南京航空航天大学
研究生系列精品教材

Advanced Quantum Mechanics

高等量子力学

李晋斌 编著

科学出版社

南京航空航天大学研究生系列精品教材

Advanced Quantum Mechanics

高等量子力学

李晋斌　编著

科学出版社

北　京

内 容 简 介

本书改编自作者在南京航空航天大学讲授 10 年的高等量子力学讲义，内容包括量子力学的数学基础(即希尔伯特空间的基本性质)、量子力学公理、薛定谔方程的近似解法等，课后的习题来自每年的作业和考题. 本书的一大特点是自成体系，尽可能少地涉及本科阶段相关知识，方便自学.

本教材适用于凝聚态、材料、光学等专业相关的学生使用，全书内容对应约 80 课时的教学需要，使用本教材作为参考书的教师可根据自己的教学需求调整.

图书在版编目(CIP)数据

高等量子力学=Advanced Quantum Mechanics: 英文/李晋斌编著. —北京：科学出版社, 2015.8

南京航空航天大学研究生系列精品教材

ISBN 978-7-03-045246-7

Ⅰ. ①高… Ⅱ. ①李… Ⅲ. ①量子力学—研究生—教材—英文 Ⅳ. ①O413.1

中国版本图书馆 CIP 数据核字(2015) 第 170509 号

责任编辑：窦京涛／责任校对：彭 涛

责任印制：徐晓晨／封面设计：迷底书装

科学出版社 出版

北京东黄城根北街 16 号

邮政编码：100717

http://www.sciencep.com

北京教图印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2015 年 8 月第 一 版 开本：787×1092 1/16

2015 年 8 月第一次印刷 印张：14 5/8

字数：337 000

定价：52.00 元

(如有印装质量问题，我社负责调换)

Preface

This book deals with advanced topics in the field of quantum mechanics, material which is usually encountered in graduate student level. The book is written in such a way as to attach importance to a rigorous presentation while, at the same time, requiring no prior knowledge, except in the field of basic quantum mechanics. The inclusion of all mathematical steps and full presentation of intermediate calculations ensures ease of understanding. A number of problems are included at the end of each chapter. Sections or parts thereof that can be omitted in a first reading are marked with a star.

It begins with a rather lengthy chapter in which the relevant mathematics of Hilbert space developed from simple ideas on vectors and matrices the student is assumed to know. The level of rigor is what I think is needed to make a practicing quantum mechanic out of the student. This chapter, which typically takes four to six lecture hours, is filled with examples from physics to keep students from getting too fidgety while they wait for the "real physics". Since the math introduced has to be taught sooner or later, and when they get to it, can give quantum theory their fullest attention without having to battle with the mathematical theorems at the same time. Also, by segregating the mathematical theorems from the physical postulates, any possible confusion as to which is nipped in the bud.

This chapter is followed by one on the postulates, with many examples from fictitious Hilbert spaces. Nonetheless, students will find it hard. It is only as they go along and see these postulates used over and over again in the rest of the book. We also introduce some basic idea and method in quantum mechanics in this chapter.

Chapter 3 introduces the formalism of second quantization and applies this to the most important problems that can be described using simple methods. These include the weakly interacting electron gas and excitations in weakly interacting Bose gases. The basic properties of the correlation and response functions of many-particle systems are also treated here.

We next discuss the Coherent States and Squeezed States, Green's Functions and Scattering Theory and Geometric Phases in different chapters. For the limitation of class hour, we usually choose one or two topic to study.

As I look back to see who all made this book possible, my thoughts first turn to my college Professor. Daning Shi and Professor. Chenping Chu, my PhD. supervisor, Yue Yu, and friends Peng Zhang and Ming Li, for they introduced me to physics in general and quantum mechanics

in particular, and discuss many problems with me. On the family front, encouragement came from my parents and most important of all from my wife, Joan, who cheerfully donated me to science and stood by me throughout. Little Henry did his bit by tearing up all my books on the subject, both as a show of support and to create a need for this one.

