In quantum mechanics, by definition, the spin of the above particles is taken to be

$$s = \frac{n-1}{2}$$

(though, strictly speaking, for reasons to become clear when a systematic theoretical study of the spin is undertaken, the total spin is $[s(s+1)]^{1/2}$). Thus, s can assume only integer and half-integer values $\frac{1}{2}$, 1, $\frac{3}{2}$, 2,...; in Fig. 3 we have depicted the case of spin $\frac{1}{2}$.

The experimental arrangement of Stern and Gerlach can be used as an apparatus for a determinative measurement of the *spin component* in the direction **H**. In that case the source of particles would originate in the interaction region, where particles of known spin are interacting. If a particle of integer spin* s leaves a mark at O, then, by definition, it has spin zero in the **H** direction; the first, second,..., $(n-1)/2$ mark above O correspond to spin components in the **H** direction equal to 1, 2,..., $(n-1)/2$, respectively, while the first, second,..., $(n-1)/2$ marks below O correspond to spin components -1, -2,..., $-(n-1)/2$, respectively. In case of a particle of half-integer spin, there will be no middle mark; the first, second,..., $(n-1)/2$ marks above or below O correspond to spin components $\frac{1}{2}$, $\frac{3}{2}$,..., $(n-1)/2$, or $-\frac{1}{2}$, $-\frac{3}{2}$,..., $-(n-1)/2$, respectively. Hence we see that according to the very definition of the spin projection onto a certain axis, that projection can assume only integer values in case of integer-spin particles, and only half-integer values in case of half-integer-spin particles.

The above experimental arrangement can be easily transformed into an apparatus for preparatory measurements of spin by replacing the photoplate with a screen which has apertures at the spots where a beam of particles from the given source had left tracks. It has to be mentioned that no simultaneous measurements of spin in two different directions can be carried out on microparticles—a feature which is in complete agreement with certain properties (noncommutativity of spin-component operators) of the formalism of quantum mechanics.

Here we end this short survey of some of the experimental procedures for measuring some of the basic observables which occur in quantum mechanics, and which will frequently appear in the pages of this book. In Chapter I we start our systematic study of the Hilbert space formalism of quantum mechanics, and related mathematics.

* This means that a beam of such particle with random-oriented spins would have $2s+1$ tracks on the photoplate.

CHAPTER I

Basic Ideas of Hilbert Space Theory

The central object of study in this chapter is the infinite-dimensional Hilbert space. The main goal is to give a rigorous analysis of the problem of expanding a vector in a Hilbert space in terms of an orthogonal basis containing a countable infinity of vectors.

We first review in §1 a few key theorems on vector spaces in general, and in §2 we investigate the basic properties of vector spaces on which an inner product is defined. In order to define convergence in an inner-product space, we introduce in §3 the concept of metric. In §4 we give the basic concepts and theorems on separable Hilbert spaces, concentrating especially on properties of orthonormal bases. We conclude the chapter by illustrating some of the physical applications of these mathematical results with the initial-value problem in wave mechanics.

1. Vector Spaces

1.1. Vector Spaces over Fields of Scalars

A mathematical space is in general a set endowed with some given structure. Such a structure can be given, for instance, by means of certain operations which are defined on the elements of that set. These operations are then required to obey certain general rules, which are called the postulates or the axioms of the mathematical space.

Definition 1.1. Any set $\mathscr{V}$ on which the operations of vector addition and multiplication by a scalar are defined is said to be a *vector* *space* (or *linear space*, or *linear manifold*). The operation of *vector addition* is a mapping,*

$$(f, g) \to f + g, \qquad (f, g) \in \mathscr{V} \times \mathscr{V}, \qquad f + g \in \mathscr{V},$$

of $\mathscr{V} \times \mathscr{V}$ into $\mathscr{V}$, while the operation of multiplication by a scalar a from a field† $\mathscr{F}$ is a mapping

$$(a, f) \to af, \qquad (a, f) \in \mathscr{F} \times \mathscr{V}, \qquad af \in \mathscr{V},$$

of $\mathscr{F} \times \mathscr{V}$ into $\mathscr{V}$. These two vector operations are required to satisfy the following axioms for any $f, g, h \in \mathscr{V}$ and any scalars $a, b \in \mathscr{F}$:

(1) $f + g = g + f$ (commutativity of vector addition).
(2) $(f + g) + h = f + (g + h)$ (associativity of vector addition).
(3) There is a vector $\mathbf{0}$, called the *zero vector*, which satisfies the relation $f + \mathbf{0} = f$ for all $f \in \mathscr{V}$.
(4) $a(f + g) = af + ag$.
(5) $(a + b)f = af + bf$.
(6) $(ab)f = a(bf)$.
(7) $1f = f$, where 1 denotes the unit element in the field.

By following a tacit convention, we denote a mathematical space constructed from a set S by the same letter S, except where ambiguities might arise. Thus, we shall denote by $\mathscr{V}$ the vector space consisting of a set $\mathscr{V}$ together with the vector operations on $\mathscr{V}$ also by $\mathscr{V}$.

When in a vector space the multiplication by a scalar is defined for scalars which are elements of the field $\mathscr{F}$, we say that we are dealing with a *vector space over the field* $\mathscr{F}$. If the field $\mathscr{F}$ is the field of real or complex numbers the vector space is called, respectively, a *real* or a *complex vector space.*

* We remind the reader that a mapping M *of* a set S *into* a set T is any unambiguous rule assigning to each element ξ of S a single element $M(\xi)$ of T; $M(\xi)$ is called the image of ξ under the mapping M. The set S is the domain of definition of M, while the subset $T_1 \subset T$ of all image points $M(\xi)$, $T_1 = \{\eta = M(\xi), \xi \in S\}$, is the range of M. If $T_1 = T$, then we say that M is a mapping of the set S *onto* the set T.

If $S_1, \ldots, S_n$ are sets, then $S_1 \times \cdots \times S_n$ denotes the family $(\xi_1, \ldots, \xi_n)$ of all n-tuples of elements $\xi_1 \in S_1, \ldots, \xi_n \in S_n$, and is called the Cartesian product of the sets $S_1, \ldots, S_n$.

† A field is a set on which field operations of summation and multiplication are defined, i.e., operations satisfying certain axioms. We do not give these axioms because in the sequel we are interested only in two special well-known fields: the field of real numbers $\mathbb{R}^1$ and the field of complex numbers $\mathbb{C}^1$ consisting, respectively, of the set of real numbers $\mathbb{R}^1$ and the set of complex numbers $\mathbb{C}^1$ on which the field operations are ordinary summation and multiplication of numbers (see Birkhoff and MacLane [1953]).

As an example (see also Exercises 1.1, 1.2, and 1.3) of a real vector space consider the family $(\mathbb{R}^n)$ of one-column real matrices and define for

$$\alpha = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \qquad \beta = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

vector summation by the mapping

$$(\alpha, \beta) \to \alpha + \beta = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}, \tag{1.1}$$

and for any scalar $a \in \mathbb{R}^1$ define multiplication of α by a as the mapping

$$(a, \alpha) \to a\alpha = \begin{pmatrix} aa_1 \\ \vdots \\ aa_n \end{pmatrix}. \tag{1.2}$$

It is easy to check that Axioms 1–7 in Definition 1.1 are satisfied.

Analogously we can define the complex vector space $(\mathbb{C}^n)$ by introducing in the set $\mathbb{C}^n$ of one-column matrices vector operations defined by the mapping (1.1) and (1.2), where now $\alpha, \beta \in \mathbb{C}^n$, and therefore $a_1, \ldots, a_n$, $b_1, \ldots, b_n$, as well as the scalar a, are complex numbers.

1.2. Linear Independence of Vectors

Theorem 1.1. Each vector space $\mathscr{V}$ has only one zero vector $\mathbf{0}$, and each element f of a vector space has one and only one inverse $(-f)$. For any $f \in \mathscr{V}$,

$$0f = \mathbf{0}, \qquad (-1)f = (-f).$$

Proof. If there are two zero vectors $\mathbf{0}_1$ and $\mathbf{0}_2$, they both have to satisfy Axiom 3 in Definition 1.1,

$$f = f + \mathbf{0}_1 = f + \mathbf{0}_2$$

for all f. Hence, by taking $f = \mathbf{0}_1$ we get $\mathbf{0}_1 = \mathbf{0}_1 + \mathbf{0}_2$, and then by taking $f = \mathbf{0}_2$ we deduce that $\mathbf{0}_2 = \mathbf{0}_2 + \mathbf{0}_1 = \mathbf{0}_1 + \mathbf{0}_2 = \mathbf{0}_1$. Now

$$f = 1f = (1 + 0)f = 1f + 0f = f + 0f$$

and therefore $0f = \mathbf{0}$. We have

$$(-1)f + f = (-1)f + 1f = (-1 + 1)f = 0f = \mathbf{0},$$

which proves the existence of an inverse $(-f) = (-1)f$ for f. This inverse $(-f)$ is unique, because if there is another $f_1 \in \mathscr{V}$ such that $f + f_1 = \mathbf{0}$, we have

$$(-f) = (-f) + \mathbf{0} = (-f) + (f + f_1) = [(-f) + f] + f_1 = \mathbf{0} + f_1 = f_1. \quad \text{Q.E.D.}$$

Definition 1.2. The vectors $f_1, \ldots, f_n$ are said to be *linearly independent* if the relation

$$c_1 f_1 + \cdots + c_n f_n = \mathbf{0}, \qquad c_1, \ldots, c_n \in \mathfrak{P},$$

has $c_1 = \cdots = c_n = 0$ as the *only* solution. A subset S (finite or infinite) of a vector space $\mathscr{V}$ is called a *set of linearly independent* vectors if any finite number of *different* vectors from S are linearly independent. The *dimension* of a vector space $\mathscr{V}$ is the least upper bound (which can be finite or positive infinite) of the set of all integers ν for which there are ν linearly independent vectors in $\mathscr{V}$.

1.3. Dimension of a Vector Space

When the maximal number of linearly independent vectors in the vector space $\mathscr{V}$ is finite and equal to n, then by the above definition $\mathscr{V}$ is n dimensional; otherwise the dimension of $\mathscr{V}$ is $+\infty$, and $\mathscr{V}$ is said to be infinite dimensional.

Theorem 1.2. If the vector space $\mathscr{V}$ is n dimensional $(n < +\infty)$, then there is at least one set $f_1, \ldots, f_n$ of linearly independent vectors, and each vector $f \in \mathscr{V}$ can be expanded in the form

$$f = a_1 f_1 + \cdots + a_n f_n, \tag{1.3}$$

where the coefficients $a_1, \ldots, a_n$ (which are scalars) are uniquely determined by f.

Proof. If $f = \mathbf{0}$, (1.3) is established by taking $a_1 = \cdots = a_n = 0$. For $f \neq \mathbf{0}$, the equation

$$cf + c_1 f_1 + \cdots + c_n f_n = 0 \tag{1.4}$$

should have a solution with $c \neq 0$ due to the assumption that $f_1, \ldots, f_n$ are linearly independent, while $f, f_1, \ldots, f_n$ have to be linearly dependent because $\mathscr{V}$ is n dimensional. From (1.4) we get

$$f = (-c_1/c) f_1 + \cdots + (-c_n/c) f_n,$$

which establishes (1.3). If we also had

$$f = b_1 f_1 + \cdots + b_n f_n, \tag{1.5}$$

then by subtracting (1.5) from (1.3) we get

$$(a_1 - b_1) f_1 + \cdots + (a_n - b_n) f_n = \mathbf{0}.$$

As $f_1, \ldots, f_n$ are linearly independent we deduce that $a_1 - b_1 = 0, \ldots, a_n - b_n = 0$, thus proving that $a_1, \ldots, a_n$ are uniquely determined when f is given. Q.E.D.

Definition 1.3. We say that the (finite or infinite) set S *spans* the vector space $\mathscr{V}$ if every vector in $\mathscr{V}$ can be written as a linear combination

$$f = a_1 h_1 + \cdots + a_n h_n, \qquad h_1, \ldots, h_n \in S$$

of a finite number of vectors belonging to S; if S is in addition a set of linearly independent vectors, then S is called a *vector basis* of $\mathscr{V}$.

