Chapter 2

Mathematical Tools of Quantum Mechanics

2.1 Introduction

We deal here with the mathematical machinery needed to study quantum mechanics. Although this chapter is mathematical in scope, no attempt is made to be mathematically complete or rigorous. We limit ourselves to those practical issues that are relevant to the formalism of quantum mechanics.

The Schrödinger equation is one of the cornerstones of the theory of quantum mechanics; it has the structure of a *linear* equation. The formalism of quantum mechanics deals with operators that are linear and wave functions that belong to an abstract Hilbert space. The mathematical properties and structure of Hilbert spaces are essential for a proper understanding of the formalism of quantum mechanics. For this, we are going to review briefly the properties of Hilbert spaces and those of linear operators. We will then consider Dirac's *bra-ket* notation.

Quantum mechanics was formulated in two different ways by Schrödinger and Heisenberg. Schrödinger's wave mechanics and Heisenberg's matrix mechanics are the representations of the general formalism of quantum mechanics in *continuous* and *discrete* basis systems, respectively. For this, we will also examine the mathematics involved in representing kets, bras, bra-kets, and operators in discrete and continuous bases.

2.2 The Hilbert Space and Wave Functions

2.2.1 The Linear Vector Space

A linear vector space consists of two sets of elements and two algebraic rules:

- a set of *vectors* ψ, ϕ, χ, $\ldots$ and a set of *scalars* a, b, c, $\ldots$;
- a rule for vector *addition* and a rule for scalar *multiplication*.

(a) Addition rule
The addition rule has the properties and structure of an *abelian* group:

- If ψ and ϕ are vectors (elements) of a space, their sum, $\psi + \phi$, is also a vector of the same space.
- Commutativity: $\psi + \phi = \phi + \psi$.
- Associativity: $(\psi + \phi) + \chi = \psi + (\phi + \chi)$.
- Existence of a zero or neutral vector: for each vector ψ, there must exist a zero vector O such that: $O + \psi = \psi + O = \psi$.
- Existence of a symmetric or inverse vector: each vector ψ must have a symmetric vector $(-\psi)$ such that $\psi + (-\psi) = (-\psi) + \psi = O$.

(b) Multiplication rule
The multiplication of vectors by scalars (scalars can be real or complex numbers) has these properties:

- The product of a scalar with a vector gives another vector. In general, if ψ and ϕ are two vectors of the space, any linear combination $a\psi + b\phi$ is also a vector of the space, a and b being scalars.
- Distributivity with respect to addition:

$$a(\psi + \phi) = a\psi + a\phi, \qquad (a + b)\psi = a\psi + b\psi, \tag{2.1}$$

- Associativity with respect to multiplication of scalars:

$$a(b\psi) = (ab)\psi \tag{2.2}$$

- For each element ψ there must exist a unitary scalar I and a zero scalar "o" such that

$$I\psi = \psi I = \psi \quad \text{and} \quad o\psi = \psi o = o. \tag{2.3}$$

2.2.2 The Hilbert Space

A Hilbert space $\mathcal{H}$ consists of a set of vectors $\psi, \phi, \chi, \ldots$ and a set of *scalars* $a, b, c, \ldots$ which satisfy the following *four* properties:

(a) $\mathcal{H}$ is a linear space

The properties of a linear space were considered in the previous section.

(b) $\mathcal{H}$ has a defined scalar product that is strictly positive

The scalar product of an element ψ with another element ϕ is in general a complex number, denoted by (ψ, ϕ), where (ψ, ϕ) = complex number. **Note:** Watch out for the order! Since the scalar product is a complex number, the quantity (ψ, ϕ) is generally not equal to (ϕ, ψ): $(\psi, \phi) = \psi^*\phi$ while $(\phi, \psi) = \phi^*\psi$. The scalar product satisfies the following properties:

- The scalar product of ψ with ϕ is equal to the complex conjugate of the scalar product of ϕ with ψ:

$$(\psi, \phi) = (\phi, \psi)^*. \tag{2.4}$$

- The scalar product of ϕ with ψ is linear with respect to the second factor if $\psi = a\psi_1 + b\psi_2$:

$$(\phi, a\psi_1 + b\psi_2) = a(\phi, \psi_1) + b(\phi, \psi_2), \tag{2.5}$$

and antilinear with respect to the first factor if $\phi = a\phi_1 + b\phi_2$:

$$(a\phi_1 + b\phi_2, \psi) = a^*(\phi_1, \psi) + b^*(\phi_2, \psi). \tag{2.6}$$

- The scalar product of a vector ψ with itself is a positive real number:

$$(\psi, \psi) = \parallel \psi \parallel^2 \geq 0, \tag{2.7}$$

where the equality holds only for $\psi = O$.