The approach to be presented here was tried many times at Nanjing University of Aeronautics and Astronautics on graduates taking a one semester course. In all cases the results were very satisfactory in the sense that the students seemed to have learned the subject well and to have enjoyed the presentation. It is, in fact, their enthusiastic response and encouragement that convinced me of the soundness of my approach and impelled me to write this book. The book also be used by nonphysicists as well. I have found that it goes well with chemistry, nano-science, material science majors in my classes.

Naturally, I am solely responsible for the hopefully few remaining errors and typos, and I invite instructors and students alike to communicate to me any suggestions for improvement, whether they be pedagogical or in reference to errors or misprints.

Jinbin Li

June 2015

Contents

Chapter 1

Mathematical Tools of Quantum Mechanics

Today quantum mechanics forms an important part of our understanding of physical phenomena. Its consequences both at the fundamental and practical levels have intrigued mathematicians, physicists, chemists, and even philosophers for the past century. A quantum system is usually described in terms of certain Hilbert spaces $\mathcal{H}$ and linear operators acting on these spaces. 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. So we will also examine the mathematics involved in representing kets, bras, bra-kets, and operators in discrete and continuous bases.

Certain mathematical topics are essential for quantum mechanics, not only as computational tools, but because they form the most effective language in terms of which the theory can be formulated. We deal with the mathematical machinery needed to study quantum mechanics in this chapter. Although it 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. These topics include the theory of linear vector spaces and linear operators. A unified theory based on that mathematical structure was first formulated by P. A. M. Dirac, and the formulation used in this book is really a modernized version of

Dirac's formalism.

The physical development of quantum mechanics begins in the Chapt.2, and the mathematically sophisticated reader may turn there at once. But since not only the results, but also the concepts and logical framework of this chapter are freely used in developing the physical theory, the reader is advised to at least skim this first chapter before proceeding to next chapter.

1.1 The Hilbert Space

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

(1) A set of vectors $\psi,\ \phi,\ \chi,\ \cdots$ and a set of scalars $a,\ b,\ c,\ \cdots$, if the scalars belong to the field of complex (real) numbers, we speak of a complex (real) linear vector space. Henceforth the scalars will be complex numbers unless otherwise stated.

(2) A rule for vector addition and a rule for scalar multiplication.

1. Addition rule

The addition rule has the properties and structure of an Abelian groups.

(1) If ψ and ϕ are vectors (elements) of a space, their sum $\psi + \phi$, is also a vector of the same space.

(2) Commutativity: $\psi + \phi = \phi + \psi$.

(3) Associativity: $(\psi + \phi) + \chi = \phi + (\psi + \chi)$.

(4) Existence of a zero or neutral vector: for each vector ψ, there must exist a zero vector ϑ such that $\psi + \vartheta = \psi$.

(5) Existence of a symmetric or inverse vector: for each vector ψ, there must exist a symmetric vector ϕ such that $\psi + \phi = \vartheta$. We write ϕ as $-\psi$ later.

2. Multiplication rule

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

(1) The product of a scalar gives another vector. In general, if ψ and ϕ are two vector of the space, any linear combination $a\psi + b\phi$ is also a vector of the space, a and b being scalars.

(2) Distributivity with respect to addition: $\begin{cases} a(\psi + \phi) = a\psi + a\phi, \\ (a + b)\psi = a\psi + b\psi. \end{cases}$

(3) Associativity with respect to multiplication of scalars: $a(b\psi) = (ab)\psi$.

(4) For each element ψ there must exist a unitary element, 1, and zero, 0, scalar such that: $1 \cdot \psi = \psi \cdot 1 = \psi,\quad 0 \cdot \psi = \psi \cdot 0 = \vartheta$.

Examples

Among the very many examples of linear vector spaces, there are two classes that are of common interest:

(1) Discrete vectors, which may be represented as columns of complex numbers, $(a_1, a_2, \cdots)^{\mathrm{T}}$.