Theorem 1.3. If the set $\{g_1, \ldots, g_m\}$ is a vector basis of the n-dimensional $(n < +\infty)$ vector space $\mathscr{V}$, then necessarily $m = n$.

Proof. As $\mathscr{V}$ is n-dimensional, there must be n linearly independent vectors $f_1, \ldots, f_n$. If the set $\{g_1, \ldots, g_m\}$ is a vector basis in $\mathscr{V}$, we can write

$$\begin{aligned} f_1 &= a_{11} g_1 + \cdots + a_{m1} g_m \\ &\vdots \\ f_n &= a_{1n} g_1 + \cdots + a_{mn} g_m. \end{aligned} \tag{1.6}$$

Thus, if we try to satisfy the equation

$$x_1 f_1 + \cdots + x_n f_n = \mathbf{0}, \tag{1.7}$$

we get by substituting $f_1, \ldots, f_n$ in (1.7) with the expressions in (1.6)

$$(a_{11} x_1 + \cdots + a_{1n} x_n) g_1 + \cdots + (a_{m1} x_1 + \cdots + a_{mn} x_n) g_m = \mathbf{0}. \tag{1.8}$$

Since $g_1, \ldots, g_m$ are assumed to be linearly independent, the above equation has a solution in $x_1, \ldots, x_n$ if and only if

$$\begin{aligned} a_{11} x_1 + \cdots + a_{1n} x_n &= 0 \\ &\vdots \\ a_{m1} x_1 + \cdots + a_{mn} x_n &= 0. \end{aligned} \tag{1.9}$$

However, as $f_1, \ldots, f_n$ are also linearly independent, (1.7) or equivalently (1.8) or (1.9) should have as the only solution the trivial one $x_1 = \cdots =$

$x_n = 0$. Now, $m \leqslant n$ because $\mathscr{V}$ is n dimensional and $g_1, \ldots, g_m$ are linearly independent (see Definition 1.2); therefore, (1.9) has only a trivial solution if and only if $m = n$. Q.E.D.

Definition 1.4. A subset $\mathscr{V}_1$ of a vector space $\mathscr{V}$ is a *vector subspace* (*linear subspace*) of $\mathscr{V}$ if it is closed under the vector operations, i.e., if $f + g \in \mathscr{V}_1$ and $af \in \mathscr{V}_1$ whenever $f, g \in \mathscr{V}_1$ and for any scalar a. A vector subspace $\mathscr{V}_1$ of $\mathscr{V}$ is said to be *nontrivial* if it is different from $\mathscr{V}$ and from the set $\{\mathbf{0}\}$.

From the very definition of the dimension of a vector space $\mathscr{V}$ we can conclude that the dimension of a vector subspace $\mathscr{V}_1$ of $\mathscr{V}$ cannot exceed the dimension of $\mathscr{V}$.

1.4. Isomorphism of Vector Spaces

Definition 1.5. Two vector spaces $\mathscr{V}_1$ and $\mathscr{V}_2$ over the same field are *isomorphic* if there is a one-to-one mapping $\mathscr{V}_1$ *onto* $\mathscr{V}_2$ which has the properties that if f_2 and g_2, $f_2, g_2 \in \mathscr{V}_2$, are the images of f_1 and g_1, $f_1, g_1 \in \mathscr{V}_1$, respectively, then for any scalar a, af_2 is the image of af_1

$$af_1 \leftrightarrow af_2,$$

and $f_2 + g_2$ is the image of $f_1 + g_1$

$$f_1 + g_1 \leftrightarrow f_2 + g_2.$$

The importance of the isomorphism of two vector spaces $\mathscr{V}_1$ and $\mathscr{V}_2$ lies in the obvious fact that two such spaces have an identical vector structure. It is easy to see that the relation of isomorphism is transitive (see Exercise 1.6), i.e., if $\mathscr{V}_1$ and $\mathscr{V}_2$ as well as $\mathscr{V}_2$ and $\mathscr{V}_3$ are isomorphic, then $\mathscr{V}_1$ and $\mathscr{V}_3$ are also isomorphic.

Theorem 1.4. All complex (real) n-dimensional ($n < +\infty$) vector spaces are isomorphic to the vector space $(\mathbb{C}^n)$ $[(\mathbb{R}^n)$ in case of real vector spaces$]$.

Proof. Consider the case of an n-dimensional vector space $\mathscr{V}$. According to Theorem 1.2 there is a vector basis consisting of n vectors $f_1, \ldots, f_n$, and each vector $f \in \mathscr{V}$ can be expanded in the form (1.3), where $a_1, \ldots, a_n \in \mathbb{C}^1$ are uniquely determined by f. Consequently

$$f \to \alpha_f = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in (\mathbb{C}^n)$$

is a mapping of $\mathscr{V}$ into $(\mathbb{C}^n)$. Furthermore, this is a one-to-one mapping of $\mathscr{V}$ *onto* $(\mathbb{C}^n)$ because to any

$$\beta = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in (\mathbb{C}^n)$$

corresponds a unique $f = b_1 f_1 + \cdots + b_n f_n$ such that $\beta = \alpha_f$. It is also easy to see that

$$f + g \to \alpha_{f+g} = \alpha_f + \alpha_g,$$
$$af \to \alpha_{af} = a\alpha_f.$$

Since isomorphism of vector spaces is a transitive relation (see Exercise 1.6) we can conclude that all n-dimensional complex vector spaces are mutually isomorphic, because each of them is isomorphic to $(\mathbb{C}^n)$. Q.E.D.

Exercises

1.1. Check that the set of all $m \times n$ complex matrices constitutes an $m \cdot n$ dimensional complex vector space if vector addition is defined as being addition of matrices, and multiplication by a scalar is multiplication of a matrix by a complex number.

1.2. Show that the set $\mathbb{C}^1$ of all complex numbers becomes a two-dimensional *real* vector space if vector addition is identical to addition of complex numbers, and multiplication by a scalar is multiplication of a complex number (the vector) by a real number (the scalar).

1.3. Show that the family $\mathscr{C}^0(\mathbb{R}^1)$ of all complex-valued continuous functions defined on the real line is an infinite-dimensional vector space if the vector sum $f + g$ of $f(x), g(x) \in \mathscr{C}^0(\mathbb{R}^1)$ is the function $(f + g)(x) = f(x) + g(x)$, and the product af of $f(x) \in \mathscr{C}^0(\mathbb{R}^1)$ with $a \in \mathbb{C}^1$ is the function $(af)(x) = af(x)$. The zero vector is taken to be the function $f(x) \equiv 0$.

1.4. Prove that if $\mathscr{K}$ is a family of linear subspaces L of a vector space $\mathscr{V}$, then their set intersection $\bigcap_{L \in \mathscr{K}} L$ is also a vector subspace of $\mathscr{V}$.

1.5. Show that if S is any subset of a vector space $\mathscr{V}$, then there is a unique smallest vector subspace $\mathscr{V}_S$ containing S (called the vector subspace spanned by S).

1.6. Verify that the relation of isomorphism of vector spaces is:

(a) reflexive, i.e., every vector space $\mathscr{V}$ is isomorphic to itself;
(b) symmetric, i.e., if $\mathscr{V}_1$ is isomorphic to $\mathscr{V}_2$, then $\mathscr{V}_2$ is isomorphic to $\mathscr{V}_1$;

(c) transitive, i.e., if $\mathscr{V}_1$ is isomorphic to $\mathscr{V}_2$ and $\mathscr{V}_2$ is isomorphic to $\mathscr{V}_3$, then $\mathscr{V}_1$ is isomorphic to $\mathscr{V}_3$.

1.7. Prove that the following subsets of the set $\mathscr{C}^0(\mathbb{R}^1)$ (see Exercise 1.3) are vector subspaces of the vector space $\mathscr{C}^0(\mathbb{R}^1)$:

(a) the set $\mathscr{P}_\infty$ of all polynomials with complex coefficients;
(b) the set $\mathscr{P}_n$ of all polynomials of *at most* degree n.

Show that $\mathscr{P}_\infty \supset \mathscr{P}_n$.

2. Euclidean Spaces

2.1. Inner Products on Vector Spaces

A *Euclidean* (or *inner product* or *unitary*) space $\mathscr{E}$ is a vector space on which an inner product is defined. The Euclidean space is called real or complex if the vector space on which the inner product is defined is, respectively, real or complex.

Definition 2.1. An *inner* (or *scalar*) product $\langle \cdot \mid \cdot \rangle$ on the complex vector space $\mathscr{V}$ is a mapping of the set $\mathscr{V} \times \mathscr{V}$ into the set $\mathbb{C}^1$ of complex numbers

$$(f, g) \to \langle f \mid g \rangle, \qquad (f, g) \in \mathscr{V} \times \mathscr{V}, \qquad \langle f \mid g \rangle \in \mathbb{C}^1,$$

which satisfies the following requirements:

(1) $\langle f \mid f \rangle > 0, \quad$ for all $\; f \neq \mathbf{0}$,
(2) $\langle f \mid g \rangle = \langle g \mid f \rangle^*$,
(3) $\langle f \mid ag \rangle = a\langle f \mid g \rangle, \qquad a \in \mathbb{C}^1$,
(4) $\langle f \mid g + h \rangle = \langle f \mid g \rangle + \langle f \mid h \rangle$.

Note that by inserting $f = g = h = \mathbf{0}$ in Point 4 we get $\langle \mathbf{0} \mid \mathbf{0} \rangle = 0$.

Following a notation first introduced by Dirac [1930] and widely adopted by physicists, we denote the inner product of f and g by $\langle f \mid g \rangle$. Mathematicians often prefer the notation (f, g) and replace Point 3 in Definition 2.1 by

$$(af, g) = a(f, g).$$

The above definition can be easily specialized to real vector spaces, in which case the inner product $\langle f \mid g \rangle$ is a real number, and Point 2 of Definition 2.1 becomes $\langle f \mid g \rangle = \langle g \mid f \rangle$. As in quantum physics we deal almost exclusively with complex Euclidean spaces, we limit ourselves from now on to the complex case. Consequently, if not otherwise stated,

whenever we talk about a Euclidean space, we shall mean a *complex* Euclidean space.

Theorem 2.1. In a Euclidean space $\mathscr{E}$, the inner product $\langle f \mid g \rangle$ satisfies the relations

(a) $\langle af \mid g \rangle = a^*\langle f \mid g \rangle$,
(b) $\langle f + g \mid h \rangle = \langle f \mid h \rangle + \langle g \mid h \rangle$.

The proof is obtained by a straightforward application of Points 1–4 in Definition 2.1:

$$\langle af \mid g \rangle = \langle g \mid af \rangle^* = [a\langle g \mid f \rangle]^* = a^*\langle g \mid f \rangle^* = a^*\langle f \mid g \rangle,$$
$$\langle f + g \mid h \rangle = \langle h \mid f + g \rangle^* = [\langle h \mid f \rangle + \langle h \mid g \rangle]^* = \langle h \mid f \rangle^* + \langle h \mid g \rangle^*,$$
$$= \langle f \mid h \rangle + \langle g \mid h \rangle.$$

As an example of a finite-dimensional Euclidean space, we can take the vector space $(\mathbb{C}^n)$ defined in the preceding section, in which we introduce as the inner product of the vectors α and β with the kth components a_k and b_k,

$$\langle \alpha \mid \beta \rangle = a_1^* b_1 + a_2^* b_2 + \cdots + a_n^* b_n .$$

It is easy to check that the above mapping of $\mathbb{C}^n \times \mathbb{C}^n$ into $\mathbb{C}^1$ satisfies the four requirements of Definition 2.1. We shall denote the above Euclidean space with the symbol $l^2(n)$.

