Mathematical Structures of Quantum Mechanics

Kow Lung Chang

Physics Department, National Taiwan University

TO FEZA AND SUHA

Preface

During the past few years, after a couple of weeks of lecturing the course of quantum mechanics that I offered at the Physics Department, National Taiwan University, some students would usually come to ask me as to what extent they had to refurbish their mathematical background in order to follow my lecture with ease and confidence. It was hard for me to provide a decent and proper answer to the question, and very often students would show reluctance to invest extra time on subjects such as group theory or functional analysis when I advised them to take some advanced mathematics courses. All these experiences that I have encountered in my class eventually motivated me to write this book.

The book is designed with the hope that it might be helpful to those students I mentioned above. It could also serve as a complementary text in quantum mechanics for students of inquiring minds who appreciate the rigor and beauty of quantum theory.

Assistance received from many sources made the appearance of this book possible. I wish to express here my great appreciation and gratitude to Dr. Yusuf Gürsey, who painstakingly went through the manuscript and responded generously by giving very helpful suggestions and comments, and made corrections line by line. I would also like to thank Mr. Paul Black who provided me with cogent suggestions and criticism of the manuscript, particularly in those sections on quantum uncertainty. I am indebted as well to Mr. Chih Han Lin who, with immense patience, compiled the whole text and drew all the figures from my suggestions. All his hard work and attention resulted in the present form of this book.

Taipei, Taiwan
March, 2011 Kow Lung Chang

Contents

Chapter 1

Postulates and Principles of Quantum Mechanics

As with many fields in physics, a precise and rigorous description of a given subject requires the use of some mathematical tools. Take Lagrange's formulation of classical mechanics for instance, one needs the basic knowledge of variational calculus in order to derive the equations of motion for a system of particles in terms of generalized coordinates. To formulate the postulates of quantum mechanics, it would also be necessary to acquire some knowledge on vector space in general, and Hilbert space in particular. It is in this chapter that we shall provide the minimum but essential mathematical preparation that allows one to perceive and understand the general framework of quantum theory and to appreciate the rigorous derivation of the quantum principles.

1.1 Vector space

A **vector space** $\mathcal{V}$ is a set of elements, called vectors with the following 2 operations:

- An operation of addition, which for each pair of vectors ψ and ϕ, corresponds to a new vector $\psi + \phi \in \mathcal{V}$, called the sum of ψ and ϕ.
- An operation of scalar multiplication, which for each vector ψ and a number a, specifies a vector $a\psi$, such that (assuming a, b are numbers and ψ, ϕ and χ are vectors)

$$\psi + \phi = \phi + \psi, \tag{1.1a}$$
$$\psi + (\phi + \chi) = (\psi + \phi) + \chi, \tag{1.1b}$$
$$\psi + \mathit{0} = \psi, \quad \mathit{0} \text{ is null vector}, \tag{1.1c}$$
$$a(\psi + \phi) = a\psi + a\phi, \tag{1.1d}$$
$$(a + b)\psi = a\psi + b\psi, \tag{1.1e}$$
$$a(b\psi) = (ab)\psi, \tag{1.1f}$$
$$1 \cdot \psi = \psi, \tag{1.1g}$$
$$0 \cdot \psi = \mathit{0}, \tag{1.1h}$$

where if a, b are real numbers, we call this vector space the real vector space, and denote it by $\mathcal{V}_r$. On the other way, complex vector space $\mathcal{V}_c$ means a, b are complex numbers.

Example

We take n-dimensional Euclidean space, $\mathcal{R}^n$-space, as an exmple. It is a vector space with the vectors ψ and ϕ specified as $\psi = (x_1, x_2, \ldots, x_i, \ldots, x_n)$ and $\phi = (y_1, y_2, \ldots, y_i, \ldots, y_n)$, where x_i and y_i ($i = 1, 2, \ldots, n$) are all taken as real numbers. The sum of ψ and ϕ becomes $(x_1 + y_1, x_2+y_2, \ldots, x_i+y_i, \ldots, x_n+y_n)$ and $a\psi = (ax_1, ax_2, \ldots, ax_i, \ldots, ax_n)$. If a and x_i are taken as complex numbers, then ψ is a vector in $\mathcal{C}^n$-space; a n-dimensional complex vector space.

It is easily understood that a set of the continuous functions $f(x)$ for $a \leqslant x \leqslant b$ forms a vector space, namely $\mathcal{L}^2(a, b)$-space.

