is an open interval, in which case the range is also an open interval and an inverse exists, the inverse also has the same properties. [Examples of such a function on the domain  $(0, \infty)$  would be  $\exp(-x)$  and  $x^2$ .]

## 2.10 Orthogonal Decomposition and Riesz Representation

A result of particular interest about Hilbert space is the projection theorem, namely, if  $M$  is any closed subspace of a Hilbert space  $H$ , then  $H$  can be decomposed into the direct sum of  $M$  and its orthogonal complement (to be defined below). This important geometric property is one of the main reasons that Hilbert spaces are easier to handle than Banach spaces.

A characterisation of a bounded linear functional [see Definition 2.10.18 below] on a Hilbert space, known as the Riesz Representation Theorem, will be studied.

P.L. Chebyshev sought the approximation of arbitrary functions by linear combinations of given ones. He considered approximations in the spaces of continuous functions,  $L^p$  spaces, etc. These have a bearing on constrained optimisation.

We deal with the approximation problem in a pre-Hilbert space  $X$ ; given a set of  $n$  linearly independent vectors  $\{v_1, v_2, \dots, v_n\}$  and an  $x \in X$ , to find a method of computing the minimum value of

$$\left\| x - \sum_{j=1}^n c_j v_j \right\|,$$

where  $c_1, c_2, \dots, c_n$  range over all scalars and to find the corresponding values of  $c_1, c_2, \dots, c_n$ . The reader will learn that this is precisely the problem of finding the distance of  $x$  from the linear span of  $\{v_1, v_2, \dots, v_n\}$ .

Recall that a set of  $M$  nonzero vectors in a pre-Hilbert space is said to be orthogonal if  $x \perp y$  whenever  $x$  and  $y$  are distinct vectors of  $M$ .

**Definition 2.10.1** Let  $X$  be a pre-Hilbert space and  $x \in X$ . We define

$$x^\perp = \{y \in X : (x, y) = 0\}.$$

and if  $S$  is a subset of  $X$ ,

$$S^\perp = \{y \in X : (x, y) = 0 \text{ for all } x \in S\}.$$

The symbol  $x \perp$  [respectively,  $S^\perp$ ] is read as  $x$  perp [respectively,  $S$  perp]. One writes  $S^{\perp\perp}$  for the perp of  $S^\perp$ ; thus  $S^{\perp\perp} = (S^\perp)^\perp$ . The set  $S^\perp$  is called the **orthogonal complement of  $S$** .

*Remarks 2.10.2*

- (i) Observe that  $x^\perp$  is a subspace of  $X$ , since  $(x, y) = 0$  and  $(x, z) = 0$  imply  $(x, \alpha y + \beta z) = 0$ , where  $\alpha, \beta$  are scalars. Also,  $x^\perp$  is precisely the set of vectors where the continuous function  $y \rightarrow (x, y)$  is zero. Hence,  $x^\perp$  is a closed subspace of  $X$ . Since

$$S^\perp = \bigcap_{x \in S} x^\perp,$$

it follows that  $S^\perp$ , being the intersection of closed subspaces of  $X$ , is itself a closed subspace of  $X$ .

$$(ii) \quad S^\perp = (\overline{S})^\perp.$$

Let  $y \in S^\perp$ . Then  $(x, y) = 0$  for all  $x \in S$ . Let  $z \in \overline{S}$ . Then there exists a sequence  $\{z_n\}_{n \geq 1}$  in  $S$  such that  $z_n \rightarrow z$ . The continuity of the mapping  $x \rightarrow (x, y)$  and the fact that  $(z_n, y) = 0$  for  $n = 1, 2, \dots$  imply  $(z, y) = 0$ . Since  $z \in \overline{S}$  is arbitrary, we conclude that  $y \in (\overline{S})^\perp$ .

On the other hand, if  $y \in (\overline{S})^\perp$ , then  $(y, x) = 0$  for all  $x \in \overline{S}$ . Since  $S \subseteq \overline{S}$ , it follows that  $(y, x) = 0$  for all  $x \in S$ , that is  $y \in S^\perp$ .

**Proposition 2.10.3** *Let  $S$  and  $S_1$  be subsets of an inner product space  $X$ . Then the following hold.*

- (a)  $S^\perp$  is a closed subspace of  $X$  and  $S \cap S^\perp \subseteq \{0\}$ ;
- (b)  $S \subseteq S^{\perp\perp}$ ;
- (c)  $S \subseteq S_1$  implies  $S^\perp \supseteq S_1^\perp$ ;
- (d)  $S^\perp = S^{\perp\perp\perp}$ .

*Proof*

- (a) In Remark 2.10.2(i), we have noted that  $S^\perp$  is a closed subspace of  $X$ . If  $x \in S \cap S^\perp$  then  $x \perp x$ , that is,  $(x, x) = 0$ , which implies  $x = 0$ .
- (b) Let  $x \in S$ . For any  $y \in S^\perp$ , one has  $(y, x) = 0$ , so that  $x \perp S^\perp$  and therefore  $x \in S^{\perp\perp}$ .
- (c) If  $x \in S_1^\perp$ , then  $(x, y) = 0$  for all  $y \in S_1$ . In particular,  $(x, y) = 0$  for all  $y \in S$ , which implies  $x \in S^\perp$ .
- (d) Applying (iii) to the relation  $S \subseteq S^{\perp\perp}$ , we have  $(S^{\perp\perp})^\perp \subseteq S^\perp$ . Also,  $S^\perp \subseteq (S^\perp)^{\perp\perp}$  by (b) above. Since  $(S^{\perp\perp})^\perp = (S^\perp)^{\perp\perp}$ , as in each case one starts with  $S$  and perps three times, it follows that  $(S^{\perp\perp})^\perp = S^\perp$ .  $\square$

*Example* Let  $S = \{f \in L^2[0, 1] : f(t) = 0 \text{ a.e. for } 0 \leq t \leq \frac{1}{2}\}$ .

Then

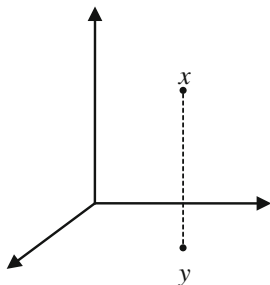
$$S^\perp = \left\{ g \in L^2[0, 1] : g(t) = 0 \text{ a.e. on } \left[ \frac{1}{2}, 1 \right] \right\}$$

and

$$S^{\perp\perp} = \left\{ g \in L^2[0, 1] : g(t) = 0 \text{ a.e. on } \left[ 0, \frac{1}{2} \right] \right\} = S.$$

Hint: To compute  $S^\perp$ , first show that  $\int_x^1 g(t)dt = 0$  for every  $x \in [\frac{1}{2}, 1]$  and then use regularity of Lebesgue measure.

If  $x$  is a point lying outside a plane in  $\mathbb{R}^3$ , then there is a unique  $y$  in the plane which is closer to  $x$  than any other point of the plane. This assertion, when translated in the language of Hilbert spaces, yields rich dividends via the Riesz Representation Theorem below. The accompanying figure illustrates the situation when the plane is a coordinate plane. However, this need not always be the case.



**Definition 2.10.4** A subset  $K$  of a vector space is **convex** if, for all  $x, y \in K$ , and all  $\lambda$  such that  $0 < \lambda < 1$ , the vector  $\lambda x + (1 - \lambda)y$  belongs to  $K$ . The set of vectors  $\{\lambda x + (1 - \lambda)y : 0 < \lambda < 1\}$  is the **line segment** joining  $x$  and  $y$ . The **convex hull** of a subset  $S$  of any vector space is the intersection of all convex subsets containing  $S$  and is denoted by  $\text{co}(S)$  or by  $\text{co}S$ .

It is sometimes neater to work with an equivalent formulation of convexity as follows: for all  $x, y \in K$ , and all  $\alpha, \beta \geq 0$  such that  $\alpha + \beta = 1$ , the vector  $\alpha x + \beta y$  belongs to  $K$ .

It is easy to see that the intersection of any family of convex sets is again convex; in particular, any convex hull is a convex set. By using the alternative formulation of convexity, the convex hull of any finite set of vectors  $\{x_1, x_2, \dots, x_n\}$  is easily seen to consist of precisely those vectors which can be written as  $\sum_{k=1}^n \lambda_k x_k$ , where  $0 \leq \lambda_k \leq 1$  for each  $k$  and  $\sum_{k=1}^n \lambda_k = 1$ . (Induction: we start with the vectors  $x_1, x_2, \dots, x_n$  and nonnegative  $\lambda_1, \lambda_2, \dots, \lambda_n$  satisfying  $\sum_{k=1}^n \lambda_k = 1$ . If  $\sum_{k=1}^{n-1} \lambda_k = 0$ , then  $\lambda_k$  is 0 for  $k = 1, \dots, n-1$  and  $\lambda_n = 1$ , which together imply  $\sum_{k=1}^n \lambda_k x_k$  is in the convex hull. Assume  $\sum_{k=1}^{n-1} \lambda_k = \beta > 0$ . Then  $\sum_{k=1}^{n-1} (\lambda_k/\beta)x_k$  is in the convex hull by the induction hypothesis. Consequently,  $\sum_{k=1}^n \lambda_k x_k = \beta \sum_{k=1}^{n-1} (\lambda_k/\beta)x_k + \lambda_n x_n$  is in the convex hull. Conversely, the vectors that can be written in this form obviously constitute a convex set that contains  $\{x_1, x_2, \dots, x_n\}$  and therefore contains the convex hull under reference.)

This description of the convex hull will now be used for arguing that it is compact when the vector space is normed.

When  $n = 1$ , there is nothing to prove. Assume as induction hypothesis that the convex hull of any  $n$  vectors is compact. Consider any set  $\{x_1, x_2, \dots, x_n, x\}$  of  $n + 1$

vectors. Let  $\{y_p\}_{p \geq 1}$  be a sequence in  $\text{co}\{x_1, x_2, \dots, x_n, x\}$ . Each  $y_p$  can be written as  $\sum_{k=1}^n \lambda_{k,p} x_k + \lambda_p x$ , where  $0 \leq \lambda_{k,p}, \lambda_p \leq 1$  for each  $k$  and  $\sum_{k=1}^n \lambda_{k,p} + \lambda_p = 1$ . If there are only finitely many such  $p$ , then we can assume that  $1 - \lambda_p > 0$  for every  $p$ . It then follows that  $\sum_{k=1}^n \lambda_{k,p} / (1 - \lambda_p) = 1$ , each term in the sum being non-negative. This means  $z_p = \sum_{k=1}^n \lambda_{k,p} x_k / (1 - \lambda_p)$  is in the convex hull of  $\{x_1, x_2, \dots, x_n\}$ . By the induction hypothesis,  $\{z_p\}$  has a subsequence  $\{z_{p(q)}\}$  converging to a limit  $z \in \text{co}\{x_1, x_2, \dots, x_n\}$ . Now the bounded sequence  $\{\lambda_{p(q)}\}$  in  $\mathbb{R}$  has a convergent subsequence  $\{\lambda_{p(q(r))}\}$ , whose limit we shall denote by  $\lambda$ . Then  $\{z_{p(q(r))}\}$  converges to  $z$  and therefore  $\sum_{k=1}^n \lambda_{k,p(q(r))} x_k = (1 - \lambda_{p(q(r))}) z_{p(q(r))}$  forms a sequence converging to  $(1 - \lambda)z$ . As  $y_{p(q(r))} = \sum_{k=1}^n \lambda_{k,p(q(r))} x_k + \lambda_{p(q(r))} x$ , the subsequence  $\{y_{p(q(r))}\}$  converges to  $(1 - \lambda)z + \lambda x$ , which belongs to  $\text{co}(\text{co}\{x_1, x_2, \dots, x_n\} \cup \{x\})$ . The latter can easily be seen to be the same as  $\text{co}\{x_1, x_2, \dots, x_n, x\}$ . This completes the induction proof that the convex hull of any finite set of vectors is compact.

If  $K$  is convex and  $x$  is any vector, then the convex hull  $\text{co}(K \cup \{x\})$  is precisely  $K_1 = \{\alpha x + \beta k : \alpha, \beta \geq 0, \alpha + \beta = 1\}$ . The convexity of  $K_1$  follows from the three computations

- (a)  $\alpha(\alpha_1 x + \beta_1 k_1) + \beta(\alpha_2 x + \beta_2 k_2) = (\alpha\alpha_1 + \beta\alpha_2)x + (\alpha\beta_1 k_1 + \beta\beta_2 k_2)$   
 $= (\alpha\alpha_1 + \beta\alpha_2)x + \gamma((\alpha\beta_1/\gamma)k_1 + (\beta\beta_2/\gamma)k_2)$ , where  $\gamma = 1 - (\alpha\alpha_1 + \beta\alpha_2)$  is nonzero;
- (b)  $(\alpha\beta_1/\gamma) + (\beta\beta_2/\gamma) = 1$  because

$$\alpha\beta_1 + \beta\beta_2 + (\alpha\alpha_1 + \beta\alpha_2) = \alpha(\alpha_1 + \beta_1) + \beta(\alpha_2 + \beta_2) = 1,$$

so that  $\alpha\beta_1 + \beta\beta_2 = 1 - (\alpha\alpha_1 + \beta\alpha_2) = \gamma$ ,

- (c)  $\alpha\alpha_1 + \beta\alpha_2 \leq \alpha + \beta = 1$  when  $0 \leq \alpha_1 \leq 1$  and  $0 \leq \alpha_2 \leq 1$ .

Regarding (a), we note that  $\gamma = 0$  implies  $\alpha_1 = \alpha_2 = 1$  and  $\beta_1 = \beta_2 = 0$ , in which case  $\alpha(\alpha_1 x + \beta_1 k_1) + \beta(\alpha_2 x + \beta_2 k_2) = x$ . Once the convexity of  $K_1$  is established, it is a trivial matter to see that it is the convex hull of  $K$ .

**Theorem 2.10.5** (Closest point property) *Let  $K$  be a nonempty closed convex set in a Hilbert space  $H$ . For every  $x \in H$ , there is a unique point  $y \in K$  which is closer to  $x$  than any other point of  $K$ , i.e.*

$$\|x - y\| = \inf_{z \in K} \|x - z\|.$$

*Proof* Let  $\delta = \inf_{z \in K} \|x - z\|$ . Since  $K \neq \emptyset$ ,  $\delta < \infty$ , therefore for each  $n \in \mathbb{N}$ , there exists  $y_n \in K$  such that

$$\|x - y_n\|^2 < \delta^2 + \frac{1}{n}. \quad (2.73)$$

We shall prove that  $\{y_n\}_{n \geq 1}$  is a Cauchy sequence in  $K$ . Consider the vectors  $x - y_n$  and  $x - y_m$ . By the Parallelogram Law [Proposition 2.2.3(c)],

$$\begin{aligned} \|(x - y_n) - (x - y_m)\|^2 + \|(x - y_n) + (x - y_m)\|^2 &= 2(\|x - y_n\|^2 + \|x - y_m\|^2) \\ &\leq 4\delta^2 + 2\left(\frac{1}{n} + \frac{1}{m}\right). \end{aligned}$$

Rearranging the left side, we obtain

$$\|y_n - y_m\|^2 + 4\left\|x - \frac{y_n + y_m}{2}\right\|^2 \leq 4\delta^2 + 2\left(\frac{1}{n} + \frac{1}{m}\right)$$

and hence

$$\|y_n - y_m\|^2 \leq 4\delta^2 + 2\left(\frac{1}{n} + \frac{1}{m}\right) - 4\left\|x - \frac{y_n + y_m}{2}\right\|^2.$$

Since  $K$  is convex,  $y_n, y_m \in K$ , we have  $\frac{y_n + y_m}{2} \in K$ , and hence

$$\left\|x - \frac{y_n + y_m}{2}\right\|^2 \geq \delta^2.$$

Consequently,

$$\begin{aligned} \|y_n - y_m\|^2 &\leq 4\delta^2 + 2\left(\frac{1}{n} + \frac{1}{m}\right) - 4\delta^2 \\ &= 2\left(\frac{1}{n} + \frac{1}{m}\right). \end{aligned}$$

Thus,  $\{y_n\}_{n \geq 1}$  is a Cauchy sequence and so converges to a limit  $y \in H$ . Since  $K$  is closed,  $y \in K$  and therefore

$$\|x - y\| \geq \delta.$$

On letting  $n \rightarrow \infty$  in (2.73), we obtain

$$\|x - y\| \leq \delta$$

and so

$$\|x - y\| = \delta.$$

We have proved that there is a closest point to  $x$  in  $K$ . It remains to show that it is unique. Suppose that  $z \in K$  ( $z \neq y$ ) is such that  $\|x - z\| = \delta$ . Then  $\frac{y+z}{2} \in K$  so that

$$\left\|x - \frac{y+z}{2}\right\| \geq \delta.$$

On applying the Parallelogram Law [Proposition 2.2.3(c)] to  $y - x$  and  $z - x$ , we get

$$\begin{aligned} \|y - z\|^2 &= 2\|y - x\|^2 + 2\|z - x\|^2 - \|y + z - 2x\|^2 \\ &= 4\delta^2 - 4\left\|\frac{y+z}{2} - x\right\|^2 \leq 0. \end{aligned}$$

Hence  $y = z$ . □

*Remarks 2.10.6*

- (i) If  $x \in H$  is such that  $x \in K$ , then the vector nearest to  $x$  is  $x$  itself.
- (ii) If  $K$  is not closed, the conclusion of Theorem 2.10.5 may not hold. In fact, in this situation, whether  $K$  is convex or not, there always exists a point in  $H$  having no closest approximation in  $K$ . Any point in the closure that does not belong to  $K$  will serve the purpose.