An example of an infinite-dimensional Euclidean space is provided by the vector space $[\mathscr{C}^0_{(2)}(\mathbb{R}^1)]$ of all continuous complex-valued functions $f(x)$ on the real line which satisfy

$$\int_{-\infty}^{+\infty} |f(x)|^2 \, dx < +\infty, \tag{2.1}$$

in which the inner product (see Exercise 2.1) is

$$\langle f \mid g \rangle = \int_{-\infty}^{+\infty} f^*(x) \, g(x) \, dx. \tag{2.2}$$

Theorem 2.2. Any two elements f, g of a Euclidean space $\mathscr{E}$ satisfy the Schwarz–Cauchy inequality

$$|\langle f \mid g \rangle|^2 \leqslant \langle f \mid f \rangle \langle g \mid g \rangle.$$

Proof. For any given $f, g \in \mathscr{E}$ and any complex number a we have, from property 1 in Definition 2.1 and the comment following it,

$$\langle f + ag \mid f + ag \rangle \geqslant 0.$$

In particular, if we take in the above inequality

$$a = \lambda \frac{\langle f \mid g\rangle^*}{|\langle f \mid g\rangle|}, \qquad \lambda = \lambda^*,$$

we easily show that the inequality

$$g(\lambda) = \lambda^2\langle g \mid g\rangle + 2\lambda\, |\langle f \mid g\rangle| + \langle f \mid f\rangle \geqslant 0$$

is true for all real values of λ. A necessary and sufficient condition that $g(\lambda) \geqslant 0$ is that the discriminant of the quadratic polynomial $g(\lambda)$ is not positive

$$|\langle f \mid g\rangle|^2 - \langle f \mid f\rangle\langle g \mid g\rangle \leqslant 0,$$

from which the Schwarz–Cauchy inequality follows immediately. Q.E.D.

2.2. The Concept of Norm

The family of all Euclidean spaces is obviously contained in the family of vector spaces. There is another family of vector spaces with special properties which is of great importance in mathematics: the family of normed spaces.

Definition 2.2. A mapping

$$f \to \|f\|, \qquad f \in \mathscr{V}, \qquad \|f\| \in \mathbb{R}^1,$$

of a complex vector space $\mathscr{V}$ into the set of real numbers is called a *norm* if it satisfies the following conditions:

(1) $\|f\| > 0$ for $f \neq \mathbf{0}$,
(2) $\|\mathbf{0}\| = 0$,
(3) $\|af\| = |a|\,\|f\|$ for all $a \in \mathbb{C}^1$,
(4) $\|f + g\| \leqslant \|f\| + \|g\|$ (the triangle inequality).

We denote the above norm by $\|\cdot\|$.

For a real vector space, we require in Point 3 that $a \in \mathbb{R}^1$.

The last requirement in Definition 2.2 is known as the triangle inequality because it represents in a two- or three-dimensional real vector space a relation satisfied by the sides of a triangle formed by three vectors f, g and $f + g$.

A real (complex) vector space on which a particular norm is given is called a real (complex) normed vector space. A Euclidean space is a special case of a normed space; this can be seen from the following theorem.

Theorem 2.3. In a Euclidean space $\mathscr{E}$ with the inner product $\langle f \mid g\rangle$ the real-valued function

$$\|f\| = \sqrt{\langle f \mid f\rangle} \tag{2.3}$$

is a norm.

Proof. The only one of the four properties of a norm which is not satisfied by (2.3) in an evident way is the triangle inequality. We easily get

$$\begin{aligned}\|f+g\|^2 &= \langle f+g \mid f+g\rangle = \langle f + f\rangle + \langle f \mid g\rangle + \langle g \mid f\rangle + \langle g \mid g\rangle \\ &= \langle f \mid f\rangle + 2\mathrm{Re}\langle f \mid g\rangle + \langle g \mid g\rangle.\end{aligned} \tag{2.4}$$

From the Schwarz–Cauchy inequality we have

$$|\mathrm{Re}\langle f \mid g\rangle| \leqslant |\langle f \mid g\rangle| \leqslant \|f\|\,\|g\|,$$

which when inserted in (2.4) yields

$$\|f+g\|^2 \leqslant \|f\|^2 + 2\|f\|\,\|g\| + \|g\|^2 = (\|f\| + \|g\|)^2.$$

The above relation leads immediately to the triangle inequality. Q.E.D.

2.3. Orthogonal Vectors and Orthonormal Bases

Some elementary geometrical concepts valid for real two- or three-dimensional Euclidean spaces can be generalized in a straightforward manner to any Euclidean space.

Definition 2.3. In a Euclidean space $\mathscr{E}$ two vectors f and g are called orthogonal, symbolically $f \perp g$, if $\langle f \mid g\rangle = 0$. Two subsets R and S of $\mathscr{E}$ are said to be *orthogonal* (symbolically, $R \perp S$) if each vector in R is orthogonal to each vector in S. A set of vectors in which any two vectors are orthogonal is called an *orthogonal system* of vectors. A vector f is said to be *normalized* if $\|f\| = 1$. An orthogonal system of vectors is called an orthonormal system if each vector in the system is normalized.

Theorem 2.4. If S is a finite or countably infinite set of vectors in a Euclidean space $\mathscr{E}$ and (S) is the vector subspace of $\mathscr{E}$ spanned by S, then there is an orthonormal system T of vectors which spans (S), i.e., for which $(T) = (S)$; T is a finite set when S is a finite set.

Proof. As the set S is at most countable we can write it in the form

$$S = \{f_1, f_2, \ldots\}$$

by assigning each vector in S to a natural number. In general some of the vectors in S might be linearly dependent. We can build from S

another set S_0 of linearly independent vectors spanning the same subspace (S), i.e., such that $(S_0) = (S)$, by the following procedure (which should be applied consecutively on $n = 1, 2, \ldots$): if f_n is the zero vector or is linearly dependent on $f_1, \ldots, f_{n-1}$, then discard it; otherwise include it in S_0. Thus we get a set S_0 of linearly independent vectors

$$S_0 = \{g_1, g_2, \ldots\}, \qquad (S_0) = (S).$$

We can obtain from S_0 an orthonormal set T such that $(T) = (S_0)$ by the so-called *Schmidt* (or *Gram–Schmidt*) *orthonormalization procedure*.

Since $g_1 \neq \mathbf{0}$, we can introduce the vector

$$e_1 = \frac{g_1}{\|g_1\|},$$

which is normalized. Proceeding by induction, assume that we have obtained the orthonormal system of vectors $e_1, \ldots, e_{n-1}$. Then e_n is given by

$$e_n = \frac{g_n - \langle e_{n-1} \mid g_n \rangle e_{n-1} - \cdots - \langle e_1 \mid g_n \rangle e_1}{\|g_n - \langle e_{n-1} \mid g_n \rangle e_{n-1} - \cdots - \langle e_1 \mid g_n \rangle e_1\|}.$$

The above vector is certainly well defined, since the denominator of the above expression is different from zero; namely, if it were zero, then we would have

$$g_n - \langle e_{n-1} \mid g_n \rangle e_{n-1} - \cdots - \langle e_1 \mid g_n \rangle e_1 = 0,$$

i.e., g_n would depend on $e_1, \ldots, e_{n-1}$. However, by solving the equations for $e_1, \ldots, e_{n-1}$, it is easy to see that we have

$$\begin{aligned} g_1 &= c_{1,1}e_1 \\ &\vdots \\ g_{n-1} &= c_{n-1,1}e_1 + c_{n-1,2}e_2 + \cdots + c_{n-1,n-1}e_{n-1}, \end{aligned}$$

and therefore if g_n depended on $e_1, \ldots, e_{n-1}$, then it would also depend on $g_1, \ldots, g_{n-1}$, contrary to the fact that S_0 consists only of linearly independent vectors.

The vectors of T are obviously normalized. In order to prove that T is an orthonormal system, assume that we have proved that $\langle e_i \mid e_j \rangle = \delta_{ij}$ for $i, j = 1, \ldots, n-1$. Then we have for $m < n$

$$\langle e_m \mid e_n \rangle = \frac{1}{\|g_n - \cdots - \langle e_1 \mid g_1 \rangle e_1\|} \left(\langle e_m \mid g_n \rangle - \sum_{k=1}^{n-1} \langle e_k \mid g_n \rangle \cdot \delta_{km} \right) = 0,$$

which proves that $\langle e_i \mid e_j \rangle = \delta_{ij}$ for $i, j = 1, \ldots, n$. Thus, by induction T is orthonormal.

As we have for any n that $e_1, \ldots, e_n$ can be expressed in terms of $g_1, \ldots, g_n$, and vice versa, we can conclude that $(T) = (S_0)$. Q.E.D.

2.4. Isomorphism of Euclidean Spaces

We introduce now a concept of isomorphism of Euclidean spaces, which makes two isomorphic Euclidean spaces identical from the point of view of their vector structure as well as from the point of view of the structure induced by the inner product.

Definition 2.4. Two Euclidean spaces $\mathscr{E}_1$ and $\mathscr{E}_2$ with inner products $\langle \cdot \mid \cdot \rangle_1$ and $\langle \cdot \mid \cdot \rangle_2$, respectively, are *isomorphic* (or *unitarily equivalent*) if there is a mapping of $\mathscr{E}_1$ *onto* $\mathscr{E}_2$

$$f_1 \to f_2, \qquad f_1 \in \mathscr{E}_1, \qquad f_2 \in \mathscr{E}_2$$

such that if for any $f_1, g_1 \in \mathscr{E}_1$ the vector $f_2 \in \mathscr{E}_2$ is the image of f_1 and the vector $g_2 \in \mathscr{E}_2$ is the image of g_1, then

$$\begin{gathered} f_1 + g_1 \to f_2 + g_2, \\ af_1 \to af_2, \qquad a \in \mathbb{C}^1, \\ \langle f_1 \mid g_1 \rangle_1 = \langle f_2 \mid g_2 \rangle_2. \end{gathered}$$

A mapping having the above properties is called a *unitary* transformation of $\mathscr{E}_1$ onto $\mathscr{E}_2$.

Theorem 2.5. All complex Euclidean n-dimensional spaces are isomorphic to $l^2(n)$, and consequently (see Exercise 2.8) mutually isomorphic.

Proof. If $\mathscr{E}$ is an n-dimensional Euclidean space, there is according to Theorem 1.2 a set of n vectors $f_1, \ldots, f_n$ spanning $\mathscr{E}$. According to Theorem 2.4, we can find an orthonormal system of n vectors $e_1, \ldots, e_n$ which also spans $\mathscr{E}$. It is easy to check (see Exercise 2.7) that the mapping

$$f \leftrightarrow \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \qquad a_1 = \langle e_1 \mid f \rangle, \ldots, a_n = \langle e_n \mid f \rangle, \tag{2.5}$$

provides an isomorphism between $\mathscr{E}$ and $l^2(n)$. Q.E.D.

Obviously, a similar theorem can be proved for real Euclidean spaces.

Theorem 2.6. A unitary transformation

$$f_1 \to f_2, \qquad f_1 \in \mathscr{E}_1, \qquad f_2 \in \mathscr{E}_2, \tag{2.6}$$

of the Euclidean space $\mathscr{E}_1$ onto the Euclidean space $\mathscr{E}_2$ has a unique inverse mapping which is a unitary transformation of $\mathscr{E}_2$ onto $\mathscr{E}_1$.

Proof. We note that since

$$\|f_1 - g_1\|_1 = \|f_2 - g_2\|_2,$$

the images f_2 and g_2 of f_1 and g_1, respectively, are distinct whenever $f_1 \neq g_1$. Since the unitary map of $\mathscr{E}_1$ is onto $\mathscr{E}_2$, we conclude that the inverse of the mapping (2.6) exists.

We leave to the reader the details of the remainder of the proof.

Exercises

2.1. Show that for a finite interval I

$$\langle f \mid g\rangle = \int_I f^*(x)\, g(x)\, dx$$

is an inner product on the vector space $\mathscr{C}^0(I)$.

2.2. Show that the vector space $\mathscr{C}^0_{(2)}(\mathbb{R}^1)$ introduced in Section 2 is a subspace of the vector space $\mathscr{C}^0(\mathbb{R}^1)$.

2.3. Prove that (2.2) is an inner product in $\mathscr{C}^0_{(2)}(\mathbb{R}^1)$.

2.4. Show that

$$|\langle f \mid g\rangle|^2 = \langle f \mid f\rangle\langle g \mid g\rangle,$$
$$\|f + g\| = \|f\| + \|g\|,$$

if and only if either f is a multiple of g, i.e., if $f = ag$, $a \in \mathbb{C}^1$, or $g = \mathbf{0}$, and if in addition $a \geqslant 0$ in case of the second relation.