Before leaving this section, we also introduce some terminologies in the following subsections that will be frequently referred to in later chapters.

1.1.1 Linearly dependent and linearly independent

Consider a set of m vectors $\{\psi_1, \psi_2, \ldots, \psi_m\}$, and we construct the linear combination of these m vectors as follows:

$$\sum_{i=1}^{m} a_i \psi_i. \tag{1.2}$$

This linear combination of m vectors is of course a vector. It becomes a **null vector** if and only if all the coefficient $a_i = 0$ for $i = 1, 2, \ldots, m$, then the set of m vectors $\{\psi_1, \psi_2, \ldots, \psi_m\}$ is called linearly independent. If at least one of the coefficient $a_l \neq 0$ such that $\sum_{i=1}^{m} a_i\psi_i = \mathit{0}$, then the set $\{\psi_1, \psi_2, \ldots, \psi_m\}$ is called linearly dependent.

1.1.2 Dimension and basis

The maximum number of linearly independent vectors in $\mathcal{V}$ is called the **dimension** of $\mathcal{V}$. Any n-linearly independent vectors in n-dimensional vector space $\mathcal{V}$ form the **basis** of the vector space.

1.2 Inner product

An **inner product**, or sometimes called **scalar product** in vector space, is a numerically valued function of the ordered pair of vectors ψ and ϕ, denoted by (ψ, ϕ), and for a scalar a, such that

$$(\psi, \phi + \chi) = (\psi, \phi) + (\psi, \chi), \tag{1.3a}$$
$$(\psi, a\phi) = a(\psi, \phi), \tag{1.3b}$$
$$(\psi, \phi) = (\phi, \psi)^*, \tag{1.3c}$$
$$(\psi, \psi) \geqslant 0, (\psi, \psi) = 0 \text{ if and only if } \psi \text{ is a null vector.} \tag{1.3d}$$

Two vectors ψ and ϕ are said to be orthogonal to each other if their corresponding inner product vanishes, namely $(\psi, \phi) = 0$.

For example, let us consider the vectors in $\mathcal{C}^n$-space $\psi = (x_1, x_2, \ldots, x_n)$ and $\phi = (y_1, y_2, \ldots, y_n)$ where x_i and y_i are complex numbers. The inner product of ψ and ϕ written as

$$(\psi, \phi) = \sum_{i=1}^{n} x_i^* y_i = x_1^* y_1 + x_2^* y_2 + \cdots + x_n^* y_n. \tag{1.4}$$

Consider the set of continuous function of $f(x)$ where $a \leqslant x \leqslant b$. An ordered pair of functions $f(x)$ and $g(x)$ define the inner product as $(f(x), g(x)) = \int_a^b f(x)^* g(x) dx$. This vector space is called $\mathcal{L}^2(a, b)$-space when $|f(x)|^2$ and $|g(x)|^2 =$ finite.

1.2.1 Schwarz inequality

We are now in the position to prove the Schwarz inequality.

Let ψ and ϕ be any two vectors. The **Schwarz inequality** reads as

$$|(\psi, \phi)| = \sqrt{(\psi, \phi)(\phi, \psi)} \leqslant \sqrt{(\psi, \psi)}\sqrt{(\phi, \phi)}. \tag{1.5}$$

Proof

Since $(\psi + \alpha\phi, \psi + \alpha\phi) \geqslant 0$, where $\alpha = \xi + i\eta$ is a complex number. Regard this inner product $(\psi + \alpha\phi, \psi + \alpha\phi) = f(\xi, \eta)$ as a function of two variables ξ and η. Then

$$f(\xi, \eta) = (\psi, \psi) + |\alpha|^2(\phi, \phi) + \alpha(\psi, \phi) + \alpha^*(\phi, \psi), \tag{1.6}$$

which is positive definite. Let us look for the minimum of $f(\xi, \eta)$ at ξ_0, η_0 by solving

$$\left.\frac{\partial f(\xi, \eta)}{\partial \xi}\right|_{\xi_0,\eta_0} = \left.\frac{\partial f(\xi, \eta)}{\partial \eta}\right|_{\xi_0,\eta_0} = 0, \tag{1.7}$$

and we obtain

$$\xi_0 = \frac{1}{2}\frac{(\psi, \phi) + (\phi, \psi)}{(\phi, \phi)}, \qquad \eta_0 = -\frac{i}{2}\frac{(\psi, \phi) - (\phi, \psi)}{(\phi, \phi)}. \tag{1.8}$$

Therefore

$$f(\xi_0, \eta_0) = (\psi, \psi) - \frac{(\psi, \phi)(\phi, \psi)}{(\phi, \phi)} \geqslant 0, \tag{1.9}$$

that can be cast into the familiar expression of Schwarz inequality.