Let  $H = \ell^2$  and  $K = \{x = \{\lambda_k\}_{k \geq 1} : \lambda_k \neq 0 \text{ for only finitely many } k \text{'s and } \sum_{k=1}^{\infty} \lambda_k = 1\}$ .  $K$  is convex. However,  $K$  is not closed. In fact, the sequence  $y_1 = (1, 0, 0, \dots)$ ,  $y_2 = (\frac{1}{2}, \frac{1}{2}, 0, 0, \dots)$ ,  $\dots$ ,  $y_n = (\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}, 0, 0, \dots)$ ,  $\dots$  is in  $K$ . However, the limit of the sequence  $\{y_n\}_{n \geq 1}$  in  $\ell^2$  is the point  $y = (0, 0, \dots)$  of  $\ell^2$ , which does not belong to  $K$ . According to the preceding paragraph, the point  $y = (0, 0, \dots)$  does not possess a closest point in  $K$ .

- (iii) The conclusion of Theorem 2.10.5 fails to hold if  $H$  is not a Hilbert space. Let  $X = \mathbb{R}^2$ , the real Banach space with  $\|(x_1, x_2)\| = \max\{|x_1|, |x_2|\}$ . Consider the closed convex set  $K = \{(x_1, x_2) : x_1 \geq 1\}$ . The minimal distance of the origin from  $K$  is attained at each of the points of the line segment  $\{(x_1, x_2) : x_1 = 1 \text{ and } |x_2| \leq 1\}$ . Even when the norm comes from an inner product and  $H$  is not complete, the existence part of the conclusion of Theorem 2.10.5 may fail to hold; an example of this will be given later in (ii) of Remarks 2.10.12. The uniqueness part however holds because its proof does not use completeness.

**Corollary 2.10.7** *Every nonempty closed convex set  $K$  in a Hilbert space  $H$  contains a unique element of smallest norm.*

*Proof* Take  $x = 0$  in Theorem 2.10.5. □

**Example 2.10.8** Let  $K = \{y = (\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{C}^n : \sum_{k=1}^n \lambda_k = 1\}$ .  $K$  is a closed convex subset of  $\mathbb{C}^n$ . The unique vector  $y_0 \in K$  of smallest norm is  $y_0 = (\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})$ ; indeed, if  $y_0$  has two unequal components, then interchanging them leads to another vector in  $K$  with smallest norm; consequently, all components of  $y_0$  are equal. The reader who is familiar with constrained optimisation can verify the claim made above independently of the corollary.

**Corollary 2.10.9** Let  $M$  be a closed subspace of a Hilbert space  $H$ . If  $x$  is a vector in  $H$ , and if  $\delta = \inf\{\|x - z\| : z \in M\}$  then there exists a unique  $y \in M$  such that  $\delta = \|x - y\|$ .

*Proof* Every subspace of a vector space is convex. □

**Theorem 2.10.10** Let  $M$  be a closed subspace of a Hilbert space  $H$  and  $x \in H$ . If  $y$  denotes the unique element in  $M$  for which  $\|x - y\| = \inf\{\|x - z\| : z \in M\}$ , then  $x - y$  is orthogonal to  $M$ . Conversely, if  $y \in M$  is such that  $x - y$  is orthogonal to  $M$ , then  $\|x - y\| = \inf\{\|x - z\| : z \in M\}$ .

*Proof* Consider  $z \in M$  with  $\|z\| = 1$ . Then  $w = y + (x - y, z)z$  lies in  $M$  and we have

$$\begin{aligned} \|x - y\|^2 &\leq \|x - w\|^2 = (x - w, x - w) \\ &= (x - y - (x - y, z)z, x - y - (x - y, z)z) \\ &= \|x - y\|^2 - |(x - y, z)|^2. \end{aligned}$$

This shows that  $(x - y, z) = 0$ , i.e.  $x - y \perp z$ . Since every vector in  $M$  is a scalar multiple of a vector in  $M$  of norm 1, it follows that  $x - y \perp M$ .

If  $z \in M$ , then  $x - y$  is orthogonal to  $y - z$ , so that  $\|x - z\|^2 = \|x - y + y - z\|^2 = \|x - y\|^2 + \|y - z\|^2 \geq \|x - y\|^2$ . Thus,  $\|x - y\| = \inf\{\|x - z\| : z \in M\}$ . □

It may be noted that  $x - y$  is closest to  $x$  in  $M^\perp$ .

**Theorem 2.10.11** (Orthogonal Decomposition Theorem) If  $M$  is a closed subspace of a Hilbert space  $H$ , then  $H = M \oplus M^\perp$  and  $M = M^{\perp\perp}$ .

*Proof* Let  $x \in H$ . Since  $M$  is a closed subspace of  $H$ , there exists a unique vector  $y \in M$  such that  $\|x - y\| = \inf\{\|x - z\| : z \in M\}$ . So  $x - y \perp M$  by Theorem 2.10.10. Hence,  $x = y + (x - y)$ , where  $y \in M$  and  $x - y \in M^\perp$ . Since  $M \cap M^\perp = \{0\}$ , it follows that  $H = M \oplus M^\perp$ .

We already know that  $M \subseteq M^{\perp\perp}$  [Proposition 2.10.3(b)]. On the other hand, let  $x \in M^{\perp\perp}$ . If  $x = y + z$ , where  $y \in M$  and  $z \in M^\perp$ , then  $x$  and  $y$  are in  $M^{\perp\perp}$ . Hence,  $z = x - y \in M^{\perp\perp} (M^{\perp\perp}$  is a subspace of  $H)$ . Since  $z \in M^\perp$ , it follows that  $(z, z) = 0$ , that is,  $z = 0$  or  $x = y$ . This shows that  $M^{\perp\perp} \subseteq M$ . The proof is complete. □

**Remarks 2.10.12**

- (i) If  $M$  is a closed subspace of a Hilbert space  $H$  and  $x \in H$ , then  $x$  can be uniquely expressed as

$$x = y + z,$$

where  $y \in M$  and  $z \in M^\perp$ .

- (ii) The condition that  $H$  is a Hilbert space for a closed subspace to satisfy  $M = M^{\perp\perp}$  cannot be omitted. Let  $\ell_0 \subseteq \ell^2$  be the inner product space consisting of sequences, each of which has only finitely many nonzero terms. Let  $M = \{x = \{\lambda_k\}_{k \geq 1} : \lambda_k \neq 0 \text{ for only finitely many } k\text{'s and } \sum_{k=1}^\infty \frac{1}{k} \lambda_k = 0\}$ . Clearly,  $M$  is a subspace of  $\ell_0$ . Moreover,  $M$  is closed, as is proved below.

Let  $\{x^{(n)}\}_{n \geq 1}$  be a sequence in  $M$  such that  $x^{(n)} \rightarrow x$  in  $\ell_0$ . By the Cauchy-Schwarz Inequality, it follows that

$$\left| \sum_{k=1}^\infty \frac{1}{k} \lambda_k \right|^2 = \left| \sum_{k=1}^\infty \frac{1}{k} (\lambda_k - \lambda_k^{(n)}) \right|^2 \leq \left( \sum_{k=1}^\infty \frac{1}{k^2} \right) \left( \sum_{k=1}^\infty |\lambda_k - \lambda_k^{(n)}|^2 \right),$$

where  $\lambda_k^{(n)}$  is the  $k$ th component of  $x_n$ . Consequently,  $\sum_{k=1}^\infty \frac{1}{k} \lambda_k = 0$ . So,  $x \in M$ .

We next show that  $M^\perp = \{0\}$ . Assume  $0 \neq z \in \ell_0$  and  $z \perp M$ . Then there exists  $k$  such that  $z = (x_1, \dots, x_k, 0, \dots)$  and  $\sum_{i=1}^k |x_i|^2 \neq 0$ . Let  $\mu = -(k+1) \sum_{j=1}^k \frac{1}{j} x_j$ . Then  $w = (x_1, \dots, x_k, \mu, 0, \dots) \in M$  in view of the definition of  $\mu$ . Hence,  $z \perp w$ , i.e.  $(z, w) = 0$ . But  $(z, w) = \|z\|^2$ . It follows that  $z = 0$ , contradicting the assumption on  $z$ . Consequently,  $M^{\perp\perp} = \ell_0 \neq M$ .

We shall use the fact that  $M^\perp = \{0\}$  to show that the closed convex subset  $M$  of the (incomplete) inner product space  $\ell_0$  has the property that every  $x \in \ell_0$  that does not lie in  $M$  fails to have a nearest element in  $M$ . In particular, it will follow that the conclusion of Theorem 2.10.5 may fail to hold in the absence of completeness. Suppose  $x \in \ell_0$  does not lie in  $M$  but has a closest element  $y \in M$ . It follows that  $x - y \perp M$  exactly as in the first paragraph of the proof of Theorem 2.10.10, considering that completeness is not needed in that paragraph. Since we have shown that  $M^\perp = \{0\}$ , we infer that  $x - y = 0$ , which is a contradiction because  $y \in M$ , whereas  $x \notin M$ .

- (iii) If  $M$  is any linear subspace of  $H$ , then  $\overline{M} = M^{\perp\perp}$ . Observe that  $M \subseteq M^{\perp\perp}$  [see Proposition 2.10.3(b)]. It follows that  $\overline{M} \subseteq M^{\perp\perp}$ , since  $M^{\perp\perp}$  is a closed subspace of  $H$ . As  $M \subseteq \overline{M}$ , it follows that  $(\overline{M})^\perp \subseteq M^\perp$  using Proposition 2.10.3(c). Another application of Proposition 2.10.3(c) yields  $M^{\perp\perp} \subseteq (\overline{M})^{\perp\perp} = \overline{M}$  since  $\overline{M}$  is a closed subspace of  $H$  [Theorem 2.10.11].
- (iv) If  $H = M \oplus N$ ,  $M \subseteq N^\perp$ , then  $M = N^\perp$  and is therefore closed, as we now show. Suppose  $x \in N^\perp$  and  $x \notin M$ . The vector  $x$  has the representation  $x = y + z$ , where  $y \in M$  and  $z \in N$ . Now

$$0 = (x, z) = (y, z) + (z, z)$$

which implies  $z = 0$ . Consequently,  $x = y$ , which is a contradiction because  $x \notin M$  and  $y \in M$ .

- (v)  $S^{\perp\perp} = (S^\perp)^\perp$  is the smallest closed subspace of the Hilbert space  $H$  which contains  $S$ .
- (vi) Let  $\{M_k\}_{k \geq 1}$  be a sequence of closed linear subspaces of a Hilbert space  $H$ . There exists a smallest closed linear subspace  $M$  such that  $M_k \subseteq M$  for all  $k$  and it has the property that  $x \perp M$  if, and only if,  $x \perp M_k$  for all  $k$ . To see why, let  $S = \{x \in H: x \in M_k \text{ for some } k\}$ . Clearly  $M_k \subseteq S$  for all  $k$ . Moreover,  $S$  is the smallest subset of  $H$  with this property. Set  $M = S^{\perp\perp}$ . If  $\mathfrak{N}$  is a closed linear subspace such that  $M_k \subseteq \mathfrak{N}$  for all  $k$ , then  $S \subseteq \mathfrak{N}$ . Hence,  $M \subseteq \mathfrak{N}$  in view of (v). The assertion that  $x \perp M$  if, and only if,  $x \perp M_k$  for all  $k$  is proved by using the following observation:

$$M^\perp = S^{\perp\perp\perp} = S^\perp.$$

*Example 2.10.13* Consider  $\mathcal{F}_o = \{f \in L^2[-1, 1]: f(t) = -f(-t)\}$  and  $\mathcal{F}_e = \{f \in L^2[-1, 1]: f(t) = f(-t)\}$ .

The set  $\mathcal{F}_o$  is an infinite-dimensional linear subspace of  $L^2[-1, 1]$ . [ $f(t) = t^{2n-1}$ ,  $n = 1, 2, \dots$  are in  $\mathcal{F}_o$ ; they are countably many and linearly independent.] Also,  $\mathcal{F}_e$  is an infinite-dimensional subspace of  $L^2[-1, 1]$ . [ $\mathcal{F}_e$  contains the functions  $f(t) = t^{2n}$ ,  $n = 0, 1, 2, \dots$ ]

For  $f \in \mathcal{F}_o$  and  $g \in \mathcal{F}_e$ , the inner product

$$(f, g) = \int_{-1}^1 f(t) \overline{g(t)} dt = 0$$

since the function  $f(t) \overline{g(t)}$  is odd. Hence  $\mathcal{F}_o \perp \mathcal{F}_e$ , i.e.  $\mathcal{F}_o \subseteq \mathcal{F}_e^\perp$ .

For any function  $f \in L^2[-1, 1]$ ,

$$f_e(t) = \frac{f(t) + f(-t)}{2}, \quad f_o(t) = \frac{f(t) - f(-t)}{2} \quad \text{and} \quad f = f_e + f_o.$$

Moreover, this representation is unique, so that  $L^2[-1, 1] = \mathcal{F}_o \oplus \mathcal{F}_e$ . Since  $\mathcal{F}_o \subseteq \mathcal{F}_e^\perp$ , it follows that  $\mathcal{F}_o = \mathcal{F}_e^\perp$ . [See Remark 2.10.12(iv)].

The following proposition provides an alternate way of computing

$$\delta = \inf \{\|x - z\| : z \in M\},$$

where  $x \in H$  and  $M$  is a closed subspace of  $H$ .

**Proposition 2.10.14** *If  $M$  is a closed subspace of a Hilbert space  $H$  and  $x \in H$ , then*

$$\delta = \inf\{\|x - u\| : u \in M\} = \max\{|(x, z)| : z \in M^\perp \text{ and } \|z\| = 1\}.$$

*Proof* Suppose  $x \in H$ . Then,  $x$  has a unique representation of the form

$$x = y + z, \quad y \in M \quad \text{and} \quad z \in M^\perp.$$

Let  $w \in M^\perp$  and  $\|w\| = 1$ . Then

$$|(x, w)| = |(y + z, w)| = |(z, w)| \leq \|z\| = \|x - y\| = \delta \text{ [Theorem 2.10.10].}$$

Hence,  $\sup\{|(x, w)| : w \in M^\perp, \|w\| = 1\} \leq \delta$ . The vector  $w = \frac{x-y}{\delta}$  is in  $M^\perp$  [Theorem 2.10.10] and  $\|w\| = \frac{\|x-y\|}{\delta} = 1$ . For this  $w$ ,

$$(x, w) = \left(x, \frac{x-y}{\delta}\right) = \frac{1}{\delta}(x, x-y) = \frac{1}{\delta}(x-y, x-y) = \frac{1}{\delta}\|x-y\|^2 = \delta,$$

using the fact that  $y \in M$  and  $x - y \in M^\perp$ .

This completes the proof.  $\square$

Let  $H$  be a Hilbert space and  $M$  is a closed subspace of  $H$ . The Orthogonal Decomposition Theorem 2.10.11 says that  $H = M \oplus M^\perp$ . Thus for each  $x \in H$ , there are unique  $y \in M$  and  $z \in M^\perp$  such that  $x = y + z$ . Note that  $z = x - y$  is the unique vector in  $M^\perp$  closest to  $x$ . The vector  $y$  is the **projection** of  $x$  on  $M$  and it follows that  $z$  is the projection of  $x$  on  $M^\perp$ . This sets up mappings from  $H$  onto  $M$  and from  $H$  onto  $M^\perp$ , respectively.

**Theorem 2.10.15** *The mapping  $P_M: H \rightarrow M$  defined by  $P_M(x) = y$ ,  $x \in H$ , where  $y$  denotes the projection of  $x$  on  $M$ . The mapping  $P_M$  has the following properties:*

- (a)  $P_M$  is linear, i.e.  $P_M(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 P_M(x_1) + \alpha_2 P_M(x_2)$ , where  $\alpha_1$  and  $\alpha_2$  are scalars;
- (b) If  $x \in M$ , then  $P_M(x) = x$ . Thus,  $P_M$  is idempotent, i.e.  $P_M^2 = P_M$ ;
- (c) If  $x \in M^\perp$ , then  $P_M(x) = 0$ ;
- (d)  $(P_M(x), x) = \|P_M(x)\|^2 \leq \|x\|^2$  for all  $x \in H$ .

*Proof*

- (a) Let  $x_i = y_i + z_i$ ,  $i = 1, 2$  be the decomposition of  $x_i$  relative to  $M$ . Then

$$P_M(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 y_1 + \alpha_2 y_2 = \alpha_1 P_M(x_1) + \alpha_2 P_M(x_2).$$

- (b) Since  $x = x + 0$  is the unique decomposition of  $x \in M$ , it follows that  $P_M(x) = x$ . If  $x \in H$ , then  $P_M(x) \in M$  and, by what has just been proved,  $P_M(x) = P_M(P_M(x)) = P_M^2(x)$ . Thus  $P_M^2 = P_M$ .
- (c) If  $x \in M^\perp$ , then  $x = 0 + x$  is the unique decomposition of  $x$ . Thus  $P_M(x) = 0$ .