(2) Spaces of functions of some type, for example the space of all differentiable functions. One can readily verify that these examples satisfy the definition of a linear vector space.

1.1.1 Definition of Hilbert Space

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

1. $\mathcal{H}$ is a linear space

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

2. $\mathcal{H}$ has a defined scalar (inner) product that is strict positive

The scalar product of an element ψ with another element ϕ is scalar, a complex number, denoted by (ψ, ϕ)=complex number. The scalar product satisfied the following properties.

(1) The scalar product of ψ with ϕ is equal to the complex conjugate of the scalar product of ϕ with ψ: $(\psi, \phi) = (\phi, \psi)^*$.

(2) The scalar product of ψ with respect to $a\phi_1 + b\phi_2$ is

$$(\psi, a\phi_1 + b\phi_2) = a(\psi, \phi_1) + b(\psi, \phi_2)$$

(3) The scalar product of ψ with itself is a non-negative number

$$(\psi, \psi) \equiv \|\psi\|^2 \geqslant 0$$

Where the equality holds only for $\psi = \vartheta$.

3. $\mathcal{H}$ is separable

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

$$\|\psi - \psi_n\| < \varepsilon$$

4. $\mathcal{H}$ is complete

Every Cauchy sequence of element $\psi_n \in \mathcal{H}$ converges to an element of $\mathcal{H}$. That is, for any ψ_n, the relation(or defines a unique limit of $\mathcal{H}$ such that)

$$\lim_{n,m\to\infty} \|\psi_m - \psi_n\| < \varepsilon \quad (\lim_{n\to\infty} \|\psi_n - \psi\| = 0)$$

Examples

We have, corresponding to our previous examples of vector spaces, the following inner products:

(1) If ψ is the column vector with elements $a_1, a_2, \cdots$ and ϕ is the column vector with elements $b_1, b_2, \cdots$, then $(\psi, \phi) = a_1^* b_1 + a_2^* b_2 + \cdots$.

(2) If ψ and ϕ are functions of x, then $(\psi, \phi) = \int \psi^*(x)\phi(x)w(x)\mathrm{d}x$ where $w(x)$ is some nonnegative weight function.

The inner product generalizes the notions of length and angle to arbitrary spaces.

1.1.2 Two Important Theorems

If the inner product of two vectors is zero, the vectors are said to be *orthogonal*. The *norm* (or length) of a vector is defined as $\|\psi\| \equiv (\psi, \psi)^{1/2}$. Norm is written as $|\psi| \equiv (\psi, \psi)^{1/2}$ in some book , and it confuse with symbol of absolute value.

The inner product and the norm satisfy two important theorems.

1. Schwarz's inequality

Schwarz's inequality: $|(\psi, \phi)| \leqslant \|\psi\| \cdot \|\phi\|$.

Proof: for given ψ and ϕ, we can define a new vector $\chi = \psi - (\phi, \psi)\phi/\|\phi\|^2$.

$$0 \leqslant (\chi, \chi) = \|\psi\|^2 - (\psi, \phi)\frac{(\phi, \psi)}{\|\phi\|^2} - \frac{(\phi, \psi)^*}{\|\phi\|^2}(\phi, \psi) + \frac{(\phi, \psi)^*(\phi, \psi)}{\|\phi\|^2\|\phi\|^2}(\phi, \phi)$$

$$= \|\psi\|^2 - \frac{1}{\|\phi\|^2}|(\psi, \phi)|^2 \tag{1.1}$$

$$\Rightarrow |(\psi, \phi)|^2 \leqslant \|\psi\|^2 \cdot \|\phi\|^2 \tag{1.2}$$

So you can, if you like, define the angle between ψ and ϕ by the formula

$$\cos\theta = \frac{|(\psi, \phi)|}{\|\psi\| \cdot \|\phi\|}$$

2. The triangle inequality

The triangle inequality: $\|(\psi + \phi)\| \leqslant \|\psi\| + \|\phi\|$.