1.2.2 Gram-Schmidt orthogonalization process

The inner product we have been considering can be applied to the orthogonalization of the basis in the n-dimensional vector space. Let

$\{\psi_1, \psi_2, \ldots, \psi_n\} \in \mathcal{V}$ be the set of n-linearly independent vectors. Since $(\psi_i, \psi_j) \neq 0$ in general, we can construct a new set of vectors $\{\psi'_1, \psi'_2, \ldots, \psi'_n\}$ such that $(\psi'_i, \psi'_j) = 0$ for all i and j unless $i = j$, namely ψ'_i and ψ'_j are orthogonal to each other for $i \neq j$ by the following procedure:

First take $\psi'_1 = \psi_1$ and construct $\psi'_2 = \psi_2 + \alpha\psi'_1$. In order to force ψ'_2 to be orthogonal to ψ'_1, we solve the α as to meet the condition $(\psi'_2, \psi'_1) = 0$, i.e.

$$(\psi'_2, \psi'_1) = (\psi_2, \psi'_1) + \alpha^*(\psi'_1, \psi'_1) = 0, \tag{1.10}$$

and we obtain $\alpha = -(\psi_2, \psi'_1)^*/(\psi'_1, \psi'_1) = -(\psi'_1, \psi_2)/(\psi'_1, \psi'_1)$, hence

$$\psi'_2 = \psi_2 - \psi'_1 \frac{(\psi'_1, \psi_2)}{(\psi'_1, \psi'_1)}. \tag{1.11}$$

The same procedure can be performed repeatedly to reach $\psi'_3 = \psi_3 + \alpha\psi'_2 + \beta\psi'_1$ which guarantees $(\psi'_3, \psi'_1) = (\psi'_3, \psi'_2) = 0$ with $\alpha = -(\psi'_2, \psi_3)/(\psi'_2, \psi'_2)$ and $\beta = -(\psi'_1, \psi_3)/(\psi'_1, \psi'_1)$. In general,

$$\psi'_i = \psi_i - \psi'_{i-1} \frac{(\psi'_{i-1}, \psi_i)}{(\psi'_{i-1}, \psi'_{i-1})} - \psi'_{i-2} \frac{(\psi'_{i-2}, \psi_i)}{(\psi'_{i-2}, \psi'_{i-2})} - \cdots - \psi'_1 \frac{(\psi'_1, \psi_i)}{(\psi'_1, \psi'_1)}. \tag{1.12}$$

The set of orthogonal basis $\{\psi'_1, \psi'_2, \ldots, \psi'_n\}$ can be normalized immediately by multiplying the inverse square root of the corresponding inner product, i.e.

$$\tilde{\psi}_i = \frac{\psi'_i}{\sqrt{(\psi'_i, \psi'_i)}}, \tag{1.13}$$

and $\{\tilde{\psi}_1, \tilde{\psi}_2, \ldots, \tilde{\psi}_n\}$ becomes the **orthonormal** set of the basis in the vector space. From now on we shall take the basis to be orthonormal without mentioning it particularly.

Example

Consider the following set of continuous functions in $C(-\infty, \infty)$

$$f_n(x) = x^n \exp\left(-\frac{x^2}{2}\right), \quad n = 0, 1, \ldots . \tag{1.14}$$

We construct the new set of orthogonal vectors by applying the Gram-Schmidt process and obtain:

$$f'_0(x) = f_0(x) = \exp\left(-\frac{x}{2}\right), \tag{1.15}$$

$$f'_1(x) = f_1 - \frac{f'_0(f'_0, f_1)}{(f'_0, f'_0)} = f_1(x) = x \exp\left(\frac{x^2}{2}\right), \tag{1.16}$$

$$f'_2(x) = f_2 - \frac{f'_1(f'_1, f_2)}{(f'_1, f'_1)} - \frac{f'_0(f'_0, f_2)}{(f'_0, f'_0)} = \left(x^2 - \frac{1}{2}\right) \exp\left(-\frac{x^2}{2}\right). \tag{1.17}$$