- (d)  $(P_M(x), x) = (y, y + z)$ , where  $x = y + z$  is the unique decomposition of  $x$ . Now  $y \perp z$  and therefore  $(y, y + z) = (y, y) = (P_M(x), P_M(x)) = \|P_M(x)\|^2$ . Also,  $\|P_M(x)\|^2 = (y, y) \leq \|y\|^2 + \|z\|^2 = \|x\|^2$ .  $\square$

**Definition 2.10.16** The map  $P_M$  is called the **orthogonal projection on  $M$** .

*Remarks 2.10.17*

- (i) The map  $P_M$  is often denoted by  $P$  when it is clear from the context on which subspace  $M$  the projection  $P_M$  is intended.
- (ii) In Theorem 2.10.15(d), we have checked that

$$\|P_M(x)\| \leq \|x\| \quad \text{for all } x \in H.$$

Since  $P_M(x) = x$  for all  $x \in M$ , it follows that  $\|P_M(x)\| = \|x\|$  and hence

$$\sup_{\|x\|=1} \|P_M(x)\| = 1 \text{ provided } M \neq \{0\}.$$

We now turn to the study of ‘linear functionals’ on Hilbert spaces.

**Definition 2.10.18** A **linear functional** on a vector space  $X$  over a field  $\mathbb{F}$  is a mapping  $f: X \rightarrow \mathbb{F}$  which satisfies  $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$  for all  $x, y \in X$  and all scalars  $\lambda, \mu$  in  $\mathbb{F}$ .

**Definition 2.10.19** Let  $X$  be a normed linear space over  $\mathbb{F}$ . A linear mapping  $f: X \rightarrow \mathbb{F}$  is called a **bounded linear functional** on  $X$  if it maps bounded subsets of  $X$  into bounded subsets of  $\mathbb{F}$ , or equivalently, if there exists a constant  $K$  such that

$$|f(x)| \leq K\|x\|, \quad x \in X.$$

The equivalence is a consequence of the linearity of the functional.

The linear functional  $f$  is said to be a **continuous** linear functional if for a sequence  $\{x_n\}_{n \geq 1}$  in  $X$ ,  $x_n \rightarrow x$  implies  $f(x_n) \rightarrow f(x)$ .

Let  $X^*$  denote the set of all bounded linear functionals on  $X$ . Define addition and scalar multiplication in  $X^*$  as follows:

$$(f_1 + f_2)(x) = f_1(x) + f_2(x) \text{ and } (\alpha f_1)(x) = \alpha f_1(x) \quad \text{for } f_1, f_2 \in X^* \text{ and } \alpha \in \mathbb{F}.$$

It can be checked that  $f_1 + f_2$  and  $\alpha f_1$  are in  $X^*$ .

Define a **norm on  $X^*$**  by setting

$$\|f\| = \sup_{x \neq 0} |f(x)| / \|x\| = \sup_{\|x\|=1} |f(x)|.$$

It can be checked that  $\|\cdot\|$  is a norm on  $X^*$ . It is immediate from the definition of the norm in  $X^*$  that

$$|f(x)| \leq \|f\| \cdot \|x\|.$$

It will be proved later that  $X^*$  with the norm described above is complete.

**Proposition 2.10.20** *A linear functional  $f: X \rightarrow \mathbb{C}$  is bounded if, and only if, it is a continuous functional.*

*Proof* Indeed, if  $f$  is bounded, then

$$\begin{aligned} |f(x_n) - f(x)| &= |f(x_n - x)| \\ &\leq \|f\| \|x_n - x\| \rightarrow 0 \end{aligned}$$

as  $x_n \rightarrow x$  and this implies the result in one direction.

Suppose  $x_n \rightarrow 0$ . Then  $f(x_n) \rightarrow 0$ . If  $f$  is not bounded, then for every  $n \in \mathbb{N}$ , there exists  $x_n$  with  $\|x_n\| = 1$  and  $|f(x_n)| \geq n$ . But in this case,  $|f(x_n/n)| \geq 1$ , whereas  $\|x_n/n\| \rightarrow 0$ .  $\square$

*Remark* The study of continuous linear functionals will be taken up in more detail later in the book.

**Proposition 2.10.21** *A linear functional  $f$  defined on  $X$  is continuous at  $x = 0$ , then it is continuous everywhere.*

*Proof* Suppose  $f$  is continuous at  $x = 0$ . Let  $\varepsilon > 0$  be given. There exists  $\delta > 0$  such that  $\|x\| < \delta$  implies  $|f(x)| < \varepsilon$ . Therefore, for every  $x \in X$ ,  $\|x - y\| < \delta$  implies  $|f(x) - f(y)| = |f(x - y)| < \varepsilon$ .  $\square$

*Remarks 2.10.22*

- (i) The point  $x = 0$  could be replaced by any other point of  $X$ .
- (ii) A slight modification of the argument in the proof above shows that  $f$  is uniformly continuous. Indeed, for every pair of points  $x, y$  in  $X$ ,  $\|x - y\| < \delta$  implies

$$|f(x) - f(y)| = |f(x - y)| < \varepsilon.$$

Thus, a linear functional is either uniformly continuous or everywhere discontinuous.

**Theorem 2.10.23** *If  $X$  is a normed linear space, then  $X^*$  is a Banach space.*

*Proof* Let  $\{f_n\}_{n \geq 1}$  be a Cauchy sequence of elements of  $X^*$ . This means that for any  $\varepsilon > 0$ , there exists  $n_0 \in \mathbb{N}$  such that  $m, n \geq n_0$  implies

$$\|f_n - f_m\| < \varepsilon,$$

that is, for any  $x \in X$ ,

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\| \|x\| < \varepsilon \|x\|. \quad (2.74)$$

In particular, for any  $x \in X$ ,  $\{f_n(x)\}_{n \geq 1}$  is a Cauchy sequence of scalars. So, the limit

$$\lim_n f_n(x) = f(x), \text{ say,}$$

must exist. The function  $f$  defined in this way is clearly linear. We next show that  $f$  is bounded. For any  $j \in \mathbb{N}$  and an appropriate  $n_1 \in \mathbb{N}$ , we have

$$\|f_{n_1+j} - f_{n_1}\| < 1,$$

which implies

$$\|f_{n_1+j}\| < (1 + \|f_{n_1}\|),$$

or

$$|f_{n_1+j}(x)| < (1 + \|f_{n_1}\|) \|x\|.$$

On letting  $j \rightarrow \infty$ , we obtain

$$|f(x)| \leq (1 + \|f_{n_1}\|) \|x\|.$$

Thus,  $f$  is a bounded linear functional in  $X$ . We next prove that  $\{f_n\}_{n \geq 1}$  converges to  $f$  in the norm of  $X^*$ . Using (2.74) again, for  $x \in X$  and  $n \geq n_0$ , we have

$$\lim_m |f_n(x) - f_m(x)| = |f_n(x) - \lim_m f_m(x)| \leq \varepsilon \|x\|.$$

Hence

$$|f_n(x) - f(x)| \leq \varepsilon \|x\|, \quad x \in X,$$

which implies  $\|f_n - f\| \leq \varepsilon$  for  $n \geq n_0$ . □

It is easy to write down the general linear functional on a finite-dimensional linear space. The description of continuous linear functionals on Banach spaces entails some efforts. However, not much effort is required to describe the continuous linear functionals on Hilbert spaces. We begin with some examples of continuous linear functionals.

#### Examples 2.10.24

- (i) Let  $H$  be a Hilbert space of finite dimension and let  $e_1, e_2, \dots, e_n$  be an orthonormal basis in  $H$ . If  $x = \sum \alpha_k e_k, \alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{C}$ , be any vector in  $H$ , then  $f(x) = \sum \alpha_k f(e_k)$  is clearly a linear functional in  $H$ . Moreover,

$$\begin{aligned}
|f(x)| &= \left| \sum_{k=1}^n \alpha_k f(e_k) \right| \leq \sum_{k=1}^n |\alpha_k| |f(e_k)| \\
&\leq \left( \sum_{k=1}^n |\alpha_k|^2 \right)^{\frac{1}{2}} \left( \sum_{k=1}^n |f(e_k)|^2 \right)^{\frac{1}{2}} \quad [\text{Cauchy-Schwarz Inequality}] \\
&= M \|x\|,
\end{aligned}$$

where  $M = \left( \sum_{k=1}^n |f(e_k)|^2 \right)^{\frac{1}{2}}$  and  $\|x\| = \left( \sum_{k=1}^n |\alpha_k|^2 \right)^{\frac{1}{2}}$ ; i.e.  $f$  is a bounded [continuous] linear functional.

- (ii) Consider the Hilbert space  $H = \ell^2$  of square summable sequences of scalars. For  $y = \{y_n\}_{n \geq 1}$  in  $\ell^2$ , define

$$f_y(x) = (x, y) = \sum_{n=1}^{\infty} x_n \overline{y_n}.$$

Observe that

$$\begin{aligned}
\left| \sum_{n=1}^{\infty} x_n \overline{y_n} \right| &\leq \sum_{n=1}^{\infty} |x_n| |y_n| \\
&\leq \left( \sum_{n=1}^{\infty} |x_n|^2 \right)^{\frac{1}{2}} \left( \sum_{n=1}^{\infty} |y_n|^2 \right)^{\frac{1}{2}} \\
&= \|x\|_2 \|y\|_2,
\end{aligned}$$

using the Cauchy-Schwarz Inequality. Thus,  $f_y$  is a bounded linear functional on  $\ell^2$  of norm at most  $\|y\|_2$ . For  $y \in \ell^2$ ,

$$f_y(y) = (y, y) = \sum_{n=1}^{\infty} |y_n|^2 = \|y\|_2^2,$$

showing that  $\|f_y\| = \|y\|_2$ .

- (iii) Consider the Hilbert space  $L^2(X, \mathfrak{M}, \mu)$  of complex-valued measurable functions  $f$  defined on  $X$  for which  $\int_X |f|^2 d\mu$  is finite. For  $g \in L^2(X, \mathfrak{M}, \mu)$ , define

$$f_g(h) = \int_X h \overline{g} d\mu, \quad h \in L^2(X, \mathfrak{M}, \mu).$$

Observe that

$$\left| \int_X h \bar{g} \, d\mu \right| \leq \left( \int_X |h|^2 \, d\mu \right)^{\frac{1}{2}} \left( \int_X |g|^2 \, d\mu \right)^{\frac{1}{2}} = \|h\|_2 \|g\|_2,$$

using Cauchy–Schwarz Inequality. Thus,  $f_g$  is a linear functional on  $L^2(X, \mathfrak{M}, \mu)$  of norm at most  $\|g\|_2$ . For  $g \in L^2(X, \mathfrak{M}, \mu)$ ,

$$f_g(g) = \int_X g \bar{g} \, d\mu = \int_X |g|^2 \, d\mu = \|g\|_2^2,$$

showing that  $\|f_g\| = \|g\|_2$ .

- (iv) Given a Hilbert space  $H$  and a vector  $y \in H$ , the function  $f_y(x) = (x, y)$ ,  $x \in H$ , is a bounded linear functional on  $H$  of norm  $\|y\|$ . Indeed,  $f$  is clearly linear and  $|f_y(x)| = |(x, y)| \leq \|x\| \|y\|$  and so,  $\|f_y\| \leq \|y\|$ . Furthermore,  $|f_y(y)| = \|y\|^2$  and so  $\|f_y\| = \|y\|$ .

Note that (i) and (ii) are special cases of (iii) with  $\mu$  as counting measure, and (iii) is a special case of (iv).

The existence of orthogonal decompositions implies that all bounded linear functionals on  $H$  can be obtained in this way.

**Theorem 2.10.25** (Riesz Representation Theorem) *Let  $H$  be a Hilbert space over  $\mathbb{C}$  and let  $f \in H^*$ , the space of all continuous linear functionals on  $H$ . Then there exists a unique vector  $y \in H$  such that  $f(x) = (x, y)$  for all  $x \in H$ .*

*Moreover, the mapping  $T: H \rightarrow H^*$  defined by  $T(y) = f_y$ , where  $f_y(x) = (x, y)$ , is onto, conjugate linear and isometric.*

*If  $H$  is a Hilbert space over  $\mathbb{R}$ , the mapping  $T$  is linear rather than conjugate linear.*

*Proof* Let  $f \in H^*$ . If  $f = 0$ , choose  $y = 0$ . Then  $f(x) = (x, y)$ ,  $x \in H$ . Furthermore,  $y = 0$  is the only such element of  $H$  since  $0 = f(y) = (y, y) = \|y\|^2$ .

Suppose that  $f \neq 0$  and let  $W = \{x \in H: f(x) = 0\}$ , known as the kernel of  $f$  and denoted by  $\ker(f)$ . Clearly,  $W$  is a linear subspace of  $H$ . Moreover,  $W$  is closed, so that  $W$  is a closed subspace of  $H$ . In fact,  $W$  is the inverse image of the closed set  $\{0\}$  under the continuous linear functional  $f$ . Since  $f \neq 0$ , we have  $W \neq H$ . So, by the Orthogonal Decomposition Theorem 2.10.11,  $H = W \oplus W^\perp$ . Since  $W^\perp \neq \{0\}$ , there exists  $y_0 \in W^\perp$ ,  $y_0 \neq 0$ . Clearly,  $f(y_0) \neq 0$  as  $y_0 \in W^\perp$ . Let  $y = [y_0 \overline{f(y_0)} / \|y_0\|^2]$ . For an arbitrary  $x \in H$ , we can form the element  $x - [f(x)/f(y_0)]y_0 \in W$ . Observe that

$$f(x - [f(x)/f(y_0)]y_0) = f(x) - [f(x)/f(y_0)]f(y_0) = 0.$$

So,  $x - [f(x)/f(y_0)]y_0 \in W$ . Consequently,

$$(x - [f(x)/f(y_0)]y_0, y_0) = 0, \text{ i.e., } (x, y_0) = [f(x)/f(y_0)]\|y_0\|^2,$$

which implies  $f(x) = (x, y)$ .

We next show that  $y$  is unique. Assuming the contrary, we have the equation

$$(x, y') = (x, y'')$$

for  $x \in H$  where  $y' \neq y''$ . But this is impossible, since the substitution  $x = y' - y''$  yields the contradiction  $\|y' - y''\| = 0$ . The fact that  $\|f\| = \|y\|$  was proved in Example 2.10.24(iv).

The mapping  $T: H \rightarrow H^*$  defined by  $T(y) = f_y$ , where  $f_y(x) = (x, y)$ , is conjugate linear:  $T(\alpha y + \beta z) = f_{\alpha y + \beta z}$ , where  $f_{\alpha y + \beta z}(x) = (x, \alpha y + \beta z) = \bar{\alpha}(x, y) + \bar{\beta}(x, z) = \bar{\alpha}f_y(x) + \bar{\beta}f_z(x)$ . Thus,  $T(\alpha y + \beta z) = \bar{\alpha}f_y + \bar{\beta}f_z$ .

The real case is left to the reader. This completes the proof.  $\square$

#### Remarks 2.10.26

- (i) The functionals defined on  $\ell^2$  and  $L^2(X, \mathbb{F}, \mu)$  in Examples 2.10.24(ii) and (iii) are the only continuous functionals on these spaces. To prove the statement without the use of the theorem will need quite an effort. The linear functionals defined in Example 2.10.24(i) are the only ones possible on that space.
- (ii) The Riesz Representation Theorem has been proved for a Hilbert space. The hypothesis that the space is complete is essential for the theorem to hold. Consider the pre-Hilbert space  $\ell_0$  of finitely nonzero sequences. Define

$$f(x) = \sum_{n=1}^{\infty} \frac{x(n)}{n}, \quad x \in \ell_0 \quad \text{and} \quad x = \{x(n)\}_{n \geq 1}.$$

Clearly,  $f$  is linear and

$$|f(x)| = \left| \sum_{n=1}^{\infty} \frac{x(n)}{n} \right| \leq \left( \sum_{n=1}^{\infty} |x(n)|^2 \right)^{\frac{1}{2}} \left( \sum_{n=1}^{\infty} \frac{1}{n^2} \right)^{\frac{1}{2}} = M\|x\|,$$

where  $M = \left( \sum_{n=1}^{\infty} \frac{1}{n^2} \right)^{\frac{1}{2}}$ , using the Cauchy-Schwarz Inequality. Thus,  $f$  is a bounded linear functional on  $\ell_0$ . However, there exists no  $y \in \ell_0$  for which  $f(x) = (x, y)$ . Indeed, for  $x = e_n = (0, 0, \dots, 0, 1, 0, \dots)$ , where 1 occurs in the  $n$ th place,  $f(x) = \frac{1}{n}$ , while  $(x, y) = \overline{y(n)}$ , so that  $y(n) = \frac{1}{n}$ . Consequently,  $y \notin \ell_0$ .

*Remark 2.10.27* In fact, every incomplete inner product space has a continuous linear functional that cannot be represented by an element of the space. Indeed, the linear functional defined by a vector in the completion but not in the incomplete space is the desired linear functional.

Let  $Y$  be a subspace of a normed linear space  $X$  and  $f$  be a bounded linear functional defined on  $Y$ . Then  $f$  can be extended to the whole of  $X$ , so that both the functional and its extension have the same norm [Theorem 5.3.2]. Apart from the fact that the procedure of extension is involved, the extension is not unique. However, the existence of an extension of a continuous linear functional defined on a subspace of a Hilbert space  $H$  to  $H$  is a direct consequence of the Riesz Representation Theorem 2.10.25. Moreover, the extension is unique.