2.5. Show that if T is an orthonormal system of vectors, then all the vectors in T are necessarily linearly independent.

2.6. Prove that a subspace of a Euclidean space is also a Euclidean space.

2.7. Show that the mapping (2.5) is a mapping of $\mathscr{E}$ *onto* $l^2(n)$, and that it satisfies the requirements of isomorphism given in Definition 2.4.

2.8. Show that the relation of isomorphism of inner-product spaces is an equivalence relation, i.e., it is (see Exercise 1.6) reflexive, symmetric, and transitive.

3. Metric Spaces

3.1. Convergence in Metric Spaces

In an n-dimensional Euclidean space $\mathscr{E}$ we can always find, due to Theorems 1.2 and 2.4, a basis of n vectors $e_1, \ldots, e_n$ which constitute an orthonormal system. We can then expand any vector f of $\mathscr{E}$ in that basis

$$f = \sum_{k=1}^{n} a_k e_k. \tag{3.1}$$

We easily see that $a_k = \langle e_k \mid f\rangle$.

In an infinite-dimensional Euclidean space not every vector can be expanded in general in terms of a finite number of vectors. We can hope, however, to replace (3.1) with the formula

$$f = \sum_{k=1}^{\infty} a_k e_k,$$

but then we meet with the problem of giving a precise meaning to the convergence of the above series. This problem is solved in its most general form in topology, but for our purposes it will be sufficient to solve it within the context of metric spaces.

Definition 3.1. If S is a given set, a function $d(\xi, \eta)$ on $S \times S$ is a *metric* (or *distance function*) if it fulfills the following requirements for any $\xi, \eta, \zeta \in S$:

(1) $d(\xi, \eta) > 0$ if $\xi \neq \eta$,
(2) $d(\xi, \xi) = 0$,
(3) $d(\xi, \eta) = d(\eta, \xi)$,
(4) $d(\xi, \zeta) \leqslant d(\xi, \eta) + d(\eta, \zeta)$ (triangle inequality).

A set S on which a metric is defined is called a *metric space*.

A metric space does not have to be a linear space. For instance, a bounded open domain in the plane becomes a metric space if the metric is taken to be the distance between each pair of points belonging to that domain; such a domain obviously is not closed under the operations of adding vectors in the plane, but it provides a metric space.

Generalizing from the case of one-, two-, or three-dimensional real Euclidean spaces, we introduce the following notions.

Definition 3.2. An infinite sequence $\xi_1, \xi_2, \ldots$ in a metric space $\mathscr{M}$ is said to converge to the point $\xi \in \mathscr{M}$ if for any $\epsilon > 0$ there is a positive number $N(\epsilon)$ such that $d(\xi, \zeta_n) < \epsilon$ for all $n > N(\epsilon)$. An infinite sequence $\xi_1, \xi_2, \ldots$ is called a *Cauchy sequence* (or a *fundamental sequence*) if for any $\epsilon > 0$ there is a positive number $M(\epsilon)$ such that $d(\xi_m, \xi_n) < \epsilon$ for all $m, n > M(\epsilon)$.

Theorem 3.1. If a sequence $\xi_1, \xi_2, \ldots$ in a metric space $\mathscr{M}$ converges to some $\xi \in \mathscr{M}$, then its limit ξ is unique, and the sequence is a Cauchy sequence.

Proof. If $\xi_1, \xi_2, \ldots$ converges to $\xi \in \mathscr{M}$ and to $\eta \in \mathscr{M}$, then by definition, for any $\epsilon > 0$ there are $N_1(\epsilon)$ and $N_2(\epsilon)$ such that $d(\xi, \xi_n) < \epsilon$ for $n > N_1(\epsilon)$ and $d(\eta, \xi_n) < \epsilon$ for $n > N_2(\epsilon)$. Consequently, for $n > \max(N_1(\epsilon), N_2(\epsilon))$ we get by applying the triangle inequality of Definition 3.1, Point 4,

$$d(\xi, \eta) \leqslant d(\xi, \xi_n) + d(\xi_n, \eta) < 2\epsilon.$$

As $\epsilon > 0$ can be chosen arbitrarily small, we get $d(\xi, \eta) = 0$, which, according to Definition 3.1, can be true only if $\xi = \eta$.

Similarly we get

$$d(\xi_m, \xi_n) \leqslant d(\xi_m, \xi) + d(\xi, \xi_n) < \epsilon$$

if $m, n > N_1(\epsilon/2)$; i.e., the sequence $\xi_1, \xi_2, \ldots$ is also a Cauchy sequence.

Q.E.D.

3.2. Complete Metric Spaces

In case of sequences of real numbers, every Cauchy sequence is convergent, i.e., the set $\mathbb{R}^1$ of all real numbers is complete. We state this generally in Definition 3.3.

Definition 3.3. A metric space $\mathscr{M}$ is *complete* if every Cauchy sequence converges to an element of $\mathscr{M}$.

Not every metric space is complete, as exemplified by the set $\mathscr{R}$ of all rational numbers with the metric $d(m_1/n_1, m_2/n_2) = | m_1/n_1 - m_2/n_2 |$, which is incomplete. However, we know that the set $\mathscr{R}$ is everywhere dense in the set $\mathbb{R}^1$; we state this generally as follows:

Definition 3.4. A subset S of a metric space $\mathscr{M}$ is (*everywhere*) *dense* in $\mathscr{M}$ if for any given $\epsilon > 0$ and any $\xi \in \mathscr{M}$ there is an element η belonging to S for which $d(\xi, \eta) < \epsilon$.

We can reexpress the above definition after introducing a few topological concepts, generalized from the case of sets in one, two, or three real dimensions.

Definition 3.5. If ξ is an element of a metric space $\mathscr{M}$, then the set of all points η satisfying the inequality $d(\xi, \eta) < \epsilon$ for some $\epsilon < 0$ is called the ϵ *neighborhood* of ξ. If S is a subset of $\mathscr{M}$, a point $\zeta \in \mathscr{M}$ is called an *accumulation* (or *cluster* or *limit*) *point* of S if every ϵ neighborhood of ζ contains a point of S. The set $\bar{S}$ consisting of all the cluster points of S is called the *closure* of S. Obviously always $S \subset \bar{S}$; if $S = \bar{S}$ then S is called a *closed* set.

We see that the subset S of a metric space $\mathscr{M}$ is everywhere dense in $\mathscr{M}$ if and only if $\mathscr{M}$ is its closure, i.e., if and only if $\bar{S} = \mathscr{M}$.

The procedure of completing the set $\mathscr{R}$ of rational numbers by embedding it in the set of all real numbers can be generalized.

Definition 3.6. A metric space $\mathscr{M}$ is said to be *densely embedded* in the metric space $\tilde{\mathscr{M}}$ if there is an isometric mapping of $\mathscr{M}$ into $\tilde{\mathscr{M}}$, and if the image set $\mathscr{M}'$ of $\mathscr{M}$ in $\tilde{\mathscr{M}}$ is everywhere dense in $\tilde{\mathscr{M}}$.

A one-to-one mapping $\xi \leftrightarrow \tilde{\xi}$ of a metric space $\mathscr{M}$ into another metric space $\tilde{\mathscr{M}}$ is called *isometric* if it preserves distances, i.e., if $d_1(\xi, \eta) = d_2(\tilde{\xi}, \tilde{\eta})$ for $\xi, \eta \in \mathscr{M}$ and $\tilde{\xi}, \tilde{\eta} \in \tilde{\mathscr{M}}$ whenever $\xi \leftrightarrow \tilde{\xi}$ and $\eta \leftrightarrow \tilde{\eta}$.

3.3. Completion of a Metric Space

***Theorem 3.2.** Every incomplete metric space $\mathscr{M}$ can be embedded in a complete metric space $\tilde{\mathscr{M}}$, called the *completion* of $\mathscr{M}$.

The proof of this theorem can be given by generalizing Cantor's construction, by which one builds the set of real numbers from the rational numbers.

Denote by $\tilde{\mathscr{M}}_s$ the family of all Cauchy sequences in $\mathscr{M}$. If $\xi' = \{\xi_1', \xi_2', \ldots\}$ and $\xi'' = \{\xi_1'', \xi_2'', \ldots\}$ are two such sequences, we say that they are equivalent if and only if

$$\lim_{n\to\infty} d(\xi_n', \xi_n'') = 0. \tag{3.2}$$

It is easy to see that we have thus introduced an equivalence relation in $\tilde{\mathscr{M}}_s$ (see Exercise 3.1) if we recall (see Exercises 1.6 and 2.8) the general definition of an equivalence relation.

Definition 3.7. A relation $\xi \sim \eta$ holding between any two ordered elements of a set S is called an *equivalence relation* if it is

(1) reflexive: $\xi \sim \xi$ for all $\xi \in S$;

(2) symmetric: $\xi \sim \eta$ implies that $\eta \sim \xi$;

(3) transitive: $\xi \sim \eta$ and $\eta \sim \zeta$ implies that $\xi \sim \zeta$.

A subset X of S having the property that all the elements of X are equivalent and that if $\eta \sim \xi$ and $\xi \in X$ then $\eta \in X$ is called an *equivalence class* (with respect to the equivalence relation $\sim$).

We denote the family of all equivalence classes in $\tilde{\mathscr{M}}_s$ [with respect to the equivalence relation given by (3.2)] by the symbol $\tilde{\mathscr{M}}$, and agree to denote the equivalence class containing the Cauchy sequence $\tilde{\xi}$ also by $\tilde{\xi}$. Consequently if $\tilde{\xi}', \tilde{\xi}'' \in \tilde{\mathscr{M}}$, then $\tilde{\xi}' = \tilde{\xi}''$ if and only if the Cauchy sequences $\tilde{\xi}', \tilde{\xi}'' \in \tilde{\mathscr{M}}_s$ are equivalent, i.e., satisfy (3.2).

We introduce the real function $d_s(\tilde{\xi}, \tilde{\eta})$ on $\tilde{\mathscr{M}}_s \times \tilde{\mathscr{M}}_s$ by defining for $\tilde{\xi} = \{\xi_1, \xi_2, ...\}$ and $\tilde{\eta} = \{\eta_1, \eta_2, ...\}$

$$d_s(\tilde{\xi}, \tilde{\eta}) = \lim_{n\to\infty} d(\xi_n, \eta_n). \tag{3.3}$$

In order to see that the above limit exists for any $\tilde{\xi}, \tilde{\eta} \in \tilde{\mathscr{M}}_s$ we employ the relation (see Exercise 3.2)

$$| d(\xi_m, \eta_m) - d(\xi_n, \eta_n)| \leqslant d(\xi_m, \xi_n) + d(\eta_m, \eta_n) \tag{3.4}$$

to show that $d(\xi_1, \eta_1), d(\xi_2, \eta_2), ...$ is a Cauchy sequence of numbers, and therefore has a limit; namely as $\xi_1, \xi_2, ...$ and $\eta_1, \eta_2, ...$ are Cauchy sequences, we can make $d(\xi_m, \xi_n) < \epsilon$ if $m, n > N_1(\epsilon)$, and $d(\eta_m, \eta_n) < \epsilon$ if $m, n > N_2(\epsilon)$, which, used in conjunction with (3.4), proves the statement.

We can show that $d_s(\tilde{\xi}, \tilde{\eta})$ also defines a real function on $\tilde{\mathscr{M}} \times \tilde{\mathscr{M}}$ by establishing that $d_s(\tilde{\xi}', \tilde{\eta}') = d_s(\tilde{\xi}'', \tilde{\eta}'')$ if $\tilde{\xi}' = \tilde{\xi}''$ and $\tilde{\eta}' = \tilde{\eta}''$ for $\tilde{\xi}', \tilde{\xi}'', \tilde{\eta}', \tilde{\eta}'' \in \tilde{\mathscr{M}}$. We first obtain that $d_s(\tilde{\xi}', \tilde{\eta}') = d_s(\tilde{\xi}'', \tilde{\eta}')$ from the inequality (see Exercise 3.3)

$$| d(\xi_n', \eta_n') - d(\xi_n'', \eta_n')| \leqslant d(\xi_n', \xi_n'')$$

because $d(\xi_n', \xi_n'') \to 0$ as $n \to \infty$ due to the fact that the Cauchy sequences $\tilde{\xi}'$ and $\tilde{\xi}''$ belong to the same equivalence class. Similarly we can show that $d_s(\tilde{\xi}'', \tilde{\eta}') = d_s(\tilde{\xi}'', \tilde{\eta}'')$, and thus prove that $d_s(\tilde{\xi}', \tilde{\eta}') = d_s(\tilde{\xi}'', \tilde{\eta}'')$.