Similarly we have $f'_3(x) = (x^3 - 3x/2)\exp(-x^2/2)$. The orthonormal functions can be calculated according to

$$\tilde{f}_n(x) = \frac{f'_n(x)}{\sqrt{(f'_n(x), f'_n(x))}} = \frac{1}{\sqrt{2^n n! \sqrt{\pi}}} \exp\left(-\frac{x^2}{2}\right) H_n(x), \tag{1.18}$$

where $H_n(x)$ are called Hermite polynomials. One also recognizes that $\tilde{f}_n(x)$ are in fact, the eigenfunctions of the Schrödinger equation for one-dimension harmonic oscillation.

1.3 Completeness and Hilbert space

Let us introduce some other terminologies in discussing Hilbert space.

1.3.1 Norm

A **norm** on a vector space is a non-negative real function such that, if ψ, ϕ are vectors, the norm of ψ is written as $\|\psi\|$, satisfying:

$$\|\psi\| \geqslant 0, \qquad \|\psi\| = 0 \quad \text{iff} \quad \psi \text{ is null vector}, \tag{1.19a}$$
$$\|a\psi\| = |a| \cdot \|\psi\|, \tag{1.19b}$$
$$\|\psi + \phi\| \leqslant \|\psi\| + \|\phi\|. \tag{1.19c}$$

Example

If $f(x) \in \mathrm{C}(a, b)$, namely if $f(x)$ is a continuous function for a variable that lies between a and b, the norm of $f(x)$ can be defined either as $\|f(x)\| = \mathrm{Max}\{|f(x)|, a \leqslant x \leqslant b\}$ or as the inner product of $f(x)$, i.e.

$$\|f(x)\|^2 = (f(x), f(x)) = \int_a^b |f(x)|^2 dx.$$

1.3.2 Cauchy sequence and convergent sequence

Consider an infinite dimensional vector space and denote the basis by $\{\phi_1, \phi_2, \phi_3, \ldots\}$. We construct the partial sum $\psi_N = \sum_i a_i \phi_i$, where i runs from 1 to N, and obtain $\ldots, \psi_j, \psi_{j+1}, \ldots, \psi_m, \psi_{m+1}, \ldots, \psi_n, \ldots$ for increasing values in N that forms an infinite sequence. The sequence is called a **Cauchy sequence** if

$$\lim_{n\to\infty} \psi_n = \lim_{m\to\infty} \psi_m,$$

or more precisely to put in terms of norm, i.e, $\lim\limits_{n,m\to\infty} \|\psi_n - \psi_m\| = 0$.

It is said that a vector ψ_m converges to ψ if

$$\lim_{m\to\infty} \psi_m = \psi, \quad \text{or} \quad \lim_{m\to\infty} \|\psi_m - \psi\| = 0,$$

then $\{\ldots, \psi_{m-1}, \psi_m, \ldots\}$ is called a **convergent sequence**.

It is easily concluded that every convergent sequence is a Cauchy sequence. Yet it is not necessary true conversely. Namely a Cauchy sequence is not always a convergent sequence.

1.3.3 Complete vector space

A vector space, in which every Cauchy sequence of a vector ψ_m converges to a limiting vector ψ, is called a complete vector space.

1.3.4 Hilbert space

A Hilbert space is a complete vector space with norm defined as the inner product. A Hilbert space, finite dimensional or infinite dimensional, is separable if its basis is countable.

1.4 Linear operator

A **linear operator** $\mathbf{A}$ on a vector space assigns to each vector ψ a new vector, i.e. $\mathbf{A}\psi = \psi'$ such that

$$\mathbf{A}(\psi + \phi) = \mathbf{A}\psi + \mathbf{A}\phi, \quad \mathbf{A}(\alpha\psi) = \alpha\mathbf{A}\psi. \tag{1.20}$$

Two operators $\mathbf{A}$, $\mathbf{B}$ are said equal if $\mathbf{A}\psi = \mathbf{B}\psi$ for all ψ in the vector space.

For convenience in later discussion, we denote

- $\mathbf{O}$: null operator such that $\mathbf{O}\psi = \mathit{0}$ for all ψ, and $\mathit{0}$ is the null vector.
- $\mathbf{I}$: unit operator or identity operator such that $\mathbf{I}\psi = \psi$.