Proof: For any complex number, its real part must be not larger than its absolute, saying $\mathrm{Re}z \leqslant |z|$. By using Schwarz's inequality, we have

$$
\begin{aligned}
(\psi+\phi,\psi+\phi) &= \|\psi\|^2 + 2\ \mathrm{Re}(\psi,\phi) + \|\phi\|^2 \leqslant \|\psi\|^2 + 2\ |(\psi,\phi)| + \|\phi\|^2 \\
&\leqslant \|\psi\|^2 + 2\ \|\psi\|\|\phi\| + \|\phi\|^2 = (\|\psi\| + \|\phi\|)^2
\end{aligned}
\tag{1.3}
$$

In both cases equality holds only if one vector is a scalar multiple of the other, i.e. $\psi = c\phi$. For Eq.(1.3) to become an equality, the scalar c must be real and positive.

1.1.3 Dimension and Basis of Vector Space

A set of N vectors $\psi_1, \psi_2, \cdots, \psi_N$ is said to be linear independent if and only if the solution of the equation

$$
\sum_{i=1}^{N} a_i \psi_i = \vartheta, \quad a_1 = a_2 = \cdots = a_N = 0 \tag{1.4}
$$

Be careful ϑ is not 0! But if there exists a set scalars, which are not all zero, so that one of the vector can be expressed as a linear combination of the others $\psi_j = \sum_{i \neq j} a_i' \psi_i$ $(a_i' = -a_i/a_j,\ a_j \neq 0)$, the set of $\{\psi_i\}$ is said to be linear dependent.

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. $\psi_1, \psi_2, \cdots, \psi_N$), this space is said to be N-dimensional. In this case, any vector ϕ of the vector space can be expressed as a linear combination: $\phi = \sum_{i=1}^{N} a_i \psi_i$.

The basis of a vector space consists of a set of the maximum possible number of linearly independent vectors belonging to that space. These vectors $\psi_1, \psi_2, \cdots, \psi_N$ to be denoted in short by $\{\psi_i\}$, are called the *base vectors.* Although the set of these linearly independent vectors is arbitrary, it is convenient to choose them *orthonormal.* A set of vectors $\{\psi_i\}$ is said to be orthonormal if the vectors are pairwise orthogonal and of unit norm; that is to say, their inner products satisfy $(\psi_i, \psi_j) = \delta_{ij}$. 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 are called the *component of the vector ϕ in the basis.* Each component is given by the scalar product of ϕ with the corresponding base vector, $a_i = (\psi_i, \phi)$.

Gram-Schmidt theorem

Theorem Given a linearly independent basis we can form linear combinations of the basis vectors to obtain an orthonormal basis.

Let us now take up the Gram-Schmidt procedure for converting a linearly independent basis into an orthonormal one. The basic idea can be seen by a simple example. Imagine the two-dimensional space of arrows in a plane. Let us take two nonparallel vectors, which qualify as a basis. To get an orthonormal basis out of these, we do the following:

(1) Rescale the first by its own length, so it becomes a unit vector. This will be the first basis vector.

(2) Subtract from the second vector its projection along the first, leaving behind only the part perpendicular to the first (Such a part will remain since by assumption the vectors are nonparallel).

(3) Rescale the left over piece by its own length. We now have the second basis vector: it is orthogonal to the first and of unit length.

1.2 Dual Spaces and the Dirac Notation

Corresponding to any linear vector space V there exists the dual space of linear functionals on V. A linear functional F assigns a scalar $F(\psi)$ to each vector ψ, such that $F(a\psi + b\phi) = aF(\psi) + bF(\phi)$ for any vectors ψ and ϕ, and any scalars a and b. The set of linear functionals may itself be regarded as forming a linear space V' if we define the sum of two functionals as $(F_1 + F_2)(\psi) = F_1(\psi) + F_2(\psi)$.