It is easy to check that the function $d_s(\tilde{\xi}, \tilde{\eta})$ defines a metric on $\tilde{\mathscr{M}}$ (see Exercise 3.4). We show now that the ensuing metric space, which we denote also by $\tilde{\mathscr{M}}$, is complete.

Assume that $\tilde{\xi}^{(1)}, \tilde{\xi}^{(2)}, ...$ is a Cauchy sequence in $\tilde{\mathscr{M}}$, where $\tilde{\xi}^{(k)}$ is the equivalence class containing the Cauchy sequence $\{\xi_1^{(k)}, \xi_2^{(k)}, ...\}$ of elements of $\mathscr{M}$. Choose for each integer k an element $\eta_k = \xi_n^{(k)} \in \mathscr{M}$ such that $d(\xi_m^{(k)}, \eta_k) = d(\xi_m^{(k)}, \xi_n^{(k)}) < 1/k$ for all m greater than some N_k; this is certainly possible because $\xi_1^{(k)}, \xi_2^{(k)}, ...$ is a Cauchy sequence in $\mathscr{M}$.

Consider now the elements $\tilde{\eta}_k = \{\eta_k, \eta_k, ...\}$ and $\tilde{\xi}_m^{(k)} = \{\xi_m^{(k)}, \xi_m^{(k)}, ...\}$ of $\tilde{\mathscr{M}}_s$. We obviously have

$$d_s(\tilde{\xi}_m^{(k)}, \tilde{\eta}_k) = d(\xi_m^{(k)}, \eta_k) < 1/k.$$

If we let in the above relation $m \to \infty$, then we find that $d_s(\tilde{\xi}_m^{(k)}, \tilde{\eta}_k) \to d_s(\tilde{\xi}^{(k)}, \tilde{\eta}_k)$ since $\lim d_s(\tilde{\xi}_m^{(k)}, \tilde{\xi}^{(k)}) = 0$ as $m \to \infty$ (see Exercises 3.5 and 3.6) and consequently

$$d_s(\tilde{\xi}^{(k)}, \tilde{\eta}_k) \leqslant 1/k.$$

We can now deduce that $\tilde{\eta} = \{\eta_1, \eta_2, ...\}$ is a Cauchy sequence in $\mathscr{M}$ by writing

$$\begin{aligned} d(\eta_m, \eta_n) &= d_s(\tilde{\eta}_m, \tilde{\eta}_n) \\ &\leqslant d_s(\tilde{\eta}_m, \tilde{\xi}^{(m)}) + d_s(\tilde{\xi}^{(m)}, \tilde{\xi}^{(n)}) + d_s(\tilde{\xi}^{(n)}, \tilde{\eta}_m) \\ &\leqslant \frac{1}{m} + d_s(\tilde{\xi}^{(m)}, \tilde{\xi}^{(n)}) + \frac{1}{n}. \end{aligned} \tag{3.5}$$

Since $\tilde{\xi}^{(1)}, \tilde{\xi}^{(2)}, ...$ is a Cauchy sequence in $\tilde{\mathscr{M}}$, we can make $d_s(\tilde{\xi}^{(m)}, \tilde{\xi}^{(n)})$, and consequently the entire right-hand side of (3.5) arbitrarily small for all sufficiently large m and n. Thus, $\tilde{\eta} \in \tilde{\mathscr{M}}_s$.

We can establish that the equivalence class $\tilde{\eta} \in \tilde{\mathscr{M}}$ containing the Cauchy sequence $\tilde{\eta} = \{\eta_1, \eta_2, ...\}$ is the limit of $\tilde{\xi}^{(1)}, \tilde{\xi}^{(2)}, ...$ if we write

$$d_s(\tilde{\eta}, \tilde{\xi}^{(k)}) \leqslant d_s(\tilde{\eta}, \tilde{\eta}_k) + d_s(\tilde{\eta}_k, \tilde{\xi}^{(k)}). \tag{3.6}$$

The right-hand side of (3.6) can be made arbitrarily small for sufficiently large k because $d_s(\tilde{\eta}_k, \tilde{\xi}^{(k)}) \leqslant 1/k$ and $\lim_{k\to\infty} d_s(\tilde{\eta}, \tilde{\eta}_k) = 0$ (see Exercise 3.6).

In order to finish the proof of the theorem, we have to embed $\mathscr{M}$ into the complete metric space $\tilde{\mathscr{M}}$. To that purpose we map $\xi \in \mathscr{M}$ into the equivalence class $\tilde{\xi}$ containing the sequence $\{\xi, \xi, ...\}$. This mapping is obviously one-to-one and isometric, as $d(\xi, \eta) = d_s(\tilde{\xi}, \tilde{\eta})$. Furthermore, the image $\mathscr{M}'$ of $\mathscr{M}$ in $\tilde{\mathscr{M}}$ is everywhere dense in $\tilde{\mathscr{M}}$; namely if $\tilde{\eta} \in \tilde{\mathscr{M}}$ contains $\{\eta_1, \eta_2, ...\} \in \tilde{\mathscr{M}}_s$, then for arbitrary $\epsilon > 0$ we can choose an $\tilde{\eta}_k$ in $\mathscr{M}'$ containing $\{\eta_k, \eta_k, ...\}$ and such that $d_s(\tilde{\eta}, \tilde{\eta}_k) < \epsilon$.

Exercises

3.1. Show that the relation $\tilde{\xi} \sim \tilde{\eta}$ between any two Cauchy sequences $\tilde{\xi} = \{\xi_1, \xi_2, ...\}$ and $\tilde{\eta} = \{\eta_1, \eta_2, ...\}$ of a metric space $\mathscr{M}$, defined to mean that $\lim_{n\to\infty} d(\xi_n, \eta_n) = 0$, satisfies the three requirements given in Definition 3.7 for an equivalence relation.

3.2. Prove that any four elements $\xi_1, \xi_2, \eta_1, \eta_2$ of a metric space $\mathscr{M}$ satisfy the relation

$$|d(\xi_1, \xi_2) - d(\eta_1, \eta_2)| \leqslant d(\xi_1, \eta_1) + d(\xi_2, \eta_2).$$

3.3. Prove that if ξ, η, ζ are elements of a metric space $\mathscr{M}$, then

$$|d(\xi, \eta) - d(\xi, \zeta)| \leqslant d(\eta, \zeta).$$

3.4. Show that the function $d_s(\tilde{\xi}, \tilde{\eta})$ defined on $\tilde{\mathscr{M}} \times \tilde{\mathscr{M}}$ by (3.3) satisfies the four requirements for a metric (those requirements are formulated in Definition 3.1).

3.5. If in a metric space $\mathscr{M}$ the sequence $\xi_1, \xi_2, \ldots$ converges to ξ, prove that for any $\eta \in \mathscr{M}$, $\lim_{n\to\infty} d(\xi_n, \eta) = d(\xi, \eta)$.

3.6. If $\tilde{\xi}$ is the equivalence class of $\tilde{\mathscr{M}}$ (introduced in Theorem 3.2) containing the Cauchy sequence $\{\xi_1, \xi_2, \ldots\} \in \tilde{\mathscr{M}}_s$, then for any $\epsilon > 0$ there is an $N(\epsilon)$ such that $d_s(\tilde{\xi}, \tilde{\xi}_k) < \epsilon$ for all $k > N(\epsilon)$, where $\tilde{\xi}_k = \{\xi_k, \xi_k, \ldots\}$. Prove this statement!

3.7. Show that if S_1 is an everywhere dense subset of a metric space $\mathscr{M}$, and S_2 is an everywhere dense subset of S_1, then S_2 is everywhere dense in $\mathscr{M}$.

4. Hilbert Space

4.1. Completion of a Euclidean Space

It is easy to establish (see Exercise 4.1) that in normed space $\mathscr{N}$

$$d(f, g) = \|f - g\| \tag{4.1}$$

is a metric. Therefore, we can define in $\mathscr{N}$ convergence, completeness, etc. in the metric (4.1), which is then called *convergence, completeness*, etc. *in the norm*. A complete normed space bears the name of *Banach space*.

The above concepts can also be applied to Euclidean spaces, because according to Theorem 2.3 we can introduce in such spaces a norm, and therefore also a metric. A Euclidean space which is complete in the norm* is called a *Hilbert space*.

Not every Euclidean space is a Hilbert space. For instance, the Euclidean space $\mathscr{C}^0_{(2)}(\mathbb{R}^1)$ introduced in §2 is not complete. To see this, note that the sequence $f_1, f_2, \ldots$ of continuous functions

$$f_n(x) = \begin{cases} 1 & \text{for } |x| \leqslant a \\ \exp[-n^2(|x| - a)^2] & \text{for } |x| > a, \end{cases} \tag{4.2}$$

is a Cauchy sequence in $\mathscr{C}^0_{(2)}(\mathbb{R}^1)$ (see Exercise 4.2) but it does not converge to an element of $\mathscr{C}^0_{(2)}(\mathbb{R}^1)$. In fact, it is easy to establish that with increasing n, the functions in the above sequence approximate arbitrarily closely in norm the discontinuous step function

$$\chi(x) = \begin{cases} 1 & \text{for } |x| \leqslant a \\ 0 & \text{for } |x| > a, \end{cases}$$

which, however, does not belong to $\mathscr{C}^0_{(2)}(\mathbb{R}^1)$.

Definition 4.1. We say that the Euclidean space $\mathscr{E}$ can be *densely embedded* in the Hilbert space $\mathscr{H}$ if there is a one-to-one mapping of $\mathscr{E}$ into $\mathscr{H}$, such that the image $\mathscr{E}'$ of $\mathscr{E}$ is everywhere dense in $\mathscr{H}$, and the mapping represents an isomorphism between the Euclidean spaces $\mathscr{E}$ and $\mathscr{E}'$.

Theorem 4.1. Any incomplete Euclidean space $\mathscr{E}$ can be densely embedded in a Hilbert space.

Proof. Denote by $\tilde{\mathscr{E}}$ the complete metric space built from the set $\tilde{\mathscr{E}}_s$ of Cauchy sequences in $\mathscr{E}$ according to the procedure used in proving Theorem 3.2. Define in $\tilde{\mathscr{E}}_s$ the operations

$$\begin{aligned} \tilde{f} + \tilde{g} &= \{f_1 + g_1, f_2 + g_2, \ldots\}, \\ a\tilde{f} &= \{af_1, af_2, \ldots\} \end{aligned} \tag{4.3}$$

for any two sequences $\tilde{f} = \{f_1, f_2, \ldots\}$, $\tilde{g} = \{g_1, g_2, \ldots\}$ from $\tilde{\mathscr{E}}_s$. It is easy to check that the above operations are operations of vector addition and multiplication by scalar. Furthermore, if $\tilde{f}' = \tilde{f}''$, where $\tilde{f}' = \{f_1', f_2', \ldots\}$ and $\tilde{f}'' = \{f_1'', f_2'', \ldots\}$, i.e., if $\tilde{f}'$ and $\tilde{f}''$ belong to the same equivalence class in $\tilde{\mathscr{E}}$ and therefore

$$\lim_{n\to\infty} \|f_n' - f_n''\| = \lim_{n\to\infty} d(f_n', f_n'') = 0,$$

then $\|af_n' - af_n''\| = |a| \, \|f_n' - f_n''\| \to 0$; thus, we also have that $\tilde{f}' + \tilde{g} = \tilde{f}'' + \tilde{g}$ and $a\tilde{f}' = a\tilde{f}''$. Consequently, (4.3) defines vector operations on $\tilde{\mathscr{E}}$, which thus becomes a vector space.