The sum of the operators $\mathbf{A}$ and $\mathbf{B}$ is an operator, such that $(\mathbf{A}+\mathbf{B})\psi = \mathbf{A}\psi + \mathbf{B}\psi$. The product of operators $\mathbf{A}$ and $\mathbf{B}$ is again an operator that one writes as $\mathbf{A}\cdot\mathbf{B}$ or $\mathbf{AB}$ such that $(\mathbf{AB})\psi = \mathbf{A}(\mathbf{B}\psi)$.

The order of the operators in the product matters greatly. It is generally that $\mathbf{AB} \neq \mathbf{BA}$. The associative rule holds for the product of the operators $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$.

1.4.1 Bounded operator

An operator $\mathbf{A}$ is called a **bounded operator** if there exists a positive number b such that

$$\|\mathbf{A}\psi\| \leqslant b\|\psi\|, \quad \text{for any vector } \psi \text{ in the vector space.}$$

The least upperbound (supremum) of $\mathbf{A}$, namely the smalleast number of b for a given operator $\mathbf{A}$ and for any ψ in $\mathcal{V}$, is denoted by

$$\|\mathbf{A}\| = \sup\left\{\frac{\|\mathbf{A}\psi\|}{\|\psi\|}, \quad \psi \neq 0\right\}, \tag{1.21}$$

then $\|\mathbf{A}\psi\| \leqslant \|\mathbf{A}\|\|\psi\|$.

We are now able to show readily that $\|\mathbf{A}+\mathbf{B}\| \leqslant \|\mathbf{A}\| + \|\mathbf{B}\|$.

Proof

Let us denote $\|\mathbf{A}\psi\| \leqslant \|\mathbf{A}\|\|\psi\|$ and $\|\mathbf{B}\psi\| \leqslant \|\mathbf{B}\|\|\psi\|$. Then

$$\begin{aligned}\|\mathbf{A}+\mathbf{B}\| &= \sup\left\{\frac{\|(\mathbf{A}+\mathbf{B})\psi\|}{\|\psi\|}, \psi \neq 0\right\} = \sup\left\{\frac{\|\mathbf{A}\psi + \mathbf{B}\psi\|}{\|\psi\|}, \psi \neq 0\right\} \\ &\leqslant \sup\left\{\frac{\|\mathbf{A}\psi\|}{\|\psi\|}, \psi \neq 0\right\} + \sup\left\{\frac{\|\mathbf{B}\psi\|}{\|\psi\|}, \psi \neq 0\right\} = \|\mathbf{A}\| + \|\mathbf{B}\|.\end{aligned}$$

Similarly, we have $\|\mathbf{AB}\| \leqslant \|\mathbf{A}\|\|\mathbf{B}\|$.

1.4.2 Continuous operator

Consider the convergent sequence $\{\ldots, \psi_m, \psi_{m+1}, \ldots, \psi_n, \ldots\}$ such that $\lim_{n\to\infty} \|\psi_n - \psi\| = 0$. If $\mathbf{A}$ is a bounded operator, then $\{\ldots, \mathbf{A}\psi_m,$ $\mathbf{A}\psi_{m+1}, \ldots, \mathbf{A}\psi_n, \ldots\}$ is also a convergent sequence because

$$\lim_{n\to\infty} \|\mathbf{A}\psi_n - \mathbf{A}\psi\| \leqslant \|\mathbf{A}\| \lim_{n\to\infty} \|\psi_n - \psi\| = 0.$$

We call operator $\mathbf{A}$ the **continuous operator**.

1.4.3 Inverse operator

An operator $\mathbf{A}$ has an **inverse operator** if there exists $\mathbf{B}_R$ such that $\mathbf{AB}_R = \mathbf{I}$, then we call operator $\mathbf{B}_R$ the right inverse of $\mathbf{A}$. Similarly an operator $\mathbf{B}_L$ such that the product operator $\mathbf{B}_L\mathbf{A} = \mathbf{I}$, then we call operator $\mathbf{B}_L$ the left inverse of $\mathbf{A}$. In fact, the left inverse operator is always equal to the right inverse operator for a given operator $\mathbf{A}$, because

$$\mathbf{B}_L = \mathbf{B}_L\mathbf{I} = \mathbf{B}_L(\mathbf{AB}_R) = (\mathbf{B}_L\mathbf{A})\mathbf{B}_R = \mathbf{IB}_R = \mathbf{B}_R. \tag{1.22}$$

The inverse operator of a given operator $\mathbf{A}$ is also unique. If operators $\mathbf{B}$ and $\mathbf{C}$ are all inverse operators of $\mathbf{A}$, then $\mathbf{C} = \mathbf{CI} = \mathbf{C}(\mathbf{AB}) = (\mathbf{CA})\mathbf{B} = \mathbf{B}$.