1.2.1 Riesz Theorem

Theorem There is a one-to-one correspondence between linear functionals F in V' and vectors f in V, such that all linear functionals have the form $F(\psi) = (f, \psi)$, f being a fixed vector, and ψ being an arbitrary vector. Thus the spaces V and V' are essentially isomorphic. For the present we shall only prove this theorem in a manner that ignores the convergence questions that arise when dealing with infinite-dimensional spaces.

Proof: It is obvious that any given vector f in V defines a linear functional, using $F(\psi) = (f, \psi)$ as the definition. So we need only prove that for an arbitrary linear functional F we can construct a unique vector f that satisfies $F(\psi) = (f, \psi)$. Let $\{\psi_i\}$ be a system of orthonormal basis vectors in V, satisfying $(\psi_i, \psi_j) = \delta_{ij}$. Let $\phi = \sum\limits_i x_i\psi_i$ be an arbitrary vector in V.

From $F(a\psi + b\phi) = aF(\psi) + bF(\phi)$ we have $F(\phi) = \sum\limits_i x_i F(\psi_i)$. Now construct the following vector: $f = \sum\limits_i [F(\psi_i)]^* \psi_i$. Its inner product with the arbitrary vector is $(f, \phi) = \sum\limits_i F(\psi_i) x_i = F(\phi)$, and hence the theorem is proved.

1.2.2 Dirac's Bra and Ket Notation

In Dirac's notation, which is very popular in quantum mechanics, the vectors in V are called *ket vectors*, and are denoted as $|\psi\rangle$. The linear functionals in the dual space V' are called *bra vectors*, and are denoted as $\langle F|$. The numerical value of the functional is denoted as $F(\psi) = \langle F|\psi\rangle$. According to the Riesz theorem, there is a one-to-one correspondence between bras and kets. Therefore we can use the same alphabetic character for the functional (a member of V') and the vector (in V) to which it corresponds, relying on the bra, $\langle F|$, or ket, $|F\rangle$, notation to determine which space is referred to. $F(\psi) = (f, \psi)$ would then be written as $\langle F|\psi\rangle = (F, \psi)$, $|F\rangle$ being the vector previously denoted as f. Note, however, that the Riesz theorem establishes, by construction, an anti-linear correspondence between bras and kets. If $\langle F| \leftrightarrow |F\rangle$, then

$$\langle c_1F_1 + c_2F_2| = c_1^*\langle F_1| + c_2^*\langle F_2| \leftrightarrow |F_1\rangle c_1 + |F_2\rangle c_2 = |c_1F_1 + c_2F_2\rangle$$

Because of the relation $\langle F|\psi\rangle = (F, \psi)$, it is possible to regard the "braket" $\langle F|\psi\rangle$ as merely another notation for the inner product. But the reader is advised that there are situations in which it is important to remember that the primary definition of the bra vector is as a linear functional on the space of ket vectors[1].

Let us now take up the Gram-Schmidt procedure to understand Dirac notation. This simple example tells the whole story behind this procedure, which will now be discussed in general terms in the Dirac notation.

Let $|I\rangle, |II\rangle, \cdots$ be a linearly independent basis. The first vector of the orthonormal basis will be

$$|1\rangle = \frac{|I\rangle}{\|I\|} \quad \Rightarrow \quad \langle 1|1\rangle = \frac{\langle I|I\rangle}{\|I\|^2} = 1 \quad (\text{Where} \quad \|I\| = \sqrt{\langle I|I\rangle}\ ) \tag{1.5}$$

As for the second vector in the basis, consider $|2'\rangle = |II\rangle - |1\rangle\langle 1|II\rangle$ which is $|II\rangle$ minus the part pointing along the first unit vector. Not surprisingly it is orthogonal to the latter: $\langle 1|2'\rangle = \langle 1|II\rangle - \langle 1|1\rangle\langle 1|II\rangle = 0$. We now divide $|2'\rangle$ by its norm to get $|2\rangle$ which will be orthogonal to the first and normalized to unity.