**Theorem 2.10.28** *Let  $H$  be a Hilbert space,  $Y$  a subspace of  $H$  and  $f$  a continuous linear functional defined on  $Y$ . Then, there exists a unique  $F \in H^*$  such that  $F|_Y = f$  and  $\|f\|_Y = \|F\|_H$ ,*

$$\|f\|_Y = \sup_{\substack{\|x\|=1 \\ x \in Y}} |f(x)| \quad \text{and} \quad \|F\|_H = \sup_{\substack{\|x\|=1 \\ x \in H}} |F(x)|.$$

*Proof* Since  $f$  is linear and continuous, it follows that it is uniformly continuous on  $Y$  [Remark 2.10.22(ii)]. Hence,  $f$  can be extended to  $\bar{Y}$ , the closure of  $Y$  with the preservation of norm. This is shown as follows:

Let  $x \in \bar{Y}$ . There exists a sequence  $\{x_n\}$  in  $Y$  such that  $x_n \rightarrow x$  and, in view of the linearity and continuity of  $f$ ,  $\lim_n f(x_n)$  exists; moreover, it is independent of the sequence chosen. We define  $f(x)$  to be  $\lim_n f(x_n)$ . Then  $|f(x_n)| \leq \|f\|_Y \|x_n\|$  and this implies  $|f(x)| \leq \|f\|_Y \|x\|$ . Consequently, the norm of the extended  $f$  is at most the norm of the given  $f$ . The reverse inequality is trivial.

We may therefore assume without loss of generality that  $Y$  is a closed subspace of  $H$ . The Riesz Representation Theorem 2.10.25 asserts the existence of a unique element  $y \in Y$  such that

$$f(x) = (x, y) \quad \text{for all } x \in Y,$$

and

$$\|f\|_Y = \|y\|.$$

We now extend  $f$  to the whole of  $H$  by defining

$$F(x) = (x, y) \quad \text{for all } x \in H,$$

i.e.  $F(x) = 0$  if  $x \in Y^\perp$  and  $F|_Y = f$ . It is clear from the definition of  $F$  that

$$\|F\|_H = \|y\| = \|f\|_Y.$$

We shall next show that any other extension of the linear functional  $f$  to the whole space increases the norm. Indeed, if  $F'$  is any other extension of  $f$  to the whole space, then

$$F'(x) = (x, z)$$

and

$$\|F'\| = \|z\|.$$

For  $x \in Y$ ,

$$(x, y) = (x, z),$$

so that  $y - z \perp Y$ . Because  $y \in Y$ ,

$$\|z\|^2 = \|y\|^2 + \|y - z\|^2,$$

which implies that

$$\|F'\| \geq \|f\|_Y,$$

where there is strict inequality if  $y \neq z$ . □

We note in passing that if  $\bar{Y}$  is not the whole space  $H$ , then there exist extensions of arbitrarily large norm.

Let  $X$  be a normed linear space over  $\mathbb{C}$ . The space  $X^*$  of all bounded linear functionals on  $X$  is a Banach space. One can then consider continuous linear functionals on  $X^*$ , that is, the space  $(X^*)^* = X^{**}$ . By the preceding remark,  $X^{**}$  is again a Banach space. The element  $x \in X$  defines a continuous linear functional on  $X^*$ ; that is,  $x$  determines an element  $\tau(x)$  of  $X^{**}$  defined by

$$\tau(x)(x^*) = x^*(x), \quad x^* \in X^*. \quad (2.75)$$

It is apparent that  $\tau(x)$  is linear. Also, the inequality

$$|\tau(x)(x^*)| = |x^*(x)| \leq \|x^*\| \|x\|$$

shows that  $\tau(x) \in X^{**}$  and  $\|\tau(x)\| \leq \|x\|$ . One learns in a course on Banach spaces that the mapping  $\tau : X \rightarrow X^{**}$  defined by (2.75) above is an isometric isomorphism. The space is said to be **reflexive** if the mapping  $\tau$  defined by (2.75) above is surjective. Not all Banach spaces are reflexive. However, all finite-dimensional

normed linear spaces are. One of the distinguishing features of a Hilbert space is that it is reflexive. We begin by showing that  $H^*$ , the dual of a Hilbert space  $H$ , is itself a Hilbert space.

**Theorem 2.10.29** *If  $H$  is a Hilbert space, then  $H^*$  is a Hilbert space. Moreover, there exists a conjugate linear map  $T:H \rightarrow H^*$  which is one-to-one, onto, norm preserving and satisfies*

$$(f_1, f_2)_{H^*} = (T^{-1}f_2, T^{-1}f_1).$$

*Proof* By Theorem 2.10.23,  $H^*$  is a complete normed linear space. Consider the mapping  $T:H \rightarrow H^*$  defined by

$$T(x)(y) = (y, x), \quad x, y \in H. \quad (2.76)$$

Note that  $T$  defined on  $H$  and given by (2.76) is conjugate linear, one-to-one, norm preserving and onto [Theorem 2.10.25]. Therefore,  $T^{-1}$  exists.

Define an inner product on  $H^*$  as follows: Given  $f_1, f_2 \in H^*$ , let

$$(f_1, f_2)_{H^*} = (T^{-1}f_2, T^{-1}f_1).$$

It is easy to check that this defines an inner product on  $H^*$ . It is related to the norm by  $(f_1, f_1)_{H^*} = \|f_1\|_{H^*}^2$ , because

$$(f_1, f_1)_{H^*} = (T^{-1}f_1, T^{-1}f_1) = \|T^{-1}f_1\|^2 = \|f_1\|_{H^*}^2.$$

Thus  $H^*$  is a Hilbert space. □

**Theorem 2.10.30** *If  $H$  is a Hilbert space and  $H^{**} = (H^*)^*$ , then the mapping  $\tau:H \rightarrow H^{**}(x \rightarrow \tau(x))$ , where the defining equation for  $\tau(x)$  is*

$$\tau(x)(f) = f(x) \quad f \in H^*,$$

*is an isometric isomorphism between  $H$  and  $H^{**}$ . Thus  $H$  is reflexive.*

*Proof* Let  $T:H \rightarrow H^*$  and  $S:H^* \rightarrow H^{**}$  be the conjugate linear maps assured by Theorem 2.10.29 (used twice). Since both are conjugate linear, one-to-one, norm preserving and onto, we know that the composition  $ST:H \rightarrow H^{**}$  is linear, one-to-one, norm preserving and onto, which is to say that it is an isometric isomorphism between  $H$  and  $H^{**}$ . Thus, we need only to prove that

$$ST(x)(f) = f(x) \quad f \in H^*,$$

and set  $\tau = ST$ .

By Theorem 2.10.29, we also have

$$T(x)(y) = (y, x)_H, \quad x, y \in H, \quad (2.77)$$

$$S(g)(f) = (f, g)_{H^*}, \quad f, g \in H^*, \quad (2.78)$$

$$(f, g)_{H^*} = (T^{-1}g, T^{-1}f)_H, \quad f, g \in H^*. \quad (2.79)$$

We compute  $ST$  from here as follows:

$$\begin{aligned} ST(x)(f) &= S(Tx)(f) \\ &= (f, Tx)_{H^*} \quad \text{by (2.78)} \\ &= (T^{-1}(Tx), T^{-1}f)_H \quad \text{by (2.79)} \\ &= (x, T^{-1}f)_H = (T(T^{-1}f))(x) \quad \text{by (2.77)} \\ &= f(x). \end{aligned}$$

□

The following result is an analogue of a familiar result from metric spaces.

**Theorem 2.10.31** *In order that the linear span of a system  $M$  of vectors is dense in  $H$ , it is necessary and sufficient that a continuous linear functional  $f \in H^*$  which vanishes for all  $x \in M$  must be identically zero.*

*Proof* Necessity: suppose the linear span of  $M$  is dense in  $H$ , i.e.  $[\overline{M}] = H$  and  $f \in H^*$  vanishes for all  $x \in M$ . By linearity,  $f$  vanishes on  $[M]$  and hence by continuity it vanishes on  $[\overline{M}]$ , which is the same as  $H$ .

Sufficiency: suppose the linear span of  $M$  is not dense in  $H$ , i.e.  $[\overline{M}] \neq H$ . Then there exists  $y \neq 0$  such that  $y \perp [\overline{M}]$ . Then the linear functional  $f$  defined by  $f(x) = (x, y)$  for all  $x \in H$  vanishes for all  $x \in M$  but is not identically zero because  $f(y) = (y, y) \neq 0$ . □

### Problem Set 2.10

2.10.P1. Let  $f \in RH^2$  have the series expansion  $f = \sum_{j=0}^{\infty} a_j z^j$ . Define  $C_n(f) = a_n$ .

Show that  $C_n$  is a continuous linear functional on  $RH^2$ .

2.10.P2. Let  $e_0(t) = 1$  and  $e_1(t) = \sqrt{3}(2t - 1)$ ,  $t \in [0, 1]$ , be vectors in the Hilbert space  $L^2[0, 1]$ . Show that  $e_0 \perp e_1$ ,  $\|e_0\| = \|e_1\| = 1$ . Compute the vector  $y$  in the linear span of  $\{e_0, e_1\}$  closest to  $t^2$  and also compute  $\min_{a,b} \int_0^1 |t^2 - a - bt|^2 dt$ .

2.10.P3. Let  $X = \mathbb{R}^2$ . Find  $M^\perp$

- (a)  $M = \{x\}$ , where  $x = (\xi_1, \xi_2) \neq 0$ ;
- (b)  $M$  is a linearly independent set  $\{x_1, x_2\} \subseteq X$ .

2.10.P4. For any subset  $M \neq \emptyset$  of a Hilbert space  $H$ ,  $\text{span}(M)$  is dense in  $H$  if, and only if,  $M^\perp = \{0\}$ .

- 2.10.P5. (a) Prove that for any two subspaces  $M_1$  and  $M_2$  of a Hilbert space  $H$ , we have  $(M_1 + M_2)^\perp = M_1^\perp \cap M_2^\perp$ .
- (b) Prove that for any two closed subspaces  $M_1$  and  $M_2$  of a Hilbert space  $H$ , we have

$$(M_1 \cap M_2)^\perp = \overline{M_1^\perp + M_2^\perp}.$$

- 2.10.P6. (a) Let  $K_1$  and  $K_2$  be the nonempty, closed and convex subsets of a Hilbert space  $H$  such that  $K_1 \subseteq K_2$ . Prove that, for all  $x \in H$ ,

$$\|y_1 - y_2\|^2 \leq 2(d(x, K_1)^2 - d(x, K_2)^2),$$

where  $y_1$  and  $y_2$  are closest points of  $x$  in  $K_1$  and  $K_2$ , respectively.

- (b) Let  $\{K_n\}_{n \geq 1}$  be an increasing sequence of nonempty closed convex subsets in  $H$  and let  $K = \overline{\bigcup_n K_n}$ . Prove that  $K$  is closed and convex. Also show that  $\lim_n y_n = y$  for all  $x \in H$ , where  $y_n$  is the projection of  $x$  onto  $K_n$ ,  $n = 1, 2, \dots$ , and  $y$  is the projection of  $x$  onto  $K$ .
- 2.10.P7. Let  $M$  be a closed subspace of a Hilbert space  $H$  and  $x_0 \in H$ . Prove that  $\min\{\|x_0 - x\|: x \in M\} = \max\{|\langle x_0, y \rangle|: y \in M^\perp \text{ and } \|y\| = 1\}$ .
- 2.10.P8. (a) Let  $a$  be a nonzero element of a Hilbert space  $H$ . Prove that, for all  $x \in H$ ,

$$d(x, \{a\}^\perp) = \frac{|\langle x, a \rangle|}{\|a\|}.$$

- (b) Let  $H = L^2[0, 1]$  and let

$$F = \left\{ f \in L^2[0, 1] : \int_0^1 f(x) dx = 0 \right\}.$$

Determine  $F^\perp$ . For  $f(x) = \exp(x)$ , determine  $d(f, F)$ .

- 2.10.P9. In the linear space  $C[0, 1]$ , consider the functional  $F(x) = \int_0^1 x(t)f(t)dt$ , where  $f$  is a continuous function defined on  $[0, 1]$ . Show that  $\|F\| = \int_0^1 |f(t)|dt$ .
- 2.10.P10. Let  $H$  be a Hilbert space and  $f$  be a nonzero continuous linear functional on  $H$ , i.e.  $f \in H^* \setminus \{0\}$ . Show that  $\dim((\ker(f))^\perp) = 1$ .
- 2.10.P11. Prove that if  $f$  is a linear functional on a Hilbert space  $H$  and  $\ker(f)$  is closed, then  $f$  is bounded.

- 2.10.P12. Show that the subspace  $M = \{x = \{x_n\}_{n \geq 1} \in \ell^2 : \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} x_n = 0\}$  is not a closed subspace of  $\ell^2$ .
- 2.10.P13. Prove that the system  $\sin nx$ ,  $n = 1, 2, \dots$ , is complete in  $L^2[0, \pi]$ .
- 2.10.P14. Let  $K$  be a nonempty closed convex set in a Hilbert space  $H$ . Show that  $K$  contains a unique vector  $k$  of smallest norm and that  $\Re(k, k - x) \leq 0$  for all  $x \in K$ . Moreover, if  $k \in K$  satisfies  $\Re(k, k - x) \leq 0$  for all  $x \in K$ , then  $k$  is the vector of smallest norm in  $K$ .
- 2.10.P15. Let  $y$  be a nonzero vector in a Hilbert space  $H$  and let

$$M = \{x \in H : (x, y) = 0\}.$$

What is  $M^\perp$ ?

## 2.11 Approximation in Hilbert Spaces

Let  $H$  be a Hilbert space and  $v_1, v_2, \dots, v_k$  be linearly independent vectors in  $H$ . Suppose that  $x \in H$ . In linear approximation, it is required to find a method of computing the minimum value of the quantity

$$\left\| x - \sum_{j=1}^n \lambda_j v_j \right\|,$$

where  $\lambda_1, \lambda_2, \dots, \lambda_n$  range over all scalars and also determine those values of  $\lambda_1, \lambda_2, \dots, \lambda_n$  for which the minimum is attained.

Let  $M$  be the closed linear space generated by linearly independent vectors  $v_1, v_2, \dots, v_n$  and  $x \in H$ . By Theorem 2.10.10, there exists a unique minimising vector  $y \in M$  and  $x - y \perp M$ . Since  $y \in M$ , we have  $(x - y, y) = 0$ .

Denote  $(v_j, v_i) = a_{ij}$  and  $b_i = (x, v_i)$ . If  $y = \sum_{j=1}^n c_j v_j$  is the minimising vector, then

$$(x - y, v_i) = 0 \quad \text{for } i = 1, 2, \dots, n,$$

which, written in full, reads as

$$b_i = \sum_{j=1}^n a_{ij} c_j, \quad i = 1, 2, \dots, n.$$

Since the vectors  $\{v_i\}$  are linearly independent, the matrix  $[a_{ij}]$  is nonsingular. Consequently, the  $n$  linear equations in  $n$  unknowns  $c_1, c_2, \dots, c_n$  have a unique solution.

Let  $\delta = \inf \left\{ \left\| x - \sum_{j=1}^n \lambda_j v_j \right\| : \lambda_1, \lambda_2, \dots, \lambda_n \text{ are scalars} \right\}$ . Then

$$\delta^2 = \|x - y\|^2 = (x - y, x - y) = \left( x, x - \sum_{j=1}^n c_j v_j \right) = \|x\|^2 - \sum_{j=1}^n \overline{c_j} b_j \quad (2.80)$$

If we replace  $v_1, v_2, \dots, v_n$  by an orthonormal set  $u_1, u_2, \dots, u_n$ , then  $a_{ij} = 1$  if  $i = j$  and 0 if  $i \neq j$ . Hence,  $c_j = b_j, j = 1, 2, \dots, n$  and it follows from (2.80) that

$$\delta^2 = \|x\|^2 - \sum_{j=1}^n |b_j|^2 = \|x\|^2 - \sum_{j=1}^n |(x, u_j)|^2.$$

We have thus proved the following theorem.

**Theorem 2.11.1** *Let  $\{u_1, u_2, \dots, u_n\}$  be an orthonormal set in  $H$  and let  $x \in H$ . Then*

$$\left\| x - \sum_{k=1}^n (x, u_k) u_k \right\| \leq \left\| x - \sum_{k=1}^n \lambda_k u_k \right\|$$

for all scalars  $\lambda_1, \lambda_2, \dots, \lambda_n$ . Equality holds if, and only if,  $\lambda_k = (x, u_k), k = 1, 2, \dots, n$ . Moreover,  $\sum_{k=1}^n \lambda_k u_k$  is the orthogonal projection of  $x$  onto the subspace  $M$  generated by  $\{u_1, u_2, \dots, u_n\}$ , and if  $\delta$  is the distance of  $x$  from  $M$ , then  $\delta^2 = \|x\|^2 - \sum_{k=1}^n |(x, u_k)|^2$ .