(c) $\mathcal{H}$ is separable

There exists a Cauchy sequence $\psi_n \in \mathcal{H}$ ($n = 1, 2, \ldots$) such that for every ψ of $\mathcal{H}$ and $\varepsilon > 0$, there exists at least one ψ_n of the sequence for which

$$\parallel \psi - \psi_n \parallel < \varepsilon. \tag{2.8}$$

(d) $\mathcal{H}$ is complete

Every Cauchy sequence $\psi_n \in \mathcal{H}$ converges to an element of $\mathcal{H}$. That is, for any ψ_n, the relation

$$\lim_{n,m\to\infty} \parallel \psi_n - \psi_m \parallel = 0, \tag{2.9}$$

defines a unique limit ψ of $\mathcal{H}$ such that

$$\lim_{n\to\infty} \parallel \psi - \psi_n \parallel = 0. \tag{2.10}$$

Remark

We should note that in a scalar product (ϕ, ψ), the second factor, ψ, belongs to the Hilbert space $\mathcal{H}$, while the first factor, ϕ, belongs to its dual Hilbert space $\mathcal{H}_d$. The distinction between $\mathcal{H}$ and $\mathcal{H}_d$ is due to the fact that, as mentioned above, the scalar product is not commutative: $(\phi, \psi) \neq (\psi, \phi)$; the order matters! From linear algebra, we know that every vector space can be associated with a dual vector space.

2.2.3 Dimension and Basis of a Vector Space

A set of N nonzero vectors $\phi_1, \phi_2, \ldots, \phi_N$ is said to be *linearly independent* if and only if the solution of the equation

$$\sum_{i=1}^{N} a_i \phi_i = 0 \tag{2.11}$$

is $a_1 = a_2 = \cdots = a_N = 0$. But if there exists a set of scalars, which are not all zero, so that one of the vectors (say ϕ_n) can be expressed as a linear combination of the others,

$$\phi_n = \sum_{i=1}^{n-1} a_i \phi_i + \sum_{i=n+1}^{N} a_i \phi_i, \tag{2.12}$$

the set $\{\phi_i\}$ is said to be *linearly dependent*.

Dimension: The *dimension* of a vector space is given by the *maximum number* of linearly independent vectors the space can have. For instance, if the maximum number of linearly independent vectors a space has is N (i.e., $\phi_1, \phi_2, \ldots, \phi_N$), this space is said to be N-dimensional. In this N-dimensional vector space, any vector ψ can be expanded as a linear combination:

$$\psi = \sum_{i=1}^{N} a_i \phi_i. \tag{2.13}$$

Basis: The *basis* of a vector space consists of a set of the maximum possible number of linearly independent vectors belonging to that space. This set of vectors, $\phi_1, \phi_2, \ldots, \phi_N$, to be denoted in short by $\{\phi_i\}$, is called the basis of the vector space, while the vectors $\phi_1, \phi_2, \ldots, \phi_N$ are called the base vectors. Although the set of these linearly independent vectors is arbitrary, it is convenient to choose them *orthonormal*; that is, their scalar products satisfy the relation $(\phi_i, \phi_j) = \delta_{ij}$ (we may recall that $\delta_{ij} = 1$ whenever $i = j$ and zero otherwise). The basis is said to be *orthonormal* if it consists of a set of orthonormal vectors. Moreover, the basis is said to be *complete* if it spans the entire space; that is, there is no need to introduce any additional base vector. The expansion coefficients a_i in (2.13) are called the *components* of the vector ψ in the basis. Each component is given by the scalar product of ψ with the corresponding base vector, $a_j = (\phi_j, \psi)$.

Examples of linear vector spaces

Let us give two examples of linear spaces that are Hilbert spaces: one having a *finite (discrete)* set of base vectors, the other an *infinite (continuous)* basis.

