

Physics 221A
Fall 2016
Notes 1
The Mathematical Formalism of Quantum Mechanics

1. Introduction

The prerequisites for Physics 221A include a full year of undergraduate quantum mechanics. In this semester we will survey that material, organize it in a more logical and coherent way than the first time you saw it, and pay special attention to fundamental principles. We will also present some new material. Physics 221B will largely consist of new material.

We begin by presenting some of the mathematical formalism of quantum mechanics. We will introduce more mathematics later in the course as needed, but for now we will concentrate on the linear algebra of spaces of wave functions, which are called *Hilbert spaces*. These notes gather together and summarize most of what you will need to know of this subject, so we can proceed with other matters. In the next set of notes we will turn to the physical postulates of quantum mechanics, which allow us to connect experimental results with the mathematics presented here.

Introductory courses on linear algebra are usually limited to finite-dimensional, real vector spaces. Making the vector spaces complex is a small change, but making them infinite-dimensional is a big step if one wishes to be rigorous. We will make no attempt to be rigorous in the following—to do so would require more than one course in mathematics and leave no time for the physics. Instead, we will follow the usual procedure in physics courses when encountering new mathematics, which is to proceed by example and analogy, attempting to gain an intuitive understanding for some of the main ideas without going into technical proofs. Specifically, in dealing with Hilbert spaces we will try to apply what we know about finite-dimensional vector spaces to the infinite-dimensional case, often using finite-dimensional intuition in infinite dimensions, and we will try to learn where things are different in infinite dimensions and where one must be careful. Fortunately, it is one of the consequences of the mathematical definition of a Hilbert space that many of its properties are the obvious generalizations of those that hold in finite dimensions, so much of finite-dimensional intuition does carry over. We will use what we know about spaces of wave functions (an example of a Hilbert space) to gain some intuition about where this is not so.

2. Hilbert Spaces

To orient ourselves, let us consider a wave function $\psi(x)$ for a one-dimensional quantum mechanical problem. In practice it is common to require $\psi(x)$ to be normalized, but for the purpose of

the following discussion we will require only that it be normalizable, that is, that

$$\int |\psi(x)|^2 dx < \infty \quad (1)$$

(which means that the integral is finite). Wave functions that are not normalizable cannot represent physically realizable states, because the probability of finding a real particle somewhere in space must be unity. Nevertheless, wave functions that are not normalizable, such as plane waves, are definitely useful for some purposes, and we will have to say something about them later. For now, however, we will stick to normalizable wave functions.

Mathematically speaking, the space of complex functions that are normalizable (or *square integrable*) in the sense of Eq. (1) constitutes a complex vector space. This is because if $\psi(x)$ is square-integrable, then so is $c\psi(x)$, where c is any complex number, and if $\psi_1(x)$ and $\psi_2(x)$ are square-integrable, then so is $\psi_1(x) + \psi_2(x)$. Moreover, the space is infinite-dimensional (see Sec. 28). This vector space is an example of a *Hilbert space*; in general, a Hilbert space is a complex, inner product vector space (a vector space upon which an inner or scalar product is defined) with certain additional properties that will not concern us in this course (see Sec. 28). Often the term “Hilbert space” is defined to be an infinite-dimensional space, but in this course we will refer to any of the vector spaces of wave functions that occur in quantum mechanics as Hilbert spaces, even when finite-dimensional.

As you know, given a normalized wave