**Remark 2.11.2** If the subspace  $M$  generated by  $(n + 1)$  orthonormal vectors and it is desired to obtain the distance of  $x \in H$  from  $M$ , then

$$\delta^2 = \|x - y\|^2 = \|x\|^2 - \sum_{k=1}^{n+1} |(x, u_k)|^2, \quad (2.81)$$

where  $y$  is the orthogonal projection of  $x$  on  $M$  and is given by

$$y = \sum_{k=1}^{n+1} (x, u_k) u_k. \quad (2.82)$$

The reader will notice that the first  $n$  components in the sums on the right of (2.81) and (2.82) remain unaltered when the dimension of the space is increased from  $n$  to  $n + 1$ . This exhibits the importance of orthonormalising the linearly independent vectors.

*Example 2.11.3* Consider the real inner product space  $C[-1, 1]$ , the inner product  $(x, y)$ ,  $x, y \in C[-1, 1]$  being defined by

$$(x, y) = \int_{-1}^1 x(t)y(t)dt.$$

Consider the three linearly independent vectors  $1, t, t^2$  (the Wronskian of the vectors  $1, t, t^2$  is  $2 \neq 0$ ) in  $C[-1, 1]$ . The Gram–Schmidt orthonormalisation process yields

$$u_0(t) = \frac{1}{\sqrt{2}}, \quad u_1(t) = \sqrt{\frac{3}{2}}t, \quad u_2(t) = \sqrt{\frac{5}{2}}\frac{1}{2}(3t^2 - 1).$$

Let  $M_2$  [respectively,  $M_3$ ] be the linear space generated by  $\{u_0, u_1\}$  [respectively  $\{u_0, u_1, u_2\}$ ]. Consider  $x(t) = e^t$  in  $C[-1, 1]$ . We shall compute the distance of  $x$  from  $M_2$  and  $M_3$ .

$$(x, u_0) = \frac{1}{\sqrt{2}} \int_{-1}^1 e^t dt = \frac{e - e^{-1}}{\sqrt{2}},$$

$$(x, u_1) = \sqrt{\frac{3}{2}} \int_{-1}^1 te^t dt = \sqrt{6}e^{-1}$$

$$(x, u_2) = \sqrt{\frac{5}{2}}\frac{1}{2} \int_{-1}^1 (3t^2 - 1)e^t dt = \sqrt{\frac{5}{2}}(e - 7e^{-1}).$$

Let  $y_2$  and  $y_3$  be the projections of  $x(t) = e^t$  on the subspaces  $M_2$  and  $M_3$ , respectively. Then

$$\begin{aligned} y_2 &= (x, u_0)u_0 + (x, u_1)u_1 \\ &= \frac{1}{\sqrt{2}} \frac{e - e^{-1}}{\sqrt{2}} + \sqrt{6}e^{-1} \sqrt{\frac{3}{2}}t \\ &= \frac{1}{2}(e - e^{-1}) + 3e^{-1}t \end{aligned}$$

and

$$\begin{aligned}
 y_3 &= (x, u_0)u_0 + (x, u_1)u_1 + (x, u_2)u_2 \\
 &= \frac{1}{2}(e - e^{-1}) + 3e^{-1}t + \sqrt{\frac{5}{2}}(e - 7e^{-1})\sqrt{\frac{5}{2}}\frac{1}{2}(3t^2 - 1) \\
 &= \frac{1}{2}(e - e^{-1}) + 3e^{-1}t + \frac{5}{4}(e - 7e^{-1})(3t^2 - 1).
 \end{aligned}$$

If  $\delta_2$  [respectively,  $\delta_3$ ] denotes the distance of  $x$  from  $M_2$  [respectively,  $M_3$ ], then by Theorem 2.11.1,

$$\begin{aligned}
 \delta_2^2 &= \|x\|^2 - |(x, u_0)|^2 - |(x, u_1)|^2 \\
 &= \frac{1}{2}(e^2 - e^{-2}) - \left(\frac{e - e^{-1}}{\sqrt{2}}\right)^2 - 6e^{-2} \\
 &= 1 - 7e^{-2}
 \end{aligned}$$

and

$$\begin{aligned}
 \delta_3^2 &= \|x\|^2 - |(x, u_0)|^2 - |(x, u_1)|^2 - |(x, u_2)|^2 \\
 &= \delta_2^2 - \frac{5}{2}(e - 7e^{-1})^2 \\
 &= 1 - 7e^{-2} - \frac{5}{2}(e^2 + 49e^{-2} - 14) \\
 &= 36 - \frac{5}{2}e^2 - \frac{259}{2}e^{-2}.
 \end{aligned}$$

### Problem Set 2.11

2.11.P1. Find  $\min_{a,b,c} \int_{-1}^1 |t^3 - a - bt - ct^2|^2 dt$  and  $\max_{-1}^1 t^3 g(t) dt$ , where  $g$  is subject to the restrictions

$$\int_{-1}^1 g(t) dt = \int_{-1}^1 t g(t) dt = \int_{-1}^1 t^2 g(t) dt = 0, \quad \int_{-1}^1 |g(t)|^2 dt = 1.$$

2.11.P2. Find the point nearest to  $(1, -1, 1)$  in the linear span of  $(1, \omega, \omega^2)$  and  $(1, \omega^2, \omega)$  in  $\mathbb{C}^3$ , where  $\omega = \exp(2\pi i/3)$ .

## 2.12 Weak Convergence

Let  $\{x_n\}_{n \geq 1}$  be a sequence in a Hilbert space  $H$ . Recall that  $\{x_n\}_{n \geq 1}$  converges to  $x$  in  $H$  if  $\|x_n - x\| = (x_n - x, x_n - x)^{\frac{1}{2}} \rightarrow 0$  as  $n \rightarrow \infty$ , and we write  $x_n \rightarrow x$ . From now on it will be called *strong convergence* to distinguish it from *weak convergence*, to be introduced shortly. The relationship between the two types of convergence will be discussed. The concepts of strong convergence and weak convergence are identical in finite-dimensional spaces. A characterisation of weak convergence in special spaces will also find a mention below.

**Definition 2.12.1** A sequence of vectors  $\{x_n\}_{n \geq 1}$  **converges weakly** to a vector  $x$  and we write  $x_n \rightarrow x$  ( $x_n \xrightarrow{w} x$ ) if

$$\lim_{n \rightarrow \infty} (x_n, y) = (x, y)$$

for all

$$y \in H.$$

The concepts of a weakly Cauchy sequence and weak completeness are defined analogously.

*Remarks 2.12.2*

- (i) A sequence cannot converge weakly to two different limits: assume that  $x_n \xrightarrow{w} x_0$  and  $x_n \xrightarrow{w} y_0$ . Then

$$(x_n, y) \rightarrow (x_0, y) \text{ and } (x_n, y) \rightarrow (y_0, y)$$

for all  $y \in H$ . Consequently,  $(x_0, y) = (y_0, y)$ , or  $(x_0 - y_0, y) = 0$ , for all  $y \in H$ . If we choose  $y = x_0 - y_0$ , we obtain  $(x_0 - y_0, x_0 - y_0) = 0$ , which implies  $x_0 = y_0$ .

- (ii) If  $x_n \xrightarrow{w} x_0$ , then every arbitrary subsequence  $\{x_{n_k}\}_{k \geq 1}$  converges weakly to  $x_0$ .
- (iii) Strong convergence of  $\{x_n\}_{n \geq 1}$  to  $x_0$  implies  $x_n \xrightarrow{w} x_0$ . Indeed, for  $y \in H$ , we have

$$|(x_n - x_0, y)| \leq \|x_n - x_0\| \|y\|,$$

by the Cauchy-Schwarz Inequality.

- (iv) The converse of (iii) is, however, not true. Indeed, let  $\{e_n\}_{n \geq 1}$  be an infinite orthonormal sequence of vectors in  $H$ . Since for any  $y \in H$ ,

$$\sum_{n=1}^{\infty} |(y, e_n)|^2 \leq \|y\|^2 \quad (\text{by Bessel's Inequality}),$$

therefore,  $\lim_{n \rightarrow \infty} (e_n, y) = 0$ . Thus, the sequence  $\{e_n\}_{n \geq 1}$  converges weakly to the vector zero, but this sequence cannot converge strongly, since

$$\|e_i - e_j\|^2 = 2 \quad (i \neq j),$$

so that  $\|e_i - e_j\| \not\rightarrow 0$  as  $i, j \rightarrow \infty$ .

However, the following theorem holds:

**Theorem 2.12.3** *If  $H$  is a finite-dimensional Hilbert space, strong convergence is equivalent to weak convergence.*

*Proof* Since we have already shown that, in any Hilbert space, strong convergence implies weak convergence [Remark 2.12.2(iii)], it is enough to show in this situation that weak convergence implies strong convergence. To this end, let  $e_1, \dots, e_k$  be an orthonormal basis for  $H$  and let

$$x_n \xrightarrow{w} x,$$

where

$$x_n = \alpha_1^{(n)} e_1 + \dots + \alpha_k^{(n)} e_k$$

for  $n = 1, \dots, k$ , and

$$x = \alpha_1 e_1 + \dots + \alpha_k e_k.$$

Since  $x_n \xrightarrow{w} x$ , it follows that

$$(x_n, e_j) \rightarrow (x, e_j), \quad \text{i.e., } \alpha_j^{(n)} \rightarrow \alpha_j$$

for  $j = 1, \dots, k$ . For any prescribed  $\varepsilon > 0$ , there must be an integer  $n_0$  such that for all  $n > n_0$  and for every  $j = 1, \dots, k$ ,

$$|\alpha_j^{(n)} - \alpha_j| < \varepsilon/k;$$

hence

$$\|x_n - x\|^2 = \left\| \sum_{j=1}^k (\alpha_j^{(n)} - \alpha_j) e_j \right\|^2 = \sum_{j=1}^k |\alpha_j^{(n)} - \alpha_j|^2 < \varepsilon.$$

Thus,  $x_n \rightarrow x$  strongly. This completes the proof.  $\square$

The next result pinpoints the relationship between weak and strong convergences.

**Theorem 2.12.4** *Let  $\{x_n\}_{n \geq 1}$  be a sequence in a Hilbert space  $H$ . Then  $x_n \rightarrow x$  if, and only if,  $x_n \xrightarrow{w} x$  and  $\limsup_{n \rightarrow \infty} \|x_n\| \leq \|x\|$ .*

*Proof* Let  $x_n \rightarrow x$ . Then  $x_n \xrightarrow{w} x$  [Remark 2.12.2(iii)]. Also,  $\limsup_{n \rightarrow \infty} \|x_n\| = \lim_{n \rightarrow \infty} \|x_n\| = \|x\|$ , since  $0 \leq \|x_n - x\| \geq \left| \|x_n\| - \|x\| \right|$ .

Conversely, let  $x_n \xrightarrow{w} x$  and  $\limsup_{n \rightarrow \infty} \|x_n\| \leq \|x\|$ . For each  $n$ ,  $0 \leq \|x_n - x\|^2 = (x_n - x, x_n - x) = \|x_n\|^2 + \|x\|^2 - 2\Re(x_n, x)$ . Since

$$\limsup_{n \rightarrow \infty} \|x_n\| \leq \|x\| \text{ and } \Re(x_n, x) \rightarrow \Re(x, x) = \|x\|^2,$$

we have

$$0 \leq \limsup_{n \rightarrow \infty} \|x_n - x\| \leq \|x\|^2 + \|x\|^2 - 2\|x\|^2 = 0,$$

so that  $\lim_{n \rightarrow \infty} \|x_n - x\|$  exists and equals zero.  $\square$

The Riesz Representation Theorem enables us to prove the following analogue of the classical Bolzano–Weierstrass Theorem.

**Theorem 2.12.5** *Any bounded sequence in  $H$  has a weakly convergent subsequence and the limit has the same bound.*

*Proof* Let  $\{x_n\}_{n \geq 1}$  be a sequence in  $H$  and an  $M > 0$  be such that  $\|x_n\| \leq M$  for all  $n$ . We need to find a weakly convergent subsequence of  $\{x_n\}_{n \geq 1}$ .

By the Cauchy–Schwarz Inequality,

$$|(x_n, x_1)| \leq \|x_n\| \|x_1\| \leq M^2$$

for all  $n$ . The classical Bolzano–Weierstrass Theorem shows that the bounded sequence  $\{(x_n, x_1)\}_{n \geq 1}$  has a convergent subsequence  $\{(x_{n(1)}, x_1)\}_{n(1) \geq 1}$ , say. Applying the preceding argument to the sequence  $\{(x_{n(1)}, x_2)\}_{n(1) \geq 1}$ , we extract a convergent subsequence  $\{(x_{n(2)}, x_2)\}_{n(2) \geq 1}$ .

Continuing inductively, we obtain for each  $k$  a convergent subsequence  $\{(x_{n(k)}, x_k)\}_{n(k) \geq 1}$  of  $\{(x_{n(k-1)}, x_k)\}_{n(k-1) \geq 1}$ .

Consider now the diagonal sequence  $\{z_p\}_{p \geq 1}$ , where  $z_p$  denotes the  $p$ th term of the sequence  $\{x_{n(p)}\}_{n(p) \geq 1}$ . We show that for each  $x \in H$ , the sequence  $\{(z_p, x)\}_{p \geq 1}$  of scalars converges.

If  $x = x_m$  for some  $m$ , then for each  $p > m$ ,  $\{(z_p, x)\}_{p \geq 1}$  is a subsequence of the convergent sequence  $\{(z_p, x_m)\}_{p \geq 1}$  and is, therefore, convergent. Hence, if  $x \in \text{span}\{x_1, x_2, \dots\}$ , then  $\{(z_p, x)\}_{p \geq 1}$  converges in the field of scalars.

Let  $x \in \overline{\text{span}\{x_1, x_2, \dots\}}$ . Consider a sequence  $\{y_r\}_{r \geq 1}$  in  $\text{span}\{x_1, x_2, \dots\}$  such that  $y_r \rightarrow x$  as  $r \rightarrow \infty$ . Then for all  $n, m$  and  $r$ , we have

$$\begin{aligned} |(z_n, x) - (z_m, x)| &= |(z_n - z_m, x)| \\ &\leq |(z_n - z_m, x - y_r)| + |(z_n - z_m, y_r)| \\ &\leq \|z_n - z_m\| \|x - y_r\| + |(z_n - z_m, y_r)| \\ &\leq 2M \|x - y_r\| + |(z_n - z_m, y_r)|. \end{aligned}$$

Since  $\|x - y_r\| \rightarrow 0$  as  $r \rightarrow \infty$  and  $|(z_n - z_m, y_r)| \rightarrow 0$  as  $n, m \rightarrow \infty$  for each  $r$ , we see that  $\{(z_n, x)\}_{n \geq 1}$  is a Cauchy sequence of scalars and is, therefore, convergent. Next, let  $x \perp \overline{\text{span}\{x_1, x_2, \dots\}}$ . Then  $(z_n, x) = 0$  for all  $n$ , since  $z_n$  is in  $\text{span}\{x_1, x_2, \dots\}$ . Thus  $(z_n, x) \rightarrow 0$  as  $n \rightarrow \infty$ .

By the Orthogonal Decomposition Theorem 2.10.11,

$$H = \overline{\text{span}\{x_1, x_2, \dots\}} \oplus \overline{\text{span}\{x_1, x_2, \dots\}}^\perp.$$

Hence,  $\{(z_n, x)\}_{n \geq 1}$  converges for each  $x \in H$ . Define

$$f(x) = \lim_{n \rightarrow \infty} (x, z_n), \quad x \in H. \quad (2.83)$$

Clearly,  $f$  is linear and

$$|f(x)| = \lim_{n \rightarrow \infty} |(x, z_n)| \leq M \|x\|$$

for all  $x \in H$ . Thus,  $f$  is a continuous linear functional on  $H$  satisfying  $\|f\| \leq M$ . By the Riesz Representation Theorem 2.10.25, there exists a unique  $y \in H$  such that

$$f(x) = (x, y), \quad x \in H. \quad (2.84)$$

and  $\|y\| = \|f\| \leq M$ . On comparing (2.83) and (2.84), we obtain

$$\lim_{n \rightarrow \infty} z_n = y \text{ (weak)}.$$

This completes the proof.  $\square$

Every convergent sequence in a normed linear space  $X$  is bounded. This is easily seen as follows: let  $\{x_n\}_{n \geq 1}$  be a sequence in  $X$  and suppose that  $\lim_{n \rightarrow \infty} x_n = x$ . For

a given  $\varepsilon > 0$ , there exists an integer  $n_0$  such that  $n \geq n_0$  implies  $\|x_n - x\| < \varepsilon$ . But since  $\|x_n\| - \|x\| \leq \|x_n - x\|$ , this implies  $\|x_n\| < \varepsilon + \|x\|$ ,  $n \geq n_0$ . It now follows that

$$\|x_n\| < \varepsilon + \|x\| + M,$$

where  $M = \max\{\|x_k\|: 1 \leq k \leq n_0\}$ .

Thus, the terms of a convergent sequence in a normed linear space, *a fortiori*, in a Hilbert space are bounded. The foregoing statement is true for a weakly convergent sequence.

**Theorem 2.12.6** *If  $H$  is a Hilbert space and  $x_n \xrightarrow{w} x$ , then there exists a positive constant  $M$  such that*

$$\|x_n\| \leq M.$$

We discuss preliminary results needed for the proof of Theorem 2.12.6.