Finally, consider $|3'\rangle = |III\rangle - |1\rangle\langle 1|III\rangle - |2\rangle\langle 2|III\rangle$ which is orthogonal to both $|1\rangle$ and $|2\rangle$. Dividing by its norm we get $|3\rangle$, the third member of the orthogonal basis. There is nothing new with the generation of the rest of the basis.

Where did we use the linear independence of the original basis? What if we had started with a linearly dependent basis? Then at some point a vector like $|2'\rangle$ or $|3'\rangle$ would have vanished,

[1] In his original presentation, Dirac assumed a one-to-one correspondence between bras and kets, and it was not entirely clear whether this was a mathematical or a physical assumption. The Riesz theorem shows that there is no need, and indeed no room, for any such assumption. Moreover, we shall eventually need to consider more general spaces (rigged-Hilbert space triplets) for which the one-to-one correspondence between bras and kets does not hold.

putting a stop to the whole procedure. On the other hand, linear independence will assure us that such a thing will never happen since it amounts to having a nontrivial linear combination of linearly independent vectors that adds up the null vector (Go back to the equations for $|2'\rangle$ or $|3'\rangle$ and satisfy yourself that these are linear combinations of the old basis vectors).

Properties of bras, kets and bra-kets

(1) Every ket has a corresponding bra: $|a\psi\rangle = |\psi\rangle a \Leftrightarrow a^* \langle\psi| = \langle a\psi|$.

(2) Properties of inner product: $\langle\phi|\psi\rangle^* = \langle\psi|\phi\rangle$.

(3) Normal is real and non-negative: $\langle\psi|\psi\rangle \geqslant 0$ (when $|\psi\rangle = |\vartheta\rangle$ it equals to 0).

(4) Schwarz's inequality: $|\langle\psi|\phi\rangle|^2 \leqslant \langle\psi|\psi\rangle\langle\phi|\phi\rangle$.

(5) The triangle inequality: $\sqrt{\langle\psi+\phi|\psi+\phi\rangle} \leqslant \sqrt{\langle\psi|\psi\rangle} + \sqrt{\langle\phi|\phi\rangle}$.

(6) Orthonormal: $\langle\psi|\phi\rangle = 0, \quad \langle\psi|\psi\rangle = 1$.

(7) Forbidden Quantities: If $|\psi\rangle$ and $|\phi\rangle$ belong to same vector space, products of $|\psi\rangle|\phi\rangle$ and $\langle\psi|\langle\phi|$ are forbidden. They are nonsense, since $|\psi\rangle|\phi\rangle$ and $\langle\psi|\langle\phi|$ are neither kets nor bras(an explicit illustration of this will be carried out later when discuss the representation in a discrete basis).

If $|\psi\rangle$ and $|\phi\rangle$ belong, however, to different vector spaces (e.g. $|\psi\rangle$ belongs to a spin space and $|\psi\rangle$ to a angular momentum space), then the products of $|\psi\rangle|\phi\rangle$ written as $|\psi\rangle \otimes |\phi\rangle$, represents a tensor product of $|\psi\rangle$ and $|\phi\rangle$. Only in these typical cases are such products meaningful. We will give its other properties in Sec.1.8.

1.3 Operators

General definition

An operator on a vector space maps vectors onto vectors; that is to say, if $\hat{A}$ is an operator and ψ is a vector, then $\phi = \hat{A}\psi$ is another vector in same space. An operator is fully defined by specifying its action on every vector in the space (or in its domain, which is the name given to the subspace on which the operator can meaningfully act, should that be smaller than the whole space).

Examples

(1) Identity operator: leaves any vectors unchanged, $\hat{I}\psi = \psi$.

(2) The parity operator: $\hat{P}_t\psi(\boldsymbol{r}) = \psi(-\boldsymbol{r})$.

(3) The gradient operator: $\hat{\nabla}\psi(\boldsymbol{r}) = \partial_x\psi(\boldsymbol{r})\boldsymbol{i} + \partial_y\psi(\boldsymbol{r})\boldsymbol{j} + \partial_z\psi(\boldsymbol{r})\boldsymbol{k}$.