The implication of uniqueness of the inverse operator of operator $\mathbf{A}$ allows us to write it in the form $\mathbf{A}^{-1}$, namely $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. It is easily verified that $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

1.4.4 Unitary operator

An operator $\mathbf{U}$ is unitary if $\|\mathbf{U}\psi\| = \|\psi\|$. A unitary operation preserves the invariant of the inner product of any pair of vectors, i.e. $(\mathbf{U}\psi, \mathbf{U}\phi) = (\psi, \phi)$. This can be proved as follows:

Let $\chi = \psi + \phi$ and we have

$$\begin{aligned}(\mathbf{U}\chi, \mathbf{U}\chi) &= (\mathbf{U}(\psi+\phi), \mathbf{U}(\psi+\phi)) \\ &= (\mathbf{U}\psi, \mathbf{U}\psi) + (\mathbf{U}\psi, \mathbf{U}\phi) + (\mathbf{U}\phi, \mathbf{U}\psi) + (\mathbf{U}\phi, \mathbf{U}\phi) \\ &= \|\mathbf{U}\psi\|^2 + \|\mathbf{U}\phi\|^2 + 2\Re\{(\mathbf{U}\psi, \mathbf{U}\phi)\},\end{aligned}$$

and on the other hand,

$$\begin{aligned}(\mathbf{U}\chi, \mathbf{U}\chi) &= (\psi+\phi, \psi+\phi) \\ &= (\chi, \chi) = (\psi, \psi) + (\psi, \phi) + (\phi, \psi) + (\phi, \phi) \\ &= \|\psi\|^2 + \|\phi\|^2 + 2\Re\{(\psi, \phi)\}.\end{aligned}$$

Since $\|\mathbf{U}\psi\| = \|\psi\|, \|\mathbf{U}\phi\| = \|\phi\|$, we have $\Re\{(\mathbf{U}\psi, \mathbf{U}\phi)\} = \Re\{(\psi, \phi)\}$. Similarly if $\chi' = \psi + i\phi$, we obtain $(\mathbf{U}\chi', \mathbf{U}\chi') = (\chi', \chi')$, that implies $\Im\{(\mathbf{U}\psi, \mathbf{U}\phi)\} = \Im\{(\psi, \phi)\}$, therefore $(\mathbf{U}\psi, \mathbf{U}\phi) = (\psi, \phi)$.

1.4.5 Adjoint operator

Consider the inner product of $(\psi, \mathbf{A}\phi)$ where $\mathbf{A}$ is a given linear operator of interest. This numerically scalar quantity certainly is a function of operator $\mathbf{A}$ and the pair of vectors ψ and ϕ, namely $(\psi, \mathbf{A}\phi) = F(\mathbf{A}, \psi, \phi)$ is a scalar quantity.

Instead of performing the above inner product straightforwardly, we shall obtain the very same scalar of $(\psi, \mathbf{A}\phi)$ by forming the following inner product $(\mathbf{A}^\dagger\psi, \phi)$ such that $(\psi, \mathbf{A}\phi) \equiv (\mathbf{A}^\dagger\psi, \phi)$. The operator $\mathbf{A}^\dagger$ is called the **adjoint operator** of $\mathbf{A}$. The following relations can be easily established (proofs left to readers):

$$(\mathbf{A} + \mathbf{B})^\dagger = \mathbf{A}^\dagger + \mathbf{B}^\dagger, \tag{1.23a}$$

$$(\alpha\mathbf{A})^\dagger = \alpha^*\mathbf{A}^\dagger, \tag{1.23b}$$

$$(\mathbf{A}\mathbf{B})^\dagger = \mathbf{B}^\dagger\mathbf{A}^\dagger, \tag{1.23c}$$

$$(\mathbf{A}^\dagger)^\dagger = \mathbf{A}, \tag{1.23d}$$

$$(\mathbf{A}^\dagger)^{-1} = (\mathbf{A}^{-1})^\dagger. \tag{1.23e}$$

It can also be shown that $\mathbf{A}^\dagger$ is a bounded operator if $\mathbf{A}$ is bounded and their norms are equal, i.e. $\|A\| = \|A^\dagger\|$.