**Definition 2.12.7** A real functional  $p(x)$  in  $H$  is said to be **convex** if for all  $x, y \in H$  and  $\alpha \in \mathbb{C}$ , the following hold:

$$p(x+y) \leq p(x) + p(y) \quad \text{and} \quad p(\alpha x) = |\alpha|p(x).$$

Observe that (i)  $p(0) = 0$  (ii)  $p(x - y) \geq |p(x) - p(y)|$  and  $p(x) \geq 0$ , where  $x, y \in H$ . Indeed,  $p(0) = p(0 \cdot x) = 0 \cdot p(x) = 0$ . Also,  $p(x - y) + p(y) \geq p(x)$  and hence  $p(x - y) \geq p(x) - p(y)$ . Since  $p(x - y) = |-1| p(y - x) \geq p(y) - p(x)$ , it follows that  $p(x - y) \geq |p(x) - p(y)|$ . On setting  $y = -x$ , we obtain  $p(2x) = 2p(x) \geq |p(x) - p(-x)| = 0$ .

That a lower semi-continuous convex functional in a Hilbert space is bounded is the content of the Lemma below. In conjunction with the observation above, it will further follow that it is uniformly continuous.

**Lemma 2.12.8** *Suppose  $p(x)$  is a convex functional in a Hilbert space  $H$  and assume that  $p(x)$  is lower semi-continuous. Then there exists  $M > 0$  such that*

$$p(x) \leq M\|x\| \quad \text{for all } x \in H.$$

*Proof* We first show that the functional  $p(x)$  is bounded in the ball  $S(0,1)$ . We assume the contrary. Then  $p(x)$  is unbounded in every ball, because every ball is obtained by dilation and/or translation of the ball  $S(0,1)$ . We choose a point  $x_1 \in S(0,1)$  such that  $p(x_1) > 1$ . The lower semi-continuity of the functional  $p(x)$  implies that there exists a ball  $S(x_1, \rho_1) \subseteq S(0,1)$  with radius  $\rho_1 < \frac{1}{2}$  in which  $p(x) > 1$ . By reducing the radius  $\rho_1$ , we may assume that  $\bar{S}(x_1, \rho_1) \subseteq S(0,1)$ . Since  $p(x)$  is unbounded in every ball, in a similar manner, we obtain a point  $x_2 \in S(x_1, \rho_1)$  and also a closed ball  $\bar{S}(x_2, \rho_2) \subseteq S(x_1, \rho_1)$  with radius  $\rho_2 < \frac{1}{2}\rho_1$ , in which  $p(x) > 2$ . Continuing the process, we obtain an infinite sequence of balls

$$S(0, 1) \supseteq \overline{S}(x_1, \rho_1) \supseteq \overline{S}(x_2, \rho_2) \supseteq \dots,$$

for which  $\rho_k < \frac{1}{2}\rho_{k-1}$  ( $k = 1, 2, \dots$  and  $\rho_0 = 1$ ) and also  $p(x) > n$  if  $x \in \overline{S}(x_n, \rho_n)$ . Observe that the sequence  $\{x_n\}_{n \geq 1}$  of the centres of the balls  $\overline{S}(x_n, \rho_n)$ ,  $n = 1, 2, \dots$  is Cauchy and since  $H$  is complete,  $\lim_{n \rightarrow \infty} x_n$  exists and equals  $x$ , say. Then  $x$  lies in the intersection of the closed balls, and hence  $p(x) > n$  for each  $n$ , which is a contradiction.

Let  $x \in H$  be arbitrary. Then  $x/2\|x\|$  is an element of  $H$  of norm  $\frac{1}{2}$  and is, therefore, in  $S(0,1)$ . Now,

$$p(x/2\|x\|) \leq M_1,$$

where  $M_1$  is an upper bound of  $p$  on  $S(0,1)$ , i.e.  $p(x) \leq 2M_1\|x\|$ . Take  $M = 2M_1$ .  $\square$

**Corollary 2.12.9** *Let  $p_k(x)$ ,  $k = 1, 2, \dots$  be a sequence of convex continuous functionals in  $H$ . If this sequence is bounded at each point  $x \in H$ , then the functional*

$$p(x) = \sup_k p_k(x)$$

*is also convex and bounded, and hence continuous.*

*Proof* Evidently,  $p(x)$  is a convex functional. On the other hand, for each  $x_0 \in H$  and each  $\varepsilon > 0$ , there exists  $N$  such that

$$p_N(x_0) > p(x_0) - \frac{1}{2}\varepsilon,$$

i.e.

$$p(x_0) - p_N(x_0) < \frac{1}{2}\varepsilon.$$

By continuity of the functional  $p_N(x)$ , there exists  $\delta > 0$  such that

$$|p_N(x) - p_N(x_0)| < \frac{1}{2}\varepsilon$$

for  $\|x - x_0\| < \delta$ . But if  $\|x - x_0\| < \delta$ , then

$$\begin{aligned} p(x) - p(x_0) &> \sup_k p_k(x) - p_N(x_0) - \frac{1}{2}\varepsilon \\ &\geq p_N(x) - p_N(x_0) - \frac{1}{2}\varepsilon > -\varepsilon. \end{aligned}$$

This implies that the functional  $p(x)$  is lower semi-continuous. By Lemma 2.12.8, it follows that  $p(x)$  is bounded. Continuity now follows from the observation preceding the lemma.  $\square$

Every weakly convergent sequence of vectors in a Hilbert space is bounded. This is an immediate consequence of the following theorem.

**Theorem 2.12.10** *Let  $\{\Phi_k\}_{k \geq 1}$  be a sequence of continuous linear functionals defined on the Hilbert space  $H$ . Suppose that the numerical sequence  $\{\Phi_k(x)\}_{k \geq 1}$  is bounded for each  $x \in H$ . Then the sequence  $\{\|\Phi_k\|\}_{k \geq 1}$  of norms of the functionals is bounded.*

*Proof* For  $x \in H$ , define

$$P_k(x) = |\Phi_k(x)|, \quad k = 1, 2, \dots$$

Then  $\{p_k\}_{k \geq 1}$  is convex and continuous. By Corollary 2.12.9, the convex functional

$$p(x) = \sup_k p_k(x)$$

is convex and bounded; i.e. there exists  $M > 0$  such that

$$\sup_{\|x\| \leq 1} p(x) \leq M.$$

Consequently,

$$\begin{aligned} \|\Phi_k\| &= \sup_{\|x\| \leq 1} |\Phi_k(x)| \\ &= \sup_{\|x\| \leq 1} p_k(x) \\ &\leq \sup_{\|x\| \leq 1} \sup_k p_k(x) \\ &= \sup_{\|x\| \leq 1} p(x) \leq M. \end{aligned}$$

This completes the proof.  $\square$

*Proof of Theorem 2.12.6* Let  $\{x_n\}_{n \geq 1}$  be a weakly convergent sequence. Each vector  $x_n$  determines a functional  $\Phi_n(x) = (x, x_n)$ . Since the sequence  $\{x_n\}_{n \geq 1}$  is weakly convergent, the numerical sequence  $\{\Phi_n(x)\}_{n \geq 1}$  converges for each  $x \in H$  and hence is bounded. Using Theorem 2.12.10, it follows that

$$\|\Phi_n\| \leq M, \quad n = 1, 2, \dots$$

As  $\|\Phi_n\| = \|x_n\|$ ,  $n = 1, 2, \dots$  [Example 2.10.24(iv)], the result follows.

**Definition 2.12.11** A sequence of vectors  $\{x_n\}_{n \geq 1}$  in an inner product space is said to be **weakly Cauchy** if, for each  $y \in H$ ,

$$\lim_{m,n} (x_m - x_n, y) = 0.$$

An inner product space is said to be **weakly complete** if every weakly Cauchy sequence converges to a weak limit in  $H$ .

**Corollary 2.12.12** *Let  $H$  be a Hilbert space. Then  $H$  is weakly complete.*

*Proof* Let  $\{x_n\}_{n \geq 1}$  be a Cauchy sequence in the sense of weak convergence, that is, for each  $y \in H$ ,

$$\lim_{m,n} (x_m - x_n, y) = 0.$$

It follows that the sequence  $\{(x_n, y)\}_{n \geq 1}$  of scalars converges for each  $y$  in  $H$ . By Theorem 2.12.10, the sequence  $\{x_n\}_{n \geq 1}$  is bounded:

$$\|x_n\| \leq M, \quad n = 1, 2, \dots$$

Therefore, the limit

$$\lim_{n \rightarrow \infty} (x, x_n)$$

defines a linear functional  $\Phi(x)$  with norm less than or equal to  $M$ . By the Riesz Representation Theorem 2.10.25,  $\Phi(x) = (x, z)$ , where  $z$  is a unique element of the Hilbert space  $H$ . This element is the weak limit of the sequence  $\{x_n\}_{n \geq 1}$ .  $\square$

We give below two applications of Corollary 2.12.9.

**Theorem 2.12.13** (F. Riesz) *If a functional  $\Phi$  is defined everywhere on  $L^2[a, b]$  by the formula*

$$\Phi(x) = \int_a^b x(t)y(t)dt, \quad x \in L^2[a, b],$$

*where  $y$  is a fixed measurable function defined on  $[a, b]$ , then  $\Phi$  is a bounded linear functional on  $L^2[a, b]$ , so that  $y \in L^2[a, b]$ .*

*Proof* Clearly,  $\Phi$  is a linear functional on  $L^2[a, b]$ . Set

$$E_n = \{t : t \in [a, b] \cap [-n, n] \text{ and } |y(t)| \leq n\}$$

and

$$p_n(x) = \int_{E_n} |x(t)y(t)| dt, \quad x \in L^2[a, b].$$

Then  $\{p_n\}_{n \geq 1}$  is a sequence of convex functionals: indeed, for  $x, z \in L^2[a, b]$  and  $\alpha \in \mathbb{C}$ ,

$$\begin{aligned} p_n(x+z) &= \int_{E_n} |[x(t)+z(t)]y(t)| dt \\ &\leq \int_{E_n} |x(t)y(t)| dt + \int_{E_n} |z(t)y(t)| dt \\ &= p_n(x) + p_n(z) \end{aligned}$$

and

$$p_n(\alpha x) = \int_{E_n} |\alpha x(t)y(t)| dt = |\alpha| \int_{E_n} |x(t)y(t)| dt = |\alpha| p_n(x).$$

Moreover,

$$\begin{aligned} p_n(x) &\leq n \int_{E_n} |x(t)| dt \leq n \left( \int_{E_n} |x(t)|^2 dt \right)^{\frac{1}{2}} \left( \int_{E_n} dt \right)^{\frac{1}{2}} \\ &= n \mu(E_n)^{\frac{1}{2}} \|x\|_2, \end{aligned}$$

using the Cauchy–Schwarz Inequality, where  $\mu$  denotes the usual Lebesgue measure.

Thus, for  $n = 1, 2, \dots$ ,  $p_n$  is a continuous convex functional on  $L^2[a, b]$ . The equality

$$p(x) = \lim_n p_n(x) = \lim_n \int_{E_n} |x(t)y(t)| dt = \int_a^b |x(t)y(t)| dt,$$

using the Monotone Convergence Theorem 1.3.6 shows that  $p(x)$  is finite for any  $x$  in  $L^2[a, b]$ . By Corollary 2.12.9, the functional  $p(x)$  is bounded; i.e. there exists  $M > 0$  such that

$$p(x) \leq M \|x\|, \quad x \in L^2[a, b].$$

Thus

$$|\Phi(x)| \leq p(x) \leq M\|x\|, \quad x \in L^2[a, b],$$

i.e.  $\Phi$  is a bounded linear functional on  $L^2[a, b]$ ; so  $y \in L^2[a, b]$  and  $\|y\|_2 = \|\Phi\|$ , using the definition of  $\Phi$  and the Riesz Representation Theorem.  $\square$

**Theorem 2.12.14** (Landau) *If  $\Phi$  is a functional defined everywhere in  $\ell^2$  by means of the formula*

$$\Phi(x) = \sum_{k=1}^{\infty} a_k x_k, \quad x = \{x_k\}_{k \geq 1} \in \ell^2,$$

where  $\{a_k\}_{k \geq 1}$  is some fixed sequence, then  $\sum_{k=1}^{\infty} |a_k|^2 < \infty$ .

*Proof* Define  $p_n(x) = \sum_{k=1}^n |a_k x_k|$ ,  $x = \{x_k\}_{k \geq 1} \in \ell^2$ . Check that  $p_n$ ,  $n = 1, 2, \dots$ , is a continuous convex functional. Then the equality

$$p(x) = \lim_n p_n(x) = \lim_n \sum_{k=1}^n |a_k x_k| = \sum_{k=1}^{\infty} |a_k x_k|$$

implies that  $p(x)$  is finite for any  $x \in \ell^2$ . So, by Corollary 2.12.9, the functional  $p(x)$  is continuous; i.e. there exists  $M > 0$  such that  $p(x) \leq M\|x\|$ ,  $x \in \ell^2$ . Consequently,

$$|\Phi(x)| \leq \sum_{k=1}^{\infty} |a_k x_k| = p(x) \leq M\|x\|.$$

So,  $\Phi$  is a bounded linear functional on  $\ell^2$ . The form of  $\Phi$  and the Riesz Representation Theorem imply that  $\sum_{k=1}^{\infty} |a_k|^2 < \infty$ .  $\square$

*Remark 2.12.15* Landau's Theorem may also be stated as follows: if  $\sum_{k=1}^{\infty} a_k x_k$  converges for every  $\{x_k\}_{k \geq 1}$  in  $\ell^2$ , then  $\sum_{k=1}^{\infty} |a_k|^2 < \infty$

### Problem Set 2.12

2.12.P1. Show that for a sequence  $\{x_n\}_{n \geq 1}$  in an inner product space  $X$  and  $x \in X$ , the conditions

$$(i) \|x_n\| \rightarrow \|x\| \quad \text{and} \quad (ii) (x_n, x) \rightarrow (x, x),$$

imply  $x_n \rightarrow x$  in  $X$ .

- 2.12.P2. (**Banach–Saks**) Let  $\{x_n\}_{n \geq 1}$  be a sequence in a Hilbert space converging weakly to  $x \in H$ . Prove that there exists a subsequence  $\{x_{n_k}\}_{k \geq 1}$  such that the sequence  $\{y_k\}_{k \geq 1}$  defined by

$$y_k = \frac{1}{k}(x_{n_1} + x_{n_2} + \cdots + x_{n_k})$$

converges strongly to  $x$ .

- 2.12.P3. (a) (**Mazur's Theorem**) Let  $\{x_n\}_{n \geq 1}$  be a weakly convergent sequence in a Hilbert space  $H$  and let  $x$  be its weak limit. Prove that  $x$  lies in the closed convex hull of the range  $\{x_n; n \geq 1\}$  of the sequence.
- (b) Let  $C$  be a convex subset of a Hilbert space  $H$ . Prove that  $C$  is closed if, and only if, it contains the weak limit of every sequence of points in it.
- 2.12.P4. (a) Let  $H$  be a separable Hilbert space and let  $\{e_n\}_{n \geq 1}$  be an orthonormal basis for  $H$ . Let  $B = \{x \in H: \|x\| \leq 1\}$ . For  $x, y \in H$ , let

$$d(x, y) = \sum_{n=1}^{\infty} 2^{-n} |(x - y, e_n)| \quad (2.85)$$

Show that  $d$  is a metric on  $B$ .

- (b) Show that the topology generated by  $d$  is the same as the one given by the weak topology, i.e.  $d(x_k, x) \rightarrow 0$  if, and only if,  $x_k \xrightarrow{w} x$ .
- (c) Show that the metric space  $(B, d)$  is compact.

## 2.13 Applications

### Müntz's Theorem

Weierstrass's Theorem for  $C[0, 1]$  says, in effect, that all linear combinations of the functions

$$1, x, x^2, \dots, x^n, \dots \quad (2.86)$$

are dense in  $C[0, 1]$ . Instead of working with all positive powers of  $x$ , let us permit gaps to occur, and consider the infinite set of functions

$$1, x^{n_1}, x^{n_2}, \dots, x^{n_k}, \dots, \quad (2.87)$$

where  $n_k$  are positive integers satisfying  $n_1 < n_2 < \dots < n_k \dots$ . The result we shall prove is called Müntz's Theorem and asserts that the linear combinations of the functions (2.87) are dense in  $C[0, 1]$  and hence in  $L^2[0, 1]$  if, and only if, the series  $\sum_{k=1}^{\infty} \frac{1}{n_k}$  diverges. The following will be needed in Sect. 2.13.

**Definition** Let  $x_1, x_2, \dots, x_n$  be any vectors in an inner product space  $X$ . Then the  $n \times n$  matrix  $G(x_1, x_2, \dots, x_n)$  whose  $(i, j)$ th entry is  $(x_i, x_j)$ , where  $(\cdot, \cdot)$  is the inner product in  $X$ , is called the **Gram matrix** of the given finite sequence of vectors.

*Its determinant is called their **Gram determinant**.*

**Proposition** The Gram matrix  $G(x_1, x_2, \dots, x_n)$  is nonsingular if, and only if, the vectors  $x_1, x_2, \dots, x_n$  are linearly independent.