(4) The linear momentum operator: $\hat{\boldsymbol{P}}\psi(\boldsymbol{r}) = -\mathrm{i}\hbar\nabla\psi(\boldsymbol{r})$.

(5) The Laplacian operator: $\Delta\psi(\boldsymbol{r}) = \partial_x^2\psi(\boldsymbol{r}) + \partial_y^2\psi(\boldsymbol{r}) + \partial_z^2\psi(\boldsymbol{r})$.

(6) Inverse operators, assuming it exists①, the *inverse* $\hat{A}^{-1}$ of a linear operator $\hat{A}$ is defined by the relation: $\hat{A}^{-1}\hat{A} = \hat{A}\hat{A}^{-1} = \hat{I}$.

1.3.1 Linear Operator

A linear operator satisfies $\hat{A}(c_1\psi_1 + c_2\psi_2) = c_1(\hat{A}\psi_1) + c_2(\hat{A}\psi_2)$. It is sufficient to define a linear operator on a set of basis vectors, since every vector can be expressed as a linear combination of the basis vectors. We shall be treating only linear operators, and so shall henceforth refer to them simply as operators.

To assert the equality of two operators, $\hat{A} = \hat{B}$, means that $\hat{A}\psi = \hat{B}\psi$ for all vectors (more precisely, for all vectors in the common domain of $\hat{A}$ and $\hat{B}$, this qualification will usually be omitted for brevity). Thus we can define the sum and product of operators, $(\hat{A} + \hat{B})\psi = \hat{A}\psi + \hat{B}\psi$, $\hat{A}\hat{B}\psi = \hat{A}(\hat{B}\psi)$, both equations holding for all. It follows from this definition that operator multiplication is necessarily associative, $\hat{A}(\hat{B}\hat{C}) = (\hat{A}\hat{B})\hat{C}$. But it need not be commutative, $\hat{A}\hat{B}$ being unequal to $\hat{B}\hat{A}$ in general.

Example 1 In a space of discrete vectors represented as columns, a linear operator is a square matrix. In fact, any operator equation in a space of N dimensions can be transformed into a matrix equation. Consider, for example, the equation $\hat{M}|\psi\rangle = |\phi\rangle$. Choose some orthonormal basis $\{|u_i\rangle,\ \ i = 1, \cdots, N\}$ in which to expand the vectors, $|\psi\rangle = \sum_j a_j|u_j\rangle$, $|\phi\rangle = \sum_j b_j|u_j\rangle$.

Operating on $\hat{M}|\psi\rangle = |\phi\rangle$ with $\langle u_i|$ yields $\sum_j \langle u_i|\hat{M}|u_j\rangle a_j = \sum_k \langle u_i|u_k\rangle b_k = b_i$, which has the form of a matrix equation, $\sum_j M_{ij}a_j = b_i$, with $M_{ij} = \langle u_i|M|u_j\rangle$ being known as a matrix element of the operator M. In this way any problem in an N-dimensional linear vector space, no matter how it arises, can be transformed into a matrix problem.

The same thing can be done formally for an infinite-dimensional vector space if it has a denumerable orthonormal basis, but one must then deal with the problem of convergence of the infinite sums, which we postpone to a later section.

Example 2 Operators in function spaces frequently take the form of differential or integral operators. An operator equation such as $\partial_x x = 1 + x\partial_x$ may appear strange if one forgets that operators are only defined by their action on vectors. Thus the above example means that $\partial_x[x\psi(x)] = \psi(x) + x\partial_x\psi(x)$.

① Not every operator has inverse, just as in the case of matrices. The inverse of a matrix exists only when its determinant is nonzero.