To prove the above equality, let us consider $\|\mathbf{A}^\dagger\psi\|^2 = (\mathbf{A}^\dagger\psi, \mathbf{A}^\dagger\psi)$, namely

$$\|\mathbf{A}^\dagger\psi\|^2 = (\mathbf{A}^\dagger\psi, \mathbf{A}^\dagger\psi) = (\mathbf{A}\mathbf{A}^\dagger\psi, \psi) \leqslant \|\psi\|\|\mathbf{A}\mathbf{A}^\dagger\psi\| \leqslant \|\psi\|\|\mathbf{A}\|\|\mathbf{A}^\dagger\psi\|,$$

therefore $\|\mathbf{A}^\dagger\psi\| \leqslant \|\mathbf{A}\|\|\psi\|$, and we have $\|\mathbf{A}^\dagger\| \leqslant \|\mathbf{A}\|$.

On the other hand, we have $\|\mathbf{A}\psi\|^2 = (\mathbf{A}\psi, \mathbf{A}\psi) = (\mathbf{A}^\dagger\mathbf{A}\psi, \psi) \leqslant \|\psi\|\|\mathbf{A}^\dagger\|\|\mathbf{A}\psi\|$, which implies $\|\mathbf{A}\| \leqslant \|\mathbf{A}^\dagger\|$. Therefore $\|A\| = \|A^\dagger\|$ is established.

1.4.6 Hermitian operator

When an operator is self-adjoint, namely an adjoint operator $\mathbf{A}^\dagger$ equals to operator $\mathbf{A}$ itself, i.e. $\mathbf{A} = \mathbf{A}^\dagger$, then we call $\mathbf{A}$ a **Hermitian operator**.

1.4.7 Projection operator

Let $\mathcal{H}$ be a Hilbert space in which we consider a **subspace** $\mathcal{M}$ and its orthogonal complement space $\mathcal{M}_\perp$ such that for each vector ψ in $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}_\perp$ that are decomposed into unique vectors $\psi_\mathcal{M}$ in $\mathcal{M}$ and $\psi_{\mathcal{M}_\perp}$ in $\mathcal{M}_\perp$ such that $\psi = \psi_\mathcal{M} + \psi_{\mathcal{M}_\perp}$, and $(\psi_\mathcal{M}, \psi_{\mathcal{M}_\perp}) = 0$.

The projection operator $\mathbf{P}_\mathcal{M}$ when acting upon vector ψ onto a subspace results in $\mathbf{P}_\mathcal{M}\psi = \psi_\mathcal{M}$. It is obvious that $\mathbf{P}_\mathcal{M}\psi = \psi$ if $\psi \in \mathcal{M}$ and $\mathbf{P}_\mathcal{M}\psi = 0$ if $\psi \in \mathcal{M}_\perp$.

One can also be easily convinced that

$$\begin{aligned}(\psi, \mathbf{P}_\mathcal{M}\phi) &= (\psi, \phi_\mathcal{M}) = (\psi_\mathcal{M} + \psi_{\mathcal{M}\perp}, \phi_\mathcal{M}) = (\psi_\mathcal{M}, \phi_\mathcal{M}) \\ &= (\psi_\mathcal{M}, \phi) = (\mathbf{P}_\mathcal{M}\psi_\mathcal{M}, \phi) = (\mathbf{P}_\mathcal{M}\psi, \phi).\end{aligned}$$

Therefore $\mathbf{P}_\mathcal{M}$ is also a Hermitian operator, i.e. $\mathbf{P}^\dagger_\mathcal{M} = \mathbf{P}_\mathcal{M}$.

Similarly we define $\mathbf{P}_{\mathcal{M}_\perp}$ such that $\mathbf{P}_{\mathcal{M}_\perp}\psi = \psi_{\mathcal{M}_\perp}$ and the sum of $\mathbf{P}_\mathcal{M}$ and $\mathbf{P}_{\mathcal{M}_\perp}$ becomes an identity operator, i.e.

$$\mathbf{P}_\mathcal{M} + \mathbf{P}_{\mathcal{M}_\perp} = \mathbf{I}.$$

1.4.8 Idempotent operator

The **projection operator** is an **idempotent operator**, namely $\mathbf{P}^2_\mathcal{M} = \mathbf{P}_\mathcal{M}$ because $\mathbf{P}^2_\mathcal{M}\psi = \mathbf{P}_\mathcal{M}\psi_\mathcal{M} = \mathbf{P}_\mathcal{M}\psi$.