We now introduce the complex function on $\tilde{\mathscr{E}}_s \times \tilde{\mathscr{E}}_s$ defined by

$$\langle \tilde{f} | \tilde{g} \rangle_s = \lim_{n\to\infty} \langle f_n | g_n \rangle. \tag{4.4}$$

* The concept of completeness can be defined and considered for other topologies besides the norm topology.

The limit in (4.4) exists because $\langle f_1 \mid g_1 \rangle$, $\langle f_2 \mid g_2 \rangle$,... is a Cauchy sequence of numbers, as can be seen from the following inequality:

$$|\langle f_m \mid g_m \rangle - \langle f_n \mid g_n \rangle| = |\langle f_m - f_n \mid g_m \rangle + \langle f_n \mid g_m - g_n \rangle| \\ \leqslant \| f_m - f_n \| \| g_m \| + \| f_n \| \| g_m - g_n \|;$$

namely, $\| f_n \| \to d_s(\tilde{f}, 0) = \| \tilde{f} \|_s$ and $\| g_n \| \to \| \tilde{g} \|_s$ for $n \to \infty$, where

$$\| \tilde{f} \|_s = \sqrt{\langle \tilde{f} \mid \tilde{f} \rangle_s},$$

while $\| f_m - f_n \|$ and $\| g_m - g_n \|$ can be made arbitrarily small for sufficiently large m and n.

Furthermore, if $\tilde{f}' = \tilde{f}'' \in \tilde{\mathscr{E}}_s$, then $\langle \tilde{f}' \mid \tilde{g} \rangle_s = \langle \tilde{f}'' \mid \tilde{g} \rangle_s$, as can be concluded from the inequality

$$|\langle f_n' \mid g_n \rangle - \langle f_n'' \mid g_n \rangle| \leqslant \| f_n' - f_n'' \| \| g_n \|.$$

Hence, $\langle \tilde{f} \mid \tilde{g} \rangle_s$ is a uniquely defined function on $\tilde{\mathscr{E}} \times \tilde{\mathscr{E}}$. This function determines an inner product on $\tilde{\mathscr{E}}$ (see Exercise 4.4). It is obvious that the mapping $f \leftrightarrow \tilde{f} = \{f, f, ...\}$ of $\mathscr{E}$ into $\tilde{\mathscr{E}}$ has an image $\mathscr{E}'$ which is a linear subspace of $\tilde{\mathscr{E}}$; according to the construction, $\mathscr{E}'$ is everywhere dense in $\tilde{\mathscr{E}}$ and the above mapping provides an isomorphism between $\mathscr{E}$ and $\mathscr{E}'$. Q.E.D.

A similar theorem can be proved for normed spaces (see Exercise 4.5).

4.2. Separable Hilbert Spaces

In quantum mechanics we deal at present almost exclusively with a special class of Hilbert spaces which are called separable.

Definition 4.2. The Euclidean space $\mathscr{E}$ is called *separable* if there is a *countable* everywhere dense subset of vectors of $\mathscr{E}$.

In the early days of research on Hilbert spaces, separability was taken to be an integral part of the concept of a Hilbert space.

In quantum mechanics we are concerned primarily with separable complex Hilbert spaces. We shall agree that in the future whenever we refer to a space as a Hilbert space we mean a complex Hilbert space, except if otherwise stated.

Theorem 4.2. Every subspace of a separable Euclidean space is a separable Euclidean space.

Proof. The fact that a subspace $\mathscr{E}_1$ of a Euclidean space $\mathscr{E}$ is also a Euclidean space is easy to check (see Exercise 2.6). In order to establish

the separability of $\mathscr{E}_1$, construct a countable subset $S = \{g_{11}, g_{12}, g_{21}, g_{13}, ...\}$ of $\mathscr{E}_1$ in the following way.

Let $R = \{f_1, f_2, ...\}$ be a countable everywhere dense subset of $\mathscr{E}$; there has to be such a set because of the separability of $\mathscr{E}$. Let g_{mn} denote a vector of $\mathscr{E}_1$ satisfying $\| g_{mn} - f_n \| < 1/m$, in case there is at least one such vector, or the zero vector in case there is no vector of $\mathscr{E}_1$ in the $1/m$ neighborhood of f_n.

The set S is everywhere dense in $\mathscr{E}_1$ because for any given $h \in \mathscr{E}_1$ and $m > 0$ we can find an $f_n \in R$ such that $\| h - f_n \| < 1/m$. Thus, by the above rule of constructing S we certainly have $g_{mn} \neq \mathbf{0}$ and therefore

$$\| h - g_{mn} \| \leqslant \| h - f_n \| + \| f_n - g_{nm} \| < 2/m.$$

This proves that S is everywhere dense in $\mathscr{E}_1$. Q.E.D.

4.3. l^2 Spaces as Examples of Separable Hilbert Spaces

As an important example of an infinite-dimensional separable Hilbert space we give the space $l^2(\infty)$, which is basic in matrix mechanics.

Theorem 4.3. The set $l^2(\infty)$ of all one-column complex matrices α with a countable number of elements

$$\alpha = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix}$$

for which

$$\sum_{k=1}^{\infty} | a_k |^2 < +\infty \tag{4.5}$$

becomes a separable Hilbert space, denoted by $l^2(\infty)$, if the vector operations are defined by

$$\alpha + \beta = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \end{pmatrix}, \tag{4.6}$$

$$a\alpha = \begin{pmatrix} aa_1 \\ aa_2 \\ \vdots \end{pmatrix}, \qquad a \in \mathbb{C}^1, \tag{4.7}$$

and the inner product by

$$\langle \alpha \mid \beta \rangle = \sum_{k=1}^{\infty} a_k^* b_k. \tag{4.8}$$

Proof. The operation (4.7) maps $\mathbb{C}^1 \times l^2(\infty)$ into $l^2(\infty)$ because $\sum_{k=1}^{\infty} | aa_k |^2 = | a |^2 \sum_{k=1}^{\infty} | a_k |^2 < +\infty$ if (4.5) is satisfied.

In order to see that (4.6) maps $l^2(\infty) \times l^2(\infty)$ into $l^2(\infty)$, apply the triangle inequality on the ν-dimensional space $l^2(\nu)$, $\nu < +\infty$, in order to obtain

$$\left[\sum_{k=1}^{\nu} | a_k + b_k |^2\right]^{1/2} \leqslant \left[\sum_{k=1}^{\nu} | a_k |^2\right]^{1/2} + \left[\sum_{k=l}^{\nu} | b_k |^2\right]^{1/2}.$$

The above inequality shows that when $\nu \to \infty$, the left-hand side converges if $\alpha, \beta \in l^2(\infty)$.

Similarly, we prove that (4.8) converges absolutely by applying the Schwarz–Cauchy inequality on the ν-dimensional space $l^2(\nu)$, $\nu < +\infty$, in order to obtain

$$\sum_{k=1}^{\nu} | a_k{}^* b_k | = \sum_{k=1}^{\nu} | a_k | \, | b_k | \leqslant \left[\sum_{k=1}^{\nu} | a_k |^2\right]^{1/2} \left[\sum_{k=1}^{\nu} | b_k |^2\right]^{1/2}.$$

We leave as an exercise for the reader (see Exercise 4.6) to check that we deal indeed with a Euclidean space.

To prove that this Euclidean space is complete, assume that $\alpha^{(1)}, \alpha^{(2)},\ldots$ is a Cauchy sequence, where

$$\alpha^{(n)} = \begin{pmatrix} a_1^{(n)} \\ a_2^{(n)} \\ \vdots \end{pmatrix}.$$

As we obviously have that for any $k = 1, 2,\ldots$

$$| a_k^{(m)} - a_k^{(n)} | \leqslant \| \alpha^{(m)} - \alpha^{(n)} \|,$$

we deduce that for fixed k the sequence $a_k^{(1)}, a_k^{(2)},\ldots$ is a Cauchy sequence of complex numbers; hence, this sequence has a limit b_k. We shall prove that the one-column infinite matrix

$$\beta = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \end{pmatrix}$$

is an element of $l^2(\infty)$, and that $\alpha^{(1)}, \alpha^{(2)},\ldots$ converges in the norm to β.

By applying again the triangle inequality on the ν-dimensional space $l^2(\nu)$, $\nu < +\infty$, we obtain

(4.9)
$$\left[\sum_{k=1}^{\nu} | b_k - a_k^{(n)} |^2\right]^{1/2} \leqslant \left[\sum_{k=1}^{\nu} | b_k - a_k^{(m)} |^2\right]^{1/2} + \left[\sum_{k=1}^{\nu} | a_k^{(m)} - a_k^{(n)} |^2\right]^{1/2};$$

the above inequality is true for any $m = 1, 2,\ldots$. As $\alpha^{(1)}, \alpha^{(2)},\ldots$ is a Cauchy sequence, there is for any given $\epsilon > 0$ an $N_0(\epsilon)$ such that for any $m, n > N_0(\epsilon)$ and any positive integer ν

$$\sum_{k=1}^{\nu} | a_k^{(m)} - a_k^{(n)} |^2 \leqslant \| \alpha^{(m)} - \alpha^{(n)} \|^2 < \epsilon^2/4.$$

On the other hand, as $b_k = \lim_{m\to\infty} a_k^{(m)}$, we can find for fixed ν an $N_\nu(\epsilon)$ such that $| b_k - a_k^{(m)} | < \epsilon/2^{(k+1)/2}$ for any $m > N_\nu(\epsilon)$ and all $k = 1, 2,\ldots, \nu$. Thus, we get from (4.9) that for all $n > N_0(\epsilon)$ and all positive integers ν

(4.10)
$$\left[\sum_{k=1}^{\nu} | b_k - a_k^{(n)} |^2\right]^{1/2} \leqslant \epsilon \left(\sum_{k=1}^{\nu} \frac{1}{2^{k+1}}\right) + \frac{\epsilon}{2}$$
$$\leqslant \frac{\epsilon}{2}\left(\sum_{k=0}^{\infty} \frac{1}{2^k}\right) + \frac{\epsilon}{2} = \epsilon.$$

As the right-hand side of the above inequality is independent of ν and the inequality itself is true for any $n > N_0(\epsilon)$, we can let $\nu \to \infty$ in (4.10) to derive

(4.11)
$$\left[\sum_{k=1}^{\infty} | b_k - a_k^{(n)} |^2\right]^{1/2} \leqslant \epsilon \qquad \text{for all} \quad n > N_0(\epsilon).$$

By returning again to $l^2(\nu)$, $\nu < +\infty$, to obtain

$$\left[\sum_{k=1}^{\nu} | b_k |^2\right]^{1/2} \leqslant \left[\sum_{k=1}^{\nu} | b_k - a_k^{(n)} |^2\right]^{1/2} + \left[\sum_{k=1}^{\nu} | a_k^{(n)} |^2\right]^{1/2}$$

we establish that $\beta \in l^2(\infty)$ by letting $\nu \to \infty$. The relation (4.11) tells us now that $\alpha^{(1)}, \alpha^{(2)},\ldots$ converges to β.

Finally, in order to prove the separability of $l^2(\infty)$, consider the set D of all the one-column matrices α from $l^2(+\infty)$ with kth components $(\alpha)_k = a_k$, where a_k has rational numbers as its real and imaginary part, i.e., $\operatorname{Re} a_k, \operatorname{Im} a_k \in \mathscr{R}$, $k = 1, 2,\ldots$, and in addition

(4.12)
$$a_{n+1} = a_{n+2} = \cdots = 0$$

for some integer n. The set D is countable (see Exercise 4.7). In order to prove that D is everywhere dense in $l^2(\infty)$, take any one-column matrix $\gamma \in l^2(+\infty)$ with kth component $(\gamma)_k = c_k$. As $\gamma \in l^2(\infty)$, there is for any given $\epsilon > 0$ an integer n such that

$$\sum_{k=n+1}^{\infty} |c_k|^2 < \epsilon^2/2.$$

Furthermore, as the set $\mathscr{R}$ of rational numbers is dense in the set $\mathbb{R}^1$ of real numbers, we can choose an $\alpha \in D$ which satisfies (4.12) and is such that $|c_k - a_k| < \epsilon/\sqrt{2n}$ for all $k = 1,..., n$. Thus, we have

$$\|\gamma - \alpha\| = \left[\sum_{k=1}^{n} |c_k - a_k|^2 + \sum_{k=n+1}^{\infty} |c_k|^2\right]^{1/2} < \epsilon,$$

which proves that $l^2(\infty)$ is separable. Q.E.D.