*Proof* Observe that for the given  $G$ , and any  $n$ -tuple of scalars  $x = (\xi_1, \xi_2, \dots, \xi_n)$ , we have

$$\begin{aligned} xG &= [\xi_1, \xi_2, \dots, \xi_n] \begin{bmatrix} (x_1, x_1) & (x_1, x_2) & \cdots & (x_1, x_n) \\ \vdots & \vdots & \vdots & \vdots \\ (x_n, x_1) & (x_n, x_2) & \cdots & (x_n, x_n) \end{bmatrix} \\ &= \left[ \sum_{i=1}^n \xi_i (x_i, x_1) \quad \sum_{i=1}^n \xi_i (x_i, x_2) \quad \cdots \quad \sum_{i=1}^n \xi_i (x_i, x_n) \right]. \end{aligned}$$

So,

$$\begin{aligned} xGx^* &= \left[ \sum_{i=1}^n \xi_i (x_i, x_1) \quad \sum_{i=1}^n \xi_i (x_i, x_2) \quad \cdots \quad \sum_{i=1}^n \xi_i (x_i, x_n) \right] \begin{bmatrix} \bar{\xi}_1 \\ \vdots \\ \bar{\xi}_n \end{bmatrix} \\ &= \sum_{i,j=1}^n (x_i, x_j) \xi_i \bar{\xi}_j = \left( \sum_{i=1}^n \xi_i x_i, \sum_{i=1}^n \xi_i x_i \right) = \left\| \sum_{i=1}^n \xi_i x_i \right\|^2 \end{aligned}$$

From the above equality, the result follows.  $\square$

**Corollary** A necessary and sufficient condition for the vectors  $x_1, x_2, \dots, x_n$  to be linearly dependent is that

$$\det G(x_1, x_2, \dots, x_n) = 0.$$

Let  $M$  be the closed subspace generated by  $x_1, x_2, \dots, x_n$ . Then  $H$  can be written as  $M \oplus M^\perp$ . If  $y \in H$ , then  $y = z + w$ , where  $z \in M$  and  $w \in M^\perp$ , so that  $y - z \in M^\perp$  [see Remark 2.10.12(i)]. The minimum distance  $\delta$  from  $y$  to the subspace  $M$  is  $\delta = \|y - z\|$ , where

$$z = \sum_{i=1}^n a_i x_i$$

[Theorem 2.10.10]. We wish to calculate the coefficients  $a_i$ ,  $i = 1, 2, \dots, n$  and the minimal distance  $\delta$ .

Since  $y - z \perp x_j$ ,  $j = 1, 2, \dots, n$ , we obtain a system of equations

$$\left( y - \sum_{i=1}^n a_i x_i, x_j \right) = 0, \quad j = 1, 2, \dots, n,$$

which, when written in full, have the form

$$\left. \begin{aligned} a_1(x_1, x_1) + a_2(x_2, x_1) + \dots + a_n(x_n, x_1) &= (y, x_1) \\ a_1(x_1, x_2) + a_2(x_2, x_2) + \dots + a_n(x_n, x_2) &= (y, x_2) \\ &\dots \\ a_1(x_1, x_n) + a_2(x_2, x_n) + \dots + a_n(x_n, x_n) &= (y, x_n) \end{aligned} \right\} \quad (2.88)$$

and represent a system of equations in the unknowns  $a_i$ ,  $i = 1, 2, \dots, n$ . The matrix of its coefficients is precisely the transpose of the Gram matrix  $G(x_1, x_2, \dots, x_n)$ . Since the vectors  $x_1, x_2, \dots, x_n$  are linearly independent, the matrix is nonsingular by Proposition 4.2.2 and the system has one and only one solution. Moreover, by Cramer's Rule, the unique solution is given by

$$a_i = \det G^{(i)} / \det G, \quad j = 1, 2, \dots, n,$$

where  $G^{(i)}$  is obtained from  $G$  by replacing its  $i$ th column by the column of constants  $(y, x_i)$ .

Now,

$$\begin{aligned} \delta^2 &= \|y - z\|^2 = (y - z, y - z) \\ &= (y, y - z) \\ &= \|y\|^2 - \left( y, \sum_{i=1}^n a_i x_i \right), \end{aligned}$$

so that

$$\left( \sum_{i=1}^n a_i x_i, y \right) = \|y\|^2 - \delta^2. \quad (2.89)$$

We combine Eq. (2.89) with the system of Eq. (2.88) and write them in the form

$$\left. \begin{aligned} a_1(x_1, x_1) + a_2(x_2, x_1) + \cdots a_n(x_n, x_1) - (y, x_1) &= 0 \\ a_1(x_1, x_2) + a_2(x_2, x_2) + \cdots a_n(x_n, x_2) - (y, x_2) &= 0 \\ &\vdots \\ a_1(x_1, x_n) + a_2(x_2, x_n) + \cdots a_n(x_n, x_n) - (y, x_n) &= 0 \\ a_1(x_1, y) + a_2(x_2, y) + \cdots a_n(x_n, y) + \delta^2 - (y, y) &= 0 \end{aligned} \right\}. \quad (2.90)$$

If we introduce a dummy value  $a_{n+1} = 1$  as a coefficient of the elements of the last column, the (2.90) becomes a system of  $n + 1$  homogeneous linear equations in the  $n + 1$  variables  $a_1, a_2, \dots, a_n, a_{n+1} (= 1)$ . The system (2.90) will possess a nontrivial solution if the determinant of the system vanishes, i.e.

$$\det \begin{bmatrix} (x_1, x_1) & (x_2, x_1) & \cdots & (x_n, x_1) & -(y, x_1) \\ (x_1, x_1) & (x_2, x_1) & \cdots & (x_n, x_2) & -(y, x_2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ (x_1, y) & (x_2, y) & \cdots & (x_n, y) & \delta^2 - (y, y) \end{bmatrix} = 0.$$

This gives

$$\delta^2 = \frac{\det G(x_1, x_2, \dots, x_n, y)}{\det G(x_1, x_2, \dots, x_n)}.$$

Apart from the above observations, the following lemmas will be needed in the proof of Müntz's Theorem.

**Lemma** Let  $\lambda_1, \lambda_2, \dots, \lambda_n$  be positive real numbers and  $A$  be the matrix whose  $(i, j)$ th entry is  $a_{ij} = \frac{1}{\lambda_i + \lambda_j}$ . Then

$$\det A = 2^{-n} \prod_{j=1}^n \frac{1}{\lambda_j} \prod_{1 \leq j < k \leq n} \left( \frac{\lambda_j - \lambda_k}{\lambda_j + \lambda_k} \right)^2.$$

*Proof* If  $A$  is a  $1 \times 1$  matrix, then  $\det A = \frac{1}{2\lambda_1} = 2^{-1} \prod_{j=1}^1 \frac{1}{\lambda_j}$ . Thus, the assertion is true for  $n = 1$ . Assume that the assertion is true for  $m$ ; i.e. if  $A$  is an  $m \times m$  matrix, then

$$\det A = 2^{-m} \prod_{j=1}^m \frac{1}{\lambda_j} \prod_{1 \leq j < k \leq m} \left( \frac{\lambda_j - \lambda_k}{\lambda_j + \lambda_k} \right)^2.$$

Consider the  $(m + 1) \times (m + 1)$  matrix whose  $(i, j)$ th entry is  $a_{ij} = \frac{1}{\lambda_i + \lambda_j}$ . Its determinant, when written in full, takes the form

$$\begin{vmatrix} \frac{1}{\lambda_1 + \lambda_1} & \frac{1}{\lambda_1 + \lambda_2} & \cdots & \frac{1}{\lambda_1 + \lambda_m} & \frac{1}{\lambda_1 + \lambda_{m+1}} \\ \frac{1}{\lambda_2 + \lambda_1} & \frac{1}{\lambda_2 + \lambda_2} & \cdots & \frac{1}{\lambda_2 + \lambda_m} & \frac{1}{\lambda_2 + \lambda_{m+1}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{\lambda_{m+1} + \lambda_1} & \frac{1}{\lambda_{m+1} + \lambda_2} & \cdots & \frac{1}{\lambda_{m+1} + \lambda_m} & \frac{1}{\lambda_{m+1} + \lambda_{m+1}} \end{vmatrix}.$$

By subtracting the last row from each of the others, removing common factors, subtracting the last column from each of the others and again removing the common factors, we obtain

$$2^{-1} \frac{1}{\lambda_{m+1}} \frac{\prod_{i=1}^m (\lambda_{m+1} - \lambda_i)^2}{\prod_{i=1}^m (\lambda_{m+1} + \lambda_i)^2} \begin{vmatrix} \frac{1}{\lambda_1 + \lambda_1} & \frac{1}{\lambda_1 + \lambda_2} & \cdots & \frac{1}{\lambda_1 + \lambda_m} & 1 \\ \frac{1}{\lambda_2 + \lambda_1} & \frac{1}{\lambda_2 + \lambda_2} & \cdots & \frac{1}{\lambda_2 + \lambda_m} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{\lambda_{m+1} + \lambda_1} & \frac{1}{\lambda_{m+1} + \lambda_2} & \cdots & \frac{1}{\lambda_{m+1} + \lambda_m} & 1 \\ 0 & 0 & \cdots & 0 & 1 \end{vmatrix}.$$

On expanding the determinant by the last row, we have

$$\det A = 2^{-m-1} \prod_{j=1}^{m+1} \frac{1}{\lambda_j} \prod_{1 \leq j < k \leq m+1} \left( \frac{\lambda_j - \lambda_k}{\lambda_j + \lambda_k} \right)^2.$$

By induction, the result follows.  $\square$

**Lemma** Let  $1, t^{n_1}, t^{n_2}, \dots, 1 \leq n_1 < n_2 < \dots$  be a set of functions defined on  $[0, 1]$ . The sequence  $\{t^{n_k}\}_{k \geq 1}$  is total (finite linear combinations are dense) in  $C[0, 1]$ , if, and only if, it is complete in  $L^2[0, 1]$ .

*Proof* Let  $x \in C[0, 1]$ . The inequality

$$\left[ \int_0^1 \left| x(t) - \sum_{i=1}^k a_i t^{n_i} \right|^2 dt \right]^{\frac{1}{2}} \leq \max_{0 \leq t \leq 1} \left| x(t) - \sum_{i=1}^k a_i t^{n_i} \right| \quad (2.91)$$

shows that the sequence is complete in  $L^2[0, 1]$  if it is total in  $C[0, 1]$ .

Conversely, suppose that the sequence  $\{t^{n_k}\}_{k \geq 1}$  is complete in  $L^2[0, 1]$ . In order to show that the finite linear combinations constitute a dense subset of  $C[0, 1]$ , it is enough to show that the inequality (2.91) in the reverse direction holds for the functions  $t^m$ ,  $m = 1, 2, \dots$ . Now

$$\begin{aligned}
\left| t^m - \sum_{i=1}^k a_i t^{n_i} \right| &= m \left| \int_0^t \left( t^{m-1} - \sum_{i=1}^k b_i t^{n_i-1} \right) dt \right| \\
&\leq m \int_0^1 \left| t^{m-1} - \sum_{i=1}^k b_i t^{n_i-1} \right| dt \\
&\leq m \left( \int_0^1 \left| t^{m-1} - \sum_{i=1}^k b_i t^{n_i-1} \right| dt \right)^{\frac{1}{2}},
\end{aligned} \tag{2.92}$$

using the Cauchy–Schwarz Inequality. The above inequality proves the assertion.  $\square$

### Remarks

- (i) The function 1 must be added in the case of  $C[0, 1]$  but is redundant in  $L^2[0, 1]$ . Indeed, if the function 1 is missing from  $\{t^{n_k}\}_{k \geq 1}$ , then the polynomial  $\sum_{i=1}^k a_i t^{n_i}$  is itself zero at  $t = 0$  and cannot, therefore, approximate the continuous function  $x(t)$  for which  $x(0) \neq 0$ .
- (ii) Since  $(x^p, x^q) = \int_0^1 t^{p+q} dt = \frac{1}{p+q+1}$ , it follows that

$$\begin{aligned}
\det G(t^{n_1}, t^{n_2}, \dots, t^{n_k}) &= \begin{vmatrix} \frac{1}{n_1+n_1+1} & \frac{1}{n_1+n_2+1} & \cdots & \frac{1}{n_1+n_k+1} \\ \frac{1}{n_2+n_1+1} & \frac{1}{n_2+n_2+1} & \cdots & \frac{1}{n_2+n_k+1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n_k+n_2+1} & \frac{1}{n_k+n_2+1} & \cdots & \frac{1}{n_k+n_k+1} \end{vmatrix} \\
&= \frac{\prod_{i>j} (n_i - n_j)^2}{\prod_{i,j} (n_i + n_j + 1)}
\end{aligned}$$

and analogously,

$$\det G(t^m, t^{n_1}, t^{n_2}, \dots, t^{n_k}) = \frac{\prod_{i>j} (n_i - n_j)^2}{\prod_{i,j} (n_i + n_j + 1)} \cdot \frac{\prod_{i=1}^k (m - n_i)^2}{\prod_{i=1}^k (m + n_i + 1)^2} \cdot \frac{1}{2m+1}.$$

From this, it follows that

$$\frac{\det G(t^m, t^{n_1}, t^{n_2}, \dots, t^{n_k})}{\det G(t^{n_1}, t^{n_2}, \dots, t^{n_k})} = \frac{1}{2m+1} \prod_{i=1}^k \left( \frac{m - n_i}{m + n_i + 1} \right)^2.$$

- (iii) The series  $\sum_{i=1}^{\infty} \ln(1+a_i)$  and the series  $\sum_{i=1}^{\infty} a_i$  converge or diverge simultaneously. This is because

$$\lim_{x \rightarrow 0} \frac{\ln(1+x)}{x} = \lim_{x \rightarrow 0} \frac{1}{1+x} = 1$$

and so, for any  $\varepsilon > 0$ ,

$$(1 - \varepsilon)a_i < \ln(1+a_i) < (1 + \varepsilon)a_i.$$

**(Müntz's Theorem)** *A necessary and sufficient condition for the set*

$$t^{n_1}, t^{n_2}, \dots, \quad 1 \leq n_1 < n_2 < \dots$$

*to be complete in  $L^2[0, 1]$  is that*

$$\sum_{i=1}^{\infty} \frac{1}{n_i} = \infty.$$

*Proof* If one of the exponents  $n_i$  coincides with  $m$ , then for  $k \geq i$ ,

$$\det G(t^m, t^{n_1}, t^{n_2}, \dots, t^{n_k}) = 0$$

and hence, the minimal distance is zero. Thus, completeness holds if, and only if, for each  $m \geq 1$  with  $m \neq n_i$ ,  $i = 1, 2, \dots$ , the minimal distance

$$\delta_k^2 = \frac{\det G(t^m, t^{n_1}, t^{n_2}, \dots, t^{n_k})}{\det G(t^{n_1}, t^{n_2}, \dots, t^{n_k})} \rightarrow 0 \quad \text{as } k \rightarrow \infty. \quad (2.93)$$

Now,

$$\frac{\det G(t^m, t^{n_1}, t^{n_2}, \dots, t^{n_k})}{\det G(t^{n_1}, t^{n_2}, \dots, t^{n_k})} = \frac{1}{2m+1} \prod_{i=1}^k \left( \frac{m - n_i}{m + n_i + 1} \right)^2. \quad (2.94)$$

In view of (2.94), the condition (2.93) becomes

$$\lim_{k \rightarrow \infty} \left[ \sum_{i=1}^k \left( \ln \left( 1 - \frac{m}{n_i} \right) - \ln \left( 1 + \frac{m+1}{n_i} \right) \right) \right] = -\infty. \quad (2.95)$$

If the series  $\sum_{i=1}^{\infty} \frac{1}{n_i}$  diverges, then by (iii) of Remarks 4.2.6

$$\sum_{i=1}^{\infty} \ln \left( 1 - \frac{m}{n_i} \right) = -\infty, \quad \sum_{i=1}^{\infty} \ln \left( 1 + \frac{m+1}{n_i} \right) = +\infty \quad (2.96)$$

and therefore, (2.95) is satisfied and hence so is (2.93). If, however, the series  $\sum_{i=1}^{\infty} \frac{1}{n_i}$  converges, then the series (2.96) also converges, so that (2.95) is not satisfied and hence (2.93) does not hold.  $\square$

### Radon–Nikodým Theorem

**Definition** Let  $(X, \Sigma)$ , where  $X$  is a nonempty set and  $\Sigma$  is a  $\sigma$ -algebra of subsets of  $X$ , be a measurable space, and let  $\nu, \mu$  be finite nonnegative measures on  $(X, \Sigma)$ . The measure  $\nu$  is said to be **absolutely continuous** with respect to  $\mu$ , in symbols,  $\nu \ll \mu$ , if  $\nu(E) = 0$  for every  $E \in \Sigma$  for which  $\mu(E) = 0$ .

For  $h \in L^1(X, \Sigma, \mu)$ , the integral

$$\nu(E) = \int_E h \, d\mu, \quad E \in \Sigma$$

defines a measure on  $\Sigma$  which is clearly absolutely continuous with respect to  $\mu$ . The point of the Radon–Nikodým Theorem is the converse: every  $\nu \ll \mu$  is obtained in this way.

von Neumann showed how to derive this from the Riesz Representation Theorem for linear functionals on a Hilbert space.