1.3.2 Hermitian Adjoint

So far we have only defined operators as acting to the right on ket vectors. We may define their action to the left on bra vectors as $(\langle\phi|\hat{A})|\psi\rangle = \langle\phi|(\hat{A}|\psi\rangle)$ for all ψ and ϕ. This appears trivial in Dirac's notation, and indeed this triviality contributes to the practical utility of his notation. However, it is worthwhile to examine the mathematical content of last equation in more detail.

A bra vector is in fact a linear functional on the space of ket vectors, and in a more detailed notation the bra $\langle\phi|$ is the functional $F_\phi(\cdot) = (\phi, \cdot)$, where ϕ is the vector that corresponds to F_ϕ via the Riesz theorem, and the dot indicates the place for the vector argument. We may define the *operation of $\hat{A}$ on the bra space of functionals* as $\hat{A}F_\phi(\psi) = F_\phi(\hat{A}\psi)$ for all ψ. The right hand side of this formula satisfies the definition of a linear functional of the vector ψ (not merely of the vector $\hat{A}\psi$), and hence it does indeed define a new functional, called $\hat{A}F_\phi$. According to the Riesz theorem there must exist a ket vector χ such that

$$\hat{A}F_\phi(\psi) = (\chi, \psi) = F_\chi(\psi) \tag{1.6}$$

Since χ is uniquely determined by ϕ (given $\hat{A}$), there must exist an operator $\hat{A}^\dagger$ such that $\chi = \hat{A}^\dagger\phi$. Thus Eq.(1.6) can be written as $\hat{A}F_\phi = F_{\hat{A}^\dagger\phi}$. From $\hat{A}F_\phi(\psi) = F_\phi(\hat{A}\psi)$ and Eq.(1.6) we have $(\phi, \hat{A}\psi) = (\chi, \psi)$, and therefore

$$(\hat{A}^\dagger\phi, \psi) = (\phi, \hat{A}\psi), \quad \text{for all } \phi \text{ and } \psi \tag{1.7}$$

This is the usual definition of the adjoint, $\hat{A}^\dagger$, of the operator $\hat{A}$. All of this nontrivial mathematics is implicit in Dirac's simple equation $(\langle\phi|\hat{A})|\psi\rangle = \langle\phi|(\hat{A}|\psi\rangle)$!

The *adjoint operator* can be formally defined within the Dirac notation by demanding that if $\langle\phi|$ and $|\phi\rangle$ are corresponding bras and kets, then $\langle\phi|\hat{A}^\dagger \equiv \langle\chi|$ and $\hat{A}|\phi\rangle \equiv |\chi\rangle$ should also be corresponding bras and kets. From the fact that $\langle\chi|\psi\rangle^* = \langle\psi|\chi\rangle$, it follows that

$$\langle\phi|\hat{A}^\dagger|\psi\rangle^* = \langle\psi|\hat{A}|\phi\rangle, \quad \text{for all } \phi \text{ and } \psi \tag{1.8}$$

this relation being equivalent to Eq.(1.7). Although simpler than the previous introduction of $\hat{A}^\dagger$ via the Riesz theorem, this formal method fails to prove the existence of the operator $\hat{A}^\dagger$.

Several useful properties of the adjoint operator that follow directly from Eq.(1.7) are $(c\hat{A})^\dagger = c^*\hat{A}^\dagger$, where c is a complex number, $(\hat{A}+\hat{B})^\dagger = \hat{A}^\dagger + \hat{B}^\dagger$, $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$. In addition to the inner product of a bra and a ket, $\langle\phi|\psi\rangle$, which is a scalar, we may define an *outer product (Dyadic product)*, which is the formal product between a ket-and a bra-vector, $|\psi\rangle\langle\phi|$. This object is an operator because, assuming associative multiplication, we have $(|\psi\rangle\langle\phi|)|\cdot\rangle = |\psi\rangle(\langle\phi|\cdot\rangle)$, it projects the vector onto the state $|\phi\rangle$ and generates a new vector in parallel to $|\psi\rangle$ with a magnitude equal to the projection $\langle\phi|\cdot\rangle$. Since an operator is defined by specifying its action on