4.4. Orthonormal Bases in Hilbert Space

In an infinite-dimensional Euclidean space it is important to distinguish between the vector space (S) spanned by a set S, and the *closed* vector space $[S]$ spanned by S.

Definition 4.3. The *vector space* (*or linear manifold*) (S) *spanned by the subset* S of a Euclidean space $\mathscr{E}$ is the smallest* subspace of $\mathscr{E}$ containing S. The *closed vector subspace* $[S]$ *spanned by* S is the smallest closed vector subspace of $\mathscr{E}$ containing S.

In the finite-dimensional case $(S) = [S]$ because all finite-dimensional Euclidean spaces are closed (see Exercise 4.8). That this is not so in the infinite-dimensional case can be deduced from the following theorem, whose simple proof we leave to the reader (see Exercise 4.9).

Theorem 4.4. The subspace (S) of the Euclidean space $\mathscr{E}$ spanned by the set S is identical with the set of all finite linear combinations $a_1 f_1 + \cdots + a_n f_n$ of vectors from S, i.e., in customary set-theoretical notation,

$$(S) = \{a_1 f_1 + \cdots + a_n f_n\colon\ f_1,..., f_n \in S,\quad a_1,..., a_n \in \mathbb{C}^1,\quad n = 1, 2,...\}.$$

The closed linear subspace $[S]$ spanned by S is identical to the closure $\overline{(S)}$ of (S).

Definition 4.4. An orthonormal system S of vectors in a Euclidean space $\mathscr{E}$ is called an *orthonormal basis* (or a *complete orthonormal system*) in the Euclidean space $\mathscr{E}$ if the closed linear space $[S]$ spanned by S is identical to the entire Euclidean space, i.e., $[S] = \mathscr{E}$.

* That is, if $\mathscr{V}$ is a subspace of $\mathscr{E}$ and $S \subset \mathscr{V}$, then necessarily $(S) \subset \mathscr{V}$.

It must be realized that an orthonormal basis T in an infinite-dimensional Euclidean space $\mathscr{E}$ is not a *vector basis* for the vector space $\mathscr{E}$, i.e., (T) is in general different from $\mathscr{E}$. For instance, this is so with the basis $e_1, e_2,...$ of $l^2(\infty)$, where e_m is the vector whose nth matrix component is $(e_m)_n = \delta_{mn}$.

Theorem 4.5. A Euclidean space $\mathscr{E}$ is separable if and only if there is a *countable* orthonormal basis in $\mathscr{E}$.

Proof. (a) To prove that in a separable Euclidean space $\mathscr{E}$ there is a countable orthonormal basis, note that due to the separability of $\mathscr{E}$ there is a countable set $S = \{f_1, f_2,...\}$ which is everywhere dense in $\mathscr{E}$. According to Theorem 2.4, there is a countable orthonormal system $T = \{e_1, e_2,...\}$ such that $(S) = (T)$. Due to Theorem 4.4 we have then

$$[T] = \overline{(T)} = \overline{(S)} = [S] = \mathscr{E}.$$

(b) Conversely, to show that if there is a countable orthonormal basis $T = \{e_1, e_2,...\}$, then $\mathscr{E}$ is separable, consider the set

$$R = \{r_1 e_1 + \cdots + r_n e_n\colon\ \operatorname{Re} r_1,..., \operatorname{Re} r_n,\quad \operatorname{Im} r_1,..., \operatorname{Im} r_n \in R,\quad n = 1, 2,...\}$$

which is countable, as can be established by using the technique for solving Exercise 4.7. The set R is also everywhere dense in $\mathscr{E}$; namely, if $f \in \mathscr{E}$ and $\epsilon > 0$ is given, then as $[e_1, e_2,...] = \mathscr{E}$, there is a vector $g = a_1 e_1 + \cdots + a_n e_n$ such that

$$\|f - a_1 e_1 - \cdots - a_n e_n\| < \epsilon/2.$$

Furthermore, we can choose complex numbers $r_1,..., r_n$ with rational real and imaginary parts so that

$$|r_k - a_k| < \epsilon/2\sqrt{n}, \qquad k = 1,..., n.$$

Thus, we have that for $h = r_1 e_1 + \cdots + r_n e_n \in R$

$$\|f - h\| \leqslant \|f - g\| + \|g - h\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \quad \text{Q.E.D.}$$

There are a few very important criteria by means of which we can establish whether an orthonormal system S is a basis in a Euclidean space.

Theorem 4.6. Each of the following statements is a *sufficient and necessary* condition that the countable orthonormal system $T = \{e_1, e_2,...\}$ is a basis in the separable Hilbert space* $\mathscr{H}$:

* The theorem applies to the finite-dimensional as well as the infinite-dimensional case, though it is stated and proved here for infinite-dimensional $\mathscr{H}$. In the finite-dimensional case, ∞ should be replaced by the dimension of $\mathscr{H}$.

Proof. (a) The only vector f satisfying the relations

$$\langle e_k | f \rangle = 0, \qquad k = 1, 2, \ldots \tag{4.13}$$

is the zero vector, i.e., (4.13) implies $f = \mathbf{0}$.

(b) For any vector $f \in \mathscr{H}$,

$$\lim_{n \to +\infty} \left\| f - \sum_{k=1}^{n} \langle e_k | f \rangle e_k \right\| = 0, \tag{4.14}$$

or symbolically written

$$f = \sum_{k=1}^{\infty} \langle e_k | f \rangle e_k ,$$

where $\langle e_k | f \rangle$ is sometimes called the *Fourier coefficient* of f.

(c) Any two vectors $f, g \in \mathscr{H}$ satisfy *Parseval's relation*:

$$\langle f | g \rangle = \sum_{k=1}^{\infty} \langle f | e_k \rangle \langle e_k | g \rangle. \tag{4.15}$$

(d) For any $f \in \mathscr{H}$

$$\| f \|^2 = \sum_{k=1}^{\infty} | \langle e_k | f \rangle |^2. \tag{4.16}$$

We start by proving that the criteria (a) and (b) are equivalent to the requirement that $T = \{e_1, e_2, \ldots\}$ is an orthonormal basis, as that requirement was formulated in Definition 4.4. To do that we shall prove that (a) implies (b), (b) implies "T is a basis" (as formulated in Definition 4.4) and "T is a basis" implies (a).

In order to show that (a) implies (b) we need the following lemma.

Lemma 4.1. For any given vector f of a Euclidean space $\mathscr{E}$ (not necessarily separable) and any countable orthonormal system $\{e_1, e_2, \ldots\}$ in $\mathscr{E}$, the sequence $f_1, f_2, \ldots$ of vectors

$$f_n = \sum_{k=1}^{n} \langle e_k | f \rangle e_k \tag{4.17}$$

is a Cauchy sequence, and the Fourier coefficients $\langle e_k | f \rangle$ satisfy Bessel's inequality

$$\| f_n \|^2 = \sum_{k=1}^{n} | \langle e_k | f \rangle |^2 \leqslant \| f \|^2. \tag{4.18}$$

Proof. Write

$$h_n = f - f_n ,$$

where f_n is given by (4.17). We have

$$\langle f_n | h_n \rangle = 0$$

because $\langle e_i | h_n \rangle = 0$ for $i = 1, 2, \ldots, n$,

$$\langle e_i | h_n \rangle = \left\langle e_i \middle| f - \sum_{j=1}^{n} \langle e_j | f \rangle e_j \right\rangle$$

$$= \langle e_i | f \rangle - \sum_{j=1}^{n} \langle e_j | f \rangle \langle e_i | e_j \rangle = 0,$$

as $\langle e_i | e_j \rangle = \delta_{ij}$. Thus

$$\langle f | f \rangle = \langle f_n + h_n | f_n + h_n \rangle = \langle f_n | f_n \rangle + \langle h_n | h_n \rangle,$$

and consequently, since $\langle h_n | h_n \rangle \geqslant 0$,

$$\langle f_n | f_n \rangle \leqslant \langle f | f \rangle. \tag{4.19}$$

By using (4.17) and $\langle e_i | e_j \rangle = \delta_{ij}$ we derive

$$\| f_n \|^2 = \langle f_n | f_n \rangle = \sum_{i,j=1}^{n} \langle e_i | f \rangle^* \langle e_i | e_j \rangle \langle e_j | f \rangle$$

$$= \sum_{i=1}^{n} | \langle e_i | f \rangle |^2,$$

which shows in conjunction with (4.19) that Bessel's inequality (4.18) is true. From (4.18) we can deduce that

$$\sum_{i=1}^{\infty} | \langle e_i | f \rangle |^2 \leqslant \| f \|^2 < +\infty, \tag{4.20}$$

i.e., the above series with nonnegative terms is bounded and therefore it converges. Since we know that for $m > n$

$$\| f_m - f_n \|^2 = \sum_{i=n+1}^{m} | \langle e_i | f \rangle |^2, \tag{4.21}$$

we easily see from (4.20) and (4.21) that $f_1, f_2, \ldots$ is a Cauchy sequence.

We return now to proving Theorem 4.6. If $T = \{e_1, e_2, \ldots\}$ is a countable orthonormal system in the Hilbert space $\mathscr{H}$, then according to Lemma 4.1 for any given $f \in \mathscr{H}$ the sequence $f_1, f_2, \ldots$, where

$$f_n = \sum_{k=1}^{n} \langle e_k | f \rangle e_k$$

is a Cauchy sequence. Since $\mathscr{H}$ is complete, this sequence has a limit $g \in \mathscr{H}$.

We can now show that if the statement (a) about T in Theorem 4.6 is true, then (b) is also true, due to the fact that (a) implies $f = g$; namely, for any $k = 1, 2, \ldots$ we have

$$\langle f - g \mid e_k \rangle = \lim_{n \to \infty} \langle f - f_n \mid e_k \rangle$$

$$= \langle f \mid e_k \rangle - \lim_{n \to +\infty} \sum_{i=1}^{n} \langle e_i | f \rangle^* \langle e_i \mid e_k \rangle = 0.$$

Thus, if (a) is true, we must have $f - g = \mathbf{0}$.

It is obvious that statement (b) implies that $T = \{e_1, e_2, \ldots\}$ is a basis, because according to Theorem 4.4 any $f \in \mathscr{H}$ is the limit of elements $f_1, f_2, \ldots$ from the linear space (T) spanned by T, where $f_n \in (T)$ is of the form (4.17).

We show now that the fact that $T = \{e_1, e_2, \ldots\}$ is an orthonormal basis implies that (a) is true. Assume that some $f \in \mathscr{H}$ is orthogonal on the system $\{e_1, e_2, \ldots\}$. Since $f \in [e_1, e_2, \ldots] = \mathscr{H}$, there is a sequence $g_1, g_2, \ldots \in (e_1, e_2, \ldots)$, i.e., for some integer s_n

$$g_n = \sum_{k=1}^{s_n} a_k e_k,$$

which converges to f. Consequently, as $\langle f \mid e_k \rangle = 0$,

$$\langle f | f \rangle = \lim_{n \to \infty} \langle f \mid g_n \rangle = \lim_{n \to \infty} \sum_{k=1}^{s_n} a_k \langle f \mid e_k \rangle = 0,$$

and therefore $f = \mathbf{0}$.

We shall demonstrate that statement (c) is equivalent to (a) or (b) by showing that (b) implies (c), and (c) implies (a), and thus finish the proof of Theorem 4.6.

If (b) is true, then we have (see Exercise 4.10)

$$\langle f \mid g \rangle = \lim_{n \to \infty} \langle f_n \mid g_n \rangle, \tag{4.22}$$

where

$$f_n = \sum_{k=1}^{n} \langle e_k | f \rangle e_k, \qquad g_n = \sum_{k=1}^{n} \langle e_k \mid g \rangle e_k.$$

From the relation

$$\langle f_n \mid g_n \rangle = \sum_{i,j=1}^{n} \langle e_i | f \rangle^* \langle e_i \mid e_j \rangle \langle e_j \mid g \rangle = \sum_{i=1}^{n} \langle f \mid e_i \rangle \langle e_i \mid g \rangle$$

we immediately obtain Parseval's relation (4.15).