- The first one is the three-dimensional Euclidean vector space; the basis of this space consists of three linearly independent vectors, usually denoted by $\vec{i}, \vec{j}, \vec{k}$. Any vector of the Euclidean space can be written in terms of the base vectors as $\vec{A} = a_1\vec{i} + a_2\vec{j} + a_3\vec{k}$, where a_1, a_2, and a_3 are the components of $\vec{A}$ in the basis; each component can be determined by taking the scalar product of $\vec{A}$ with the corresponding base vector: $a_1 = \vec{i}\cdot\vec{A}$, $a_2 = \vec{j}\cdot\vec{A}$, and $a_3 = \vec{k}\cdot\vec{A}$. Note that the scalar product in the Euclidean space is real and hence symmetric. The norm in this space is the usual length of vectors $\| \vec{A} \| = A$. Note also that whenever $a_1\vec{i} + a_2\vec{j} + a_3\vec{k} = 0$ we have $a_1 = a_2 = a_3 = 0$ and that none of the unit vectors $\vec{i}, \vec{j}, \vec{k}$ can be expressed as a linear combination of the other two.
- The second example is the space of the entire complex functions $\psi(x)$; the dimension of this space is infinite for it has an infinite number of linearly independent basis vectors.

Example 2.1

Check whether the following sets of functions are linearly independent or dependent on the real x-axis.

(a) $f(x) = 4$, $g(x) = x^2$, $h(x) = e^{2x}$
(b) $f(x) = x$, $g(x) = x^2$, $h(x) = x^3$
(c) $f(x) = x$, $g(x) = 5x$, $h(x) = x^2$
(d) $f(x) = 2 + x^2$, $g(x) = 3 - x + 4x^3$, $h(x) = 2x + 3x^2 - 8x^3$

Solution

(a) The first set is clearly linearly independent since $a_1 f(x) + a_2 g(x) + a_3 h(x) = 4a_1 + a_2 x^2 + a_3 e^{2x} = 0$ implies that $a_1 = a_2 = a_3 = 0$ for any value of x.

(b) The functions $f(x) = x$, $g(x) = x^2$, $h(x) = x^3$ are also linearly independent since $a_1 x + a_2 x^2 + a_3 x^3 = 0$ implies that $a_1 = a_2 = a_3 = 0$ no matter what the value of x. For instance, taking $x = -1, 1, 3$, the following system of three equations

$$-a_1 + a_2 - a_3 = 0, \qquad a_1 + a_2 + a_3 = 0, \qquad 3a_1 + 9a_2 + 27a_3 = 0 \tag{2.14}$$

yields $a_1 = a_2 = a_3 = 0$.

(c) The functions $f(x) = x$, $g(x) = 5x$, $h(x) = x^2$ are not linearly independent, since $g(x) = 5f(x) + 0 \times h(x)$.

(d) The functions $f(x) = 2 + x^2$, $g(x) = 3 - x + 4x^3$, $h(x) = 2x + 3x^2 - 8x^3$ are not linearly independent since $h(x) = 3f(x) - 2g(x)$.

Example 2.2

Are the following sets of vectors (in the three-dimensional Euclidean space) linearly independent or dependent?

(a) $\vec{A} = (3, 0, 0)$, $\vec{B} = (0, -2, 0)$, $\vec{C} = (0, 0, -1)$

(b) $\vec{A} = (6, -9, 0)$, $\vec{B} = (-2, 3, 0)$

(c) $\vec{A} = (2, 3, -1)$, $\vec{B} = (0, 1, 2)$, $\vec{C} = (0, 0, -5)$

(d) $\vec{A} = (1, -2, 3)$, $\vec{B} = (-4, 1, 7)$, $\vec{C} = (0, 10, 11)$, and $\vec{D} = (14, 3, -4)$

Solution

(a) The three vectors $\vec{A} = (3, 0, 0)$, $\vec{B} = (0, -2, 0)$, $\vec{C} = (0, 0, -1)$ are linearly independent, since

$$a_1\vec{A} + a_2\vec{B} + a_3\vec{C} = 0 \Longrightarrow 3a_1\vec{i} - 2a_2\vec{j} - a_3\vec{k} = 0 \tag{2.15}$$

leads to

$$3a_1 = 0, \qquad -2a_2 = 0, \qquad -a_3 = 0, \tag{2.16}$$

which yields $a_1 = a_2 = a_3 = 0$.