1.5 The postulates of quantum mechanics

We start to formulate the postulates of quantum mechanics. We shall treat the first three postulates in this chapter, and leave the 4th postulate for the next chapter when we investigate the time evolution of a quantum system.

1*st postulate of quantum mechanics:*

For every physical system, there exists an abstract entity, called the state (or the state function or wave function that shall be discussed later), which provides the information of the dynamical quantities of the system; such as coordinates, momenta, energy, angular momentum, charge or isospin, etc. All the states for a given physical system are elements of a Hilbert space, i.e.

physical system	$\longleftrightarrow$	*Hilbert space* $\mathcal{H}$
physical state	$\longleftrightarrow$	*state vector* ψ *in* $\mathcal{H}$

Furthermore for each physical observable, such as the 3rd component of the angular momentum or the total energy of the system and so forth, there associates a unique Hermitian operator in the Hilbert space, i.e.

physical (dynamical) observable		*corresponding hermitean operator*
total energy E	$\longleftrightarrow$	$\mathbf{H} = \mathbf{H}^\dagger$
coordinate $\vec{x}$	$\longleftrightarrow$	$\mathbf{X} = \mathbf{X}^\dagger$
angular momentum $\vec{l}$	$\longleftrightarrow$	$\mathbf{L} = \mathbf{L}^\dagger$

The physical quantity measured in the system for the corresponding observable is obtained by taking the inner product of the pair ψ and $\mathbf{A}\psi$, i.e.

$$\langle \mathbf{A} \rangle = (\psi, \mathbf{A}\psi), \tag{1.24}$$

which is called the **expectation value** of dynamical quantity $\mathbf{A}$ for the system in the state ψ, which is normalized, i.e. $\|\psi\| = 1$.

Since the action of operator $\mathbf{A}$ upon the vector ψ changes it into another vector ϕ, which implies that the action of the measurement of the dynamical quantity in a certain state usually would disturb the physical system and the original state is changed into another state due to the external disturbance accompanying the measurement.

In particular, if an operator $\mathbf{A}$ such that $\mathbf{A}\psi_a = a\psi_a$, i.e. when $\mathbf{A}$ acts upon a particular physical state ψ_a, the resultant state is the same as the one before, then it is said that the physical state is prepared for the measurement of the **dynamical observable** associated with the operator $\mathbf{A}$. We shall name:

- ψ_a : the state particularly prepared in the system for the measurement of the dynamic quantity, called the **eigenstate** of the operator $\mathbf{A}$.
- a : the value of the measurement of the dynamical quantity in the particular prepared state, called the **eigenvalue** of the operator $\mathbf{A}$.

We shall now explore some properties concerning the eigenvectors and the eigenvalues through a few propositions.

Proposition 1.

The eigenvalues for a Hermitian operator are all real.

Let $\mathbf{A}\psi_a = a\psi_a$ and $\mathbf{A}^\dagger\psi_a = a\psi_a$, and consider the inner product $\langle\mathbf{A}\rangle_{\psi_a} = (\psi_a, \mathbf{A}\psi_a) = (\psi_a, a\psi_a) = a(\psi_a, \psi_a) = a$. On the other hand, we have $\langle\mathbf{A}\rangle_{\psi_a} = (\mathbf{A}^\dagger\psi_a, \psi_a) = (\mathbf{A}\psi_a, \psi_a) = a^*(\psi_a, \psi_a) = a^*$ which implies $a = a^*$ if ψ_a is not a null vector.

Proposition 2.

Two eigenvectors of a Hermitian operator are orthogonal to each other if the corresponding eigenvalues are unequal.

Let $\mathbf{A}\psi_a = a\psi_a$ and $\mathbf{A}\psi_b = b\psi_b$ where $\psi_a \neq \psi_b$, and since

$$(\psi_a, \mathbf{A}\psi_b) = b(\psi_a, \psi_b) = (\mathbf{A}^\dagger\psi_a, \psi_b) = a^*(\psi_a, \psi_b) = a(\psi_a, \psi_b),$$