**(Radon–Nikodým Theorem)** *Let  $\nu$  and  $\mu$  be finite nonnegative measures on  $(X, \Sigma)$ . If  $\nu \ll \mu$ , then there exists a unique nonnegative measurable function  $h$  such that*

$$\nu(E) = \int_E h \, d\mu, \quad E \in \Sigma.$$

*In particular,  $h \in L^1(X, \Sigma, \mu)$ .*

*Proof* For any  $E \in \Sigma$ , put  $\phi(E) = \nu(E) + \mu(E)$ . Since  $\nu$  and  $\mu$  are finite nonnegative measures, so is  $\phi$ . Moreover,

$$\int_X x \, d\phi = \int_X x \, d\nu + \int_X x \, d\mu \quad (2.97)$$

holds for  $x = \chi_E$ ,  $E \in \Sigma$ . Hence, (2.97) holds for simple functions and consequently for any nonnegative measurable function  $x$ .

Let  $H$  be the real Hilbert space  $L^2(X, \Sigma, \phi)$  with the norm  $\|x\|^2 = \int_X |x|^2 \, d\phi$ . For  $x \in H$ , the Cauchy–Schwarz Inequality gives

$$\left| \int_X x \, dv \right| \leq \int_X |x| \, dv \leq \int_X |x| \, d\varphi \leq \left( \int_X |x|^2 \, d\varphi \right)^{\frac{1}{2}} (\varphi(X))^{\frac{1}{2}} < \infty$$

since  $\varphi(X) < \infty$ . Thus, the mapping

$$L : x \rightarrow \int_X x \, dv$$

is seen to be defined and finite on  $H$ . It is clear that  $L(\alpha x + \beta y) = \alpha L(x) + \beta L(y)$  for all  $\alpha, \beta$  scalars and  $x, y \in L^2(X, \Sigma, \varphi) = H$ . Thus,  $L$  is a bounded linear functional on  $H$  and so, by Theorem 2.10.25, there is a function  $y \in H$  such that

$$\int_X x \, dv = \int_X xy \, d\varphi = \int_X xy \, dv + \int_X xy \, d\mu,$$

where we have used (2.97) in the last equality. It is easy to discern that  $y$  is non-negative a.e. with respect to  $\varphi$  and hence with respect to  $\mu$  and  $\nu$  as well. This may be written as

$$\int_X x(1-y) \, dv = \int_X xy \, d\mu. \quad (2.98)$$

Let  $E = \{s \in X : y(s) \geq 1\}$ . Since  $\chi_E \in L^2(X, \Sigma, \varphi)$ , we apply (2.98) to  $x = \chi_E$  to obtain

$$0 \leq \mu(E) = \int_X \chi_E \, d\mu \leq \int_X \chi_E y \, d\mu = \int_X \chi_E (1-y) \, dv \leq 0.$$

Thus, we have  $\mu(E) = 0$  and since  $\nu \ll \mu$ ,  $\nu(E) = 0$ .

Let  $z = y\chi_{E^c}$ . Then  $z(s) \in [0, 1)$  and  $z = y$  a.e. with respect to both  $\nu$  and  $\mu$ . The equality (2.98) then becomes

$$\int_X x(1-z) \, dv = \int_X xz \, d\mu. \quad (2.99)$$

Consider any bounded, nonnegative, measurable function  $x$ . Let  $z$  be as above. Since both  $x$  and  $z$  are bounded and  $\varphi$  is a finite measure, the function  $(1 + z + z^2 + \cdots + z^{n-1})x$  is in  $L^2(X, \Sigma, \varphi)$  for every positive integer  $n$  and hence by (2.99)

$$\int_X (1 + z + z^2 + \cdots + z^{n-1})x(1 - z)d\nu = \int_X (1 + z + z^2 + \cdots + z^{n-1})xz d\mu$$

holds. In view of the fact that  $z(s) \neq 1$  for any  $s$ , the above equality can be written as

$$\int_X (1 - z^n)x d\nu = \int_X \frac{z(1 - z^n)}{1 - z}x d\mu.$$

Since  $0 \leq z(s) < 1$  for all  $s \in X$ , the sequences  $(1 - z^n)x$  and  $\frac{z(1 - z^n)}{1 - z}x$  increase to  $x$  and  $\frac{z}{1 - z}x$ , respectively, as  $n \rightarrow \infty$ . By the Monotone Convergence Theorem 1.3.6, we obtain

$$\int_X x d\nu = \int_X \frac{z}{(1 - z)}x d\mu.$$

Now define  $h = \frac{z}{1 - z}$ ; then we have

$$\int_X x d\nu = \int_X hx d\mu.$$

In particular, for  $E \in \mathfrak{M}$  and  $x = \chi_E$ , we obtain

$$\nu(E) = \int_E h d\mu.$$

The uniqueness is obvious. □

### Remarks

- (i) The construction of  $h$  shows that  $h \geq 0$ .
- (ii) The Radon–Nikodým Theorem is valid if  $\nu$  and  $\mu$  are  $\sigma$ -finite measures. For details, the reader may consult [26].

### Bergman Kernel and Conformal Mappings

Let  $\Omega$  be a bounded domain in the  $z = x + iy$  plane, whose boundary consists of a finite number of smooth simple closed curves. The class of all holomorphic functions in  $\Omega$  for which the integral  $\iint_{\Omega} |f(z)|^2 dx dy < \infty$  is denoted by  $A(\Omega)$ . The integral is understood as the limit of Riemann integrals

$$\lim_n \iint_{K_n} |f(z)|^2 dx dy,$$

where  $\{K_n\}_{n \geq 1}$  is a nondecreasing sequence of compact subsets of  $\Omega$  whose union is  $\Omega$ . It has been proved [see 2.6.2, 2.6.3, 2.6.4, 2.6.5] that  $A(\Omega)$  is a Hilbert space.

Consider the linear functional  $L(f) = f(\zeta)$ , where  $\zeta \in \Omega$  is fixed and  $f \in A(\Omega)$ . Observe that  $|L(f)| = |f(\zeta)| \leq \frac{\|f\|}{\sqrt{\pi d_\zeta}}$ , where  $d_\zeta = \text{dist}(\zeta, \partial\Omega)$  and  $\partial\Omega$  denotes the boundary of  $\Omega$  [see Proposition 2.6.3]. It follows on using Theorem 2.10.25 that there exists a uniquely determined  $u_\zeta \in A(\Omega)$  such that

$$f(\zeta) = (f, u_\zeta), \quad f \in A(\Omega). \quad (2.100)$$

The traditional notation is  $u_\zeta(z) = K(z, \zeta)$  and  $K$  is called the **Bergman kernel** of  $\Omega$ . For each  $\zeta \in \Omega$ , the function has the reproducing property

$$f(\zeta) = (f, K(\cdot, \zeta)) = \iint_{\Omega} f(z) \overline{K(z, \zeta)} dx dy, \quad f \in A(\Omega). \quad (2.101)$$

The following two assertions are immediate from (2.101).

(a) If one substitutes  $f = K(\cdot, \zeta)$  in (2.101), one finds that

$$\begin{aligned} \|K(\cdot, \zeta)\|^2 &= \iint_{\Omega} K(z, \zeta) \overline{K(z, \zeta)} dx dy \\ &= K(\zeta, \zeta), \quad \zeta \in \Omega. \end{aligned}$$

(b) For  $z_1, z_2 \in \Omega$ , the relation  $K(z_1, z_2) = \overline{K(z_2, z_1)}$  holds. To see this, we let  $f = K(\cdot, z_2)$  and  $\zeta = z_1$  in (2.101) and we obtain

$$\begin{aligned} K(z_1, z_2) &= \iint_{\Omega} K(z, z_2) \overline{K(z, z_1)} dx dy \\ &= \iint_{\Omega} \overline{\overline{K(z, z_2)} K(z, z_1)} dx dy \\ &= \overline{K(z_2, z_1)} \end{aligned}$$

The relation between the kernel function and a certain minimum problem in  $A(\Omega)$  is also important. Suppose  $\zeta \in \Omega$  is fixed, and write

$$M = \{f \in A(\Omega) : f(\zeta) = 1\}.$$

*There is exactly one solution  $f_0 \in M$  such that  $\min_{f \in M} \|f\| = \|f_0\|$ . Moreover, the function  $f_0$  is connected with the Bergman kernel function as follows:*

$$f_0(z) = \frac{K(z, \zeta)}{K(\zeta, \zeta)} \quad \text{and} \quad K(z, \zeta) = \frac{f_0(z)}{\|f_0\|^2}.$$

*Proof* Since  $A(\Omega)$  is a Hilbert space and  $M \subseteq A(\Omega)$  is its closed subspace, the first assertion follows on using Corollary 2.10.7.

For each  $f \in A(\Omega)$ , we have  $f(\zeta) = (f, K(\cdot, \zeta))$ . For  $f \in M$ , on using the Cauchy–Schwarz inequality, we have

$$1 = (f, K(\cdot, \zeta)) \leq \|f\| \|K(\cdot, \zeta)\| = \|f\| \cdot \sqrt{K(\zeta, \zeta)}. \quad (2.102)$$

Equality in the above inequality occurs provided

$$f = f_0 = CK(\cdot, \zeta), \quad \text{where } C \text{ is a constant.} \quad (2.103)$$

Since  $1 = f_0(\zeta) = CK(\zeta, \zeta)$  (therefore  $C = \frac{1}{K(\zeta, \zeta)}$ ), it follows that

$$f_0(z) = \frac{K(z, \zeta)}{K(\zeta, \zeta)}.$$

This implies

$$K(z, \zeta) = f_0(z)K(\zeta, \zeta).$$

Also,

$$K(z, \zeta) = \frac{f_0(z)}{\|f_0\|^2},$$

since  $\|f\|^2 = \|f_0\|^2 = \frac{1}{K(\zeta, \zeta)}$ , using (2.102) and (2.103).  $\square$

Recall that the **Riemann Mapping Theorem** asserts: if  $\Omega$  is a simply connected domain having more than one boundary point, then there exists a holomorphic function in  $\Omega$  which maps  $\Omega$  bijectively onto  $D = \{z : |z| < 1\}$ . If  $\zeta$  is fixed, then the mapping function  $f(z) = f(z, \zeta)$  for which  $f(\zeta) = 0$  and  $f'(\zeta) > 0$  is unique.

*The mapping function  $f$  and the Bergman kernel  $K$  of  $\Omega$  are related as follows:*

$$f'(z) = \sqrt{\frac{\pi}{K(\zeta, \zeta)}} k(z, \zeta) \quad \text{and} \quad k(z, \zeta) = \frac{1}{\pi} f'(z) f'(\zeta), \quad z \in \Omega$$

*Proof* Let  $\Omega_r$  denote the subdomain of  $\Omega$  which is mapped by  $f$  onto the disc  $\{\omega : |\omega| < r\}$ , where  $r < 1$  and  $\omega = f(z)$ . Denote the boundary of  $\Omega_r$  by  $\gamma_r$ . If  $g \in A(\Omega)$ , then  $\frac{g(z)}{f(z)}$  has a simple pole at  $z = \zeta$  and the residue at this pole is

$$\lim_{z \rightarrow \zeta} \frac{(z - \zeta)g(z)}{f(z)} = \frac{g(\zeta)}{f'(\zeta)}.$$

By the Residue Theorem,

$$\frac{g(\zeta)}{f'(\zeta)} = \frac{1}{2\pi i} \int_{\gamma_r} \frac{g(z)}{f(z)} dz = \frac{1}{2\pi i r^2} \int_{\gamma_r} \overline{f(z)} g(z) dz.$$

since  $|f(z)|^2 = r^2$  for  $z \in \gamma_r$ . Using Green's formula, we obtain

$$\frac{g(\zeta)}{f'(\zeta)} = \frac{1}{\pi r^2} \iint_{\Omega_r} \overline{f'(z)} g(z) dx dy.$$

Letting  $r \rightarrow 1$ , we get

$$g(\zeta) = \iint_{\Omega} \frac{\overline{f'(z)} f'(\zeta)}{\pi} g(z) dx dy.$$

In other words, the function

$$K(z, \zeta) = \frac{f'(z) \overline{f'(\zeta)}}{\pi} \quad (2.104)$$

has the reproducing property for  $A(\Omega)$  and is therefore the Bergman kernel. For  $z = \zeta$ , it follows that

$$K(\zeta, \zeta) = \frac{f'(\zeta)^2}{\pi},$$

which implies on using (2.104)

$$f'(z) = \sqrt{\frac{\pi}{K(\zeta, \zeta)}} k(z, \zeta).$$

This completes the proof. □

*Remarks*

- (i) In only a few cases, it is possible to obtain a representation for the kernel function in closed form. It is easy to find a series representation with respect to some complete orthonormal system  $\{\phi_j\}$ , because by (2.101), the Fourier coefficients are

$$u_j = (K(\cdot, \zeta), \phi_j) = \overline{\phi_j(\zeta)}, \quad j = 1, 2, \dots$$

and the Bergman kernel has the series representation [see Theorem 2.9.15(iii)]

$$K(z, \zeta) = \sum_{j=1}^{\infty} \overline{\phi_j(\zeta)} \phi_j(z), \quad z, \zeta \in \Omega.$$

- (ii) Consider the special case where  $\Omega = D\{z : |z| < 1\}$ . According to (vi) of Examples 2.9.16, the set  $\phi_n(z) = \sqrt{\frac{n}{\pi}} z^{n-1}$   $n = 1, 2, \dots$  is an orthonormal system in  $A(D)$ . Thus,

$$K(z, \zeta) = \sum_{n=1}^{\infty} \frac{n}{\pi} \overline{\zeta}^{n-1} z^{n-1} = \frac{1}{\pi} \frac{1}{(1 - z\bar{\zeta})^2}, \quad z, \zeta \in \Omega.$$

is the kernel function of  $D$ . The series converges uniformly in  $|\zeta| \leq r, r < 1$ .

The reproducing property becomes

$$f(\zeta) = \frac{1}{\pi} \iint_D \frac{f(z)}{(1 - \bar{z}\zeta)^2} dx dy.$$

### Special Case of Browder Fixed Point Theorem

Let  $C$  be a nonempty convex, closed and bounded subset of a Hilbert space  $H$  and let  $T$  be a map from  $C$  into  $C$  such that

$$\|Tx - Ty\| \leq \|x - y\| \quad \text{for all } x, y \in C.$$

Then  $T$  has at least one fixed point.

Solution: for each  $n \in \mathbb{N}$ , let

$$T_n(x) = \frac{1}{n} a + \frac{n-1}{n} T(x),$$

where  $a \in C$  is fixed. Then  $T_n$  is a contraction and therefore has a fixed point  $x_n \in C$ . Indeed, for  $x, y \in C$ ,  $\|T_n(x) - T_n(y)\| = \frac{n-1}{n} \|Tx - Ty\| \leq \frac{n-1}{n} \|x - y\|$ .

Since  $C$  is a bounded subset of  $H$  and  $\{x_n\}_{n \geq 1}$  is in  $C$ , it follows that there exists a subsequence  $\{x_{n_j}\}_{j \geq 1}$  such that  $x_{n_j} \xrightarrow{w} x$ , say [Theorem 3.1.5]. By the Banach–Saks Theorem [Problem 2.12.P2],  $\{x_{n_j}\}_{j \geq 1}$  has subsequence such that a sequence of certain convex combinations of its terms converges strongly to  $x$ . Consequently,  $x \in C$  as  $C$  is convex and (strongly) closed. We shall prove that  $x$  is a fixed point of  $T$ .

For any point  $y$  in  $H$ , we note that

$$\|x_{n_j} - y\|^2 = \|x_{n_j} - x\|^2 + \|x - y\|^2 + 2\Re(x_{n_j} - x, x - y), \quad (2.105)$$

where  $2\Re(x_{n_j} - x, x - y) \rightarrow 0$  as  $j \rightarrow \infty$ , since  $x_{n_j} - x \rightarrow 0$  weakly in  $H$ .

Observe that

$$\begin{aligned} T(x_{n_j}) - x_{n_j} &= T(x_{n_j}) - T_{n_j}(x_{n_j}) = T(x_{n_j}) - \frac{1}{n_j}a - \frac{n_j - 1}{n_j}T(x_{n_j}) \\ &= \frac{1}{n_j}(T(x_{n_j}) - a) \rightarrow 0 \text{ as } j \rightarrow \infty. \end{aligned} \quad (2.106)$$

Setting  $y = T(x)$  in (2.105), we have

$$\lim_{j \rightarrow \infty} \left\{ \|x_{n_j} - T(x)\|^2 - \|x_{n_j} - x\|^2 \right\} = \|x - T(x)\|^2. \quad (2.107)$$

On the other hand, using the hypothesis,

$$\|T(x_{n_j}) - T(x)\| \leq \|x_{n_j} - x\|.$$

Hence

$$\begin{aligned} \|x_{n_j} - T(x)\| &\leq \|x_{n_j} - T(x_{n_j})\| + \|T(x_{n_j}) - T(x)\| \\ &\leq \|x_{n_j} - T(x_{n_j})\| + \|x_{n_j} - x\|. \end{aligned}$$

Thus on using (2.106), we obtain

$$\limsup_{j \rightarrow \infty} [\|x_{n_j} - T(x)\| - \|x_{n_j} - x\|] \leq 0$$

and therefore

$$\limsup_{j \rightarrow \infty} \left\{ \|x_{n_j} - T(x)\|^2 - \|x_{n_j} - x\|^2 \right\} \leq 0,$$

which implies, on using (2.107),

$$\|x - T(x)\| = 0.$$

Elements of Hilbert Spaces and Operator Theory

Vasudeva, H.L.

2017, XIII, 522 p. 5 illus., Hardcover

ISBN: 978-981-10-3019-2