(b) The vectors $\vec{A} = (6, -9, 0)$, $\vec{B} = (-2, 3, 0)$ are linearly dependent, since the solution to

$$a_1\vec{A} + a_2\vec{B} = 0 \quad \Longrightarrow \quad (6a_1 - 2a_2)\vec{i} + (-9a_1 + 3a_2)\vec{j} = 0 \tag{2.17}$$

is $a_1 = a_2/3$. The first vector is equal to -3 times the second one: $\vec{A} = -3\vec{B}$.

(c) The vectors $\vec{A} = (2, 3, -1)$, $\vec{B} = (0, 1, 2)$, $\vec{C} = (0, 0, -5)$ are linearly independent, since

$$a_1\vec{A} + a_2\vec{B} + a_3\vec{C} = 0 \Longrightarrow 2a_1\vec{i} + (3a_1 + a_2)\vec{j} + (-a_1 + 2a_2 - 5a_3)\vec{k} = 0 \tag{2.18}$$

leads to

$$2a_1 = 0, \qquad 3a_1 + a_2 = 0, \qquad -a_1 + 2a_2 - 5a_3 = 0. \tag{2.19}$$

The only solution of this system is $a_1 = a_2 = a_3 = 0$.

(d) The vectors $\vec{A} = (1, -2, 3)$, $\vec{B} = (-4, 1, 7)$, $\vec{C} = (0, 10, 11)$, and $\vec{D} = (14, 3, -4)$ are not linearly independent, because $\vec{D}$ can be expressed in terms of the other vectors:

$$\vec{D} = 2\vec{A} - 3\vec{B} + \vec{C}. \tag{2.20}$$

2.2.4 Square-Integrable Functions: Wave Functions

In the case of function spaces, a "vector" element is given by a *complex function* and the *scalar product* by *integrals*. That is, the scalar product of two functions $\psi(x)$ and $\phi(x)$ is given by

$$(\psi, \phi) = \int \psi^*(x)\phi(x)\, dx. \tag{2.21}$$

If this integral *diverges*, the scalar product *does not exist*. As a result, if we want the function space to possess a scalar product, we must select only those functions for which (ψ, ϕ) is *finite*. In particular, a function $\psi(x)$ is said to be *square integrable* if the scalar product of ψ with itself,

$$(\psi, \psi) = \int |\psi(x)|^2\, dx, \tag{2.22}$$

is *finite*.

It is easy to verify that the space of square-integrable functions possesses the properties of a Hilbert space. For instance, any linear combination of square-integrable functions is also a square-integrable function and (2.21) satisfies all the properties of the scalar product of a Hilbert space.

Note that the dimension of the Hilbert space of square-integrable functions is infinite, since each wave function can be expanded in terms of an infinite number of linearly independent functions. The dimension of a space is given by the maximum number of linearly independent basis vectors required to span that space.

A good example of square-integrable functions is the *wave function* of quantum mechanics, $\psi(\vec{r}, t)$. We have seen in Chapter 1 that, according to Born's probabilistic interpretation of $\psi(\vec{r}, t)$, the quantity $|\ \psi(\vec{r}, t)\ |^2\ d^3r$ represents the probability of finding, at time t, the particle in a volume d^3r, centered around the point $\vec{r}$. The probability of finding the particle somewhere in space must then be equal to 1:

$$\int |\ \psi(\vec{r}, t)\ |^2\ d^3r = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} |\ \psi(\vec{r}, t)\ |^2\ dz = 1; \tag{2.23}$$

hence the wave functions of quantum mechanics are square-integrable. Wave functions satisfying (2.23) are said to be normalized or square-integrable. As wave mechanics deals with square-integrable functions, any wave function which is not square-integrable has no physical meaning in quantum mechanics.

2.3 Dirac Notation

The physical state of a system is represented in quantum mechanics by elements of a Hilbert space; these elements are called state vectors. We can represent the state vectors in different bases by means of function expansions. This is analogous to specifying an ordinary (Euclidean) vector by its components in various coordinate systems. For instance, we can represent equivalently a vector by its components in a Cartesian coordinate system, in a spherical coordinate system, or in a cylindrical coordinate system. *The meaning of a vector is, of course, independent of the coordinate system chosen to represent its components.* Similarly, the state of a microscopic system has a meaning independent of the basis in which it is expanded.

To free state vectors from coordinate meaning, Dirac introduced what was to become an invaluable notation in quantum mechanics; it allows one to manipulate the formalism of quantum