If we assume (c) to be true, then (a) is also true, because if some vector f is orthogonal on $\{e_1, e_2, \ldots\}$, i.e., $\langle f \mid e_k \rangle = 0$, $k = 1, 2, \ldots$, then by inserting $f = g$ in (4.15) we get

$$\langle f | f \rangle = \sum_{k=1}^{\infty} \langle f \mid e_k \rangle \langle e_k | f \rangle = 0,$$

which implies that $f = \mathbf{0}$.

Finally, (d) follows from (c) by taking again in (4.15) that $f = g$. Vice versa, if (d) is true then (a) has to be true, because if $\langle f \mid e_k \rangle = 0$ for $k = 1, 2, \ldots$, then we get from (4.16) that $\|f\|^2 = 0$, which implies that $f = \mathbf{0}$. Q.E.D.

It is easy to see that, due to the fact that every Euclidean space can be embedded in a Hilbert space (Theorem 4.1), the criteria (b) (c), and (d) are also necessary and sufficient criteria for T to be an orthonormal basis in a Euclidean space in general; while (a) is necessary but not sufficient (see Exercise 4.13).

4.5. Isomorphism of Separable Hilbert Spaces

We can now demonstrate for infinite-dimensional Hilbert spaces a theorem analogous to the Theorem 2.5 for finite-dimensional Hilbert spaces.

Theorem 4.7. All complex infinite-dimensional separable Hilbert spaces are isomorphic to $l^2(\infty)$, and consequently mutually isomorphic.

Proof. If $\mathscr{H}$ is separable, there is, according to Theorem 4.5, an orthonormal countable basis $\{e_1, e_2, \ldots\}$ in $\mathscr{H}$, which is infinite when $\mathscr{H}$ is infinite dimensional. According to Theorem 4.6 we can write for any $f \subset \mathscr{H}$

$$f = \sum_{k=1}^{\infty} c_k e_k, \qquad c_k = \langle e_k | f \rangle,$$

where by (4.16)

$$\sum_{k=1}^{\infty} |c_k|^2 = \|f\|^2 < +\infty.$$

Therefore

$$\alpha_f = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \end{pmatrix} \in l^2(\infty).$$

Vice versa, if

$$\beta = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \end{pmatrix} \in l^2(\infty),$$

then $f_1, f_2, \ldots$

$$f_n = \sum_{k=1}^{n} b_k e_k$$

is a Cauchy sequence, because for any $m > n$

$$\|f_m - f_n\|^2 = \sum_{k=n+1}^{m} |b_k|^2$$

and $\sum_{k=1}^{\infty} |b_k|^2$ converges. Thus, due to the completeness of $\mathscr{H}$ $f_1, f_2, \ldots$ converges to a vector $f \in \mathscr{H}$ and we have

$$c_k = \langle e_k | f \rangle = \lim_{n\to\infty} \left\langle e_k \,\middle|\, \sum_{i=1}^{n} b_i e_i \right\rangle = b_k .$$

Therefore, the inverse mapping of the mapping $f \to \alpha_f$ of $\mathscr{H}$ into $l^2(\infty)$ exists, and has $l^2(\infty)$ as its domain of definition. Hence the mapping $f \to \alpha_f$ is a one-to-one mapping of $\mathscr{H}$ *onto* $l^2(\infty)$. It can be easily checked (see Exercise 4.11) that this mapping supplies an isomorphism between $\mathscr{H}$ and $l^2(\infty)$. Q.E.D.

As we shall see later, the above theorem provides the basis of the equivalence of Heisenberg's matrix formulation and Schroedinger's wave formulation of quantum mechanics.

Theorem 4.8. If the mapping

$$f \to f', \qquad f \in \mathscr{E}, \qquad f' \in \mathscr{E}'$$

is a unitary transformation of the separable Euclidean space into the Euclidean space $\mathscr{E}'$, and if $\{e_1, e_2, \ldots\}$ is an orthonormal basis in $\mathscr{E}$,

then $\{e_1', e_2', \ldots\}$ is an orthonormal basis in $\mathscr{E}'$, where e_n' denotes the image of e_n.

Proof. Let $\mathscr{E}$ be infinite dimensional, and denote by $\langle \cdot | \cdot \rangle_1$ and $\langle \cdot | \cdot \rangle_2$ the inner products in $\mathscr{E}$ and $\mathscr{E}'$ respectively. Then

$$\langle e_i' | e_j' \rangle_2 = \langle e_i | e_j \rangle_1 = \delta_{ij},$$

i.e., $\{e_1', e_2', \ldots\}$ is an orthonormal system in $\mathscr{E}'$. Since each $f' \in \mathscr{E}'$ has a unique inverse image $f \in \mathscr{E}$, we have

$$\lim_{n\to\infty} \left\| f' - \sum_{k=1}^{n} \langle e_n' | f' \rangle_2 e_n' \right\|_2 = \lim_{n\to\infty} \left\| f - \sum_{k=1}^{n} \langle e_n | f \rangle_1 e_k \right\|_1 = 0,$$

which by Theorem 4.6(b) proves that $\{e_1', e_2', \ldots\}$ is a basis.

The case when $\mathscr{E}$ is finite dimensional can be treated in a similar manner. Q.E.D.

Exercises

4.1. Show that in a normed space $\mathscr{N}$ the real function $d(f, g) = \|f - g\|$ on $\mathscr{N} \times \mathscr{N}$ is a metric, i.e., it satisfies all the requirements of Definition 3.1.

4.2. Prove that for any $\epsilon > 0$ there is an $N(\epsilon)$ such that

$$\|f_m - f_n\| = \left(\int_{-\infty}^{+\infty} |f_m(x) - f_n(x)|^2 \, dx \right)^{1/2} < \epsilon$$

for $m, n > N(\epsilon)$, where f_n is given by (4.2).

4.3. Check that the operations (4.3) satisfy the axioms in Definition 1.1.

4.4. Check that (4.4) satisfies the requirements of Definition 2.1.

4.5. Show that if $\mathscr{N}$ is a normed space and $\tilde{\mathscr{N}}$ is the completion of $\mathscr{N}$ in the norm, then:

(a) $\tilde{\mathscr{N}}$ is a linear space with respect to the operations

$$\tilde{f} + \tilde{g} = \{f_1 + g_1, f_2 + g_2, \ldots\},$$
$$a\tilde{f} = \{af_1, af_2, \ldots\};$$

(b) the limit $\|\tilde{f}\|_s = \lim_{n\to\infty} \|f_n\|$ exists for every Cauchy sequence $\{f_1, f_2, \ldots\}$ and defines a norm in $\tilde{\mathscr{N}}$;

(c) $\tilde{\mathscr{N}}$ is a Banach space and the image $\mathscr{N}'$ of $\mathscr{N}$ in $\tilde{\mathscr{N}}$ defined by the mapping $f \leftrightarrow \{f, f, \ldots\}$ is a linear subspace of $\tilde{\mathscr{N}}$ which is everywhere dense in $\mathscr{N}$.

4.6. Show that (4.6), (4.7), and (4.8) satisfy the axioms for vector addition, multiplication by a scalar, and inner product respectively.

4.7. Show that the subset D of $l^2(\infty)$ is countable, where D consists of all vectors α which have the properties: (1) a finite number of components $a_1, \ldots, a_n$ (for some integer $n = 1, 2, \ldots$) of α are complex numbers with real and imaginary parts which are rational numbers; (2) the rest of the components vanish.

4.8. Show that every finite-dimensional Euclidean space is a separable Hilbert space.

4.9. Prove Theorem 4.4.

4.10. Show that if in a Euclidean space $f_1, f_2, \ldots$ converges in the norm f and $g_1, g_2, \ldots$ to g, then $\langle f \mid g \rangle = \lim_{n\to\infty} \langle f_n \mid g_n \rangle$.

4.11. Show that the mapping $f \leftrightarrow \alpha_f$ of $\mathscr{H}$ onto $l^2(\infty)$ satisfies the requirements for an isomorphism, given in Definition 2.4.

4.12. Prove that if one orthonormal system $\{e_1, e_2, \ldots\}$ in a Euclidean space $\mathscr{E}$ satisfies either (4.14), or (4.15), or (4.16), for every vector f (or, in case of (4.15), for any two vectors f and g) from $\mathscr{E}$, then $\{e_1, e_2, \ldots\}$ is a basis in $\mathscr{E}$.

4.13. Verify that the criterion of Theorem 4.6(a) is not sufficient to insure that an orthonormal system $\{e_1, e_2, \ldots\}$ in a Euclidean space $\mathscr{E}$ satisfying that criterion is a basis by showing the following:

Let $\{h_1, h_2, \ldots\}$ be an orthonormal basis in a Hilbert space $\mathscr{H}$, and let $\mathscr{E}$ be the vector subspace spanned by $(\sum_{k=1}^{\infty} (1/k) h_k)$, h_2, h_3, ..., i.e., $\mathscr{E} = (\sum_{k=1}^{\infty} (1/k) h_k, h_2, \ldots)$; then $\mathscr{E}$ is a Euclidean space. Prove that:

(a) $\{e_1 = h_2, e_2 = h_3, \ldots, e_n = h_{n+1}, \ldots\}$ is not an orthonormal basis in $\mathscr{E}$.

(b) If $f \in \mathscr{E}$ is orthogonal to $\{e_1, e_2, \ldots\} \subset \mathscr{E}$, then $f = \mathbf{0}$.

5. Wave-Mechanical Treatment of a Single Particle Moving in One Dimension

5.1. The Formalism and Its Partial Physical Interpretation

As an illustration of a physical application of the preceding results, we shall consider the case of a particle restricted to move in only one space dimension within a potential well. We denote the space-coordinate variable by x and the time variable by t. Assume that on our system there acts a force field $F(x)$ which can be derived from a potential $V(x)$, i.e., $F(x) = -(d/dx)\, V(x)$. In classical mechanics, if we denote the momentum of the particle by p, we have the following expression for the total energy E of a particle of mass m:

$$E = p^2/2m + V(x). \tag{5.1}$$

Classically the state of the particle is described by its trajectory $x(t)$, where at any moment t, $x(t) \in \mathbb{R}^1$.

As we mentioned in the Introduction, one of the postulates of quantum mechanics is that the state of a system is described by a function $\Psi(t)$, where $\Psi(t)$ is a vector in a Hilbert space. In the wave mechanics version of quantum mechanics, the state of a one-particle system is postulated to be described at time t by a "wave function" $\psi(x, t)$ which is required to satisfy the condition

$$\int_{-\infty}^{+\infty} |\psi(x, t)|^2 \, dx = 1. \tag{5.2}$$

As a function of t, $\psi(x, t)$ is assumed to be once continuously differentiable in t; in addition we require for the present that $\psi(x, t)$ have a piecewise continuous second derivative in x. Thus, we can consider $\psi(x, t)$ to be at any fixed time t an element of the Euclidean space $\mathscr{C}^1_{(2)}(\mathbb{R}^1)$ (see Exercise 5.2) of all complex functions $f(x)$ which are square integrable, i.e.,

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

and once continuously differentiable. In $\mathscr{C}^1_{(2)}(\mathbb{R}^1)$ the inner product is taken to be

$$\langle f \mid g \rangle = \int_{-\infty}^{+\infty} f^*(x)\, g(x)\, dx,$$

and consequently we recognize (5.2) to be the normalization condition

$$\| \Psi(t) \|^2 = \int_{-\infty}^{+\infty} |\psi(x, t)|^2 \, dx = 1,$$

where $\Psi(t) \in \mathscr{C}^1_{(2)}(\mathbb{R}^1)$ denotes the vector represented by the function $f_t(x) = \psi(x, t)$.

As a dynamical law we have in classical mechanics an equation of motion derivable from Newton's second law, which in the present case is

$$-\frac{dV}{dx} = m\ddot{x}, \qquad \ddot{x} = \frac{d^2x(t)}{dt^2}. \tag{5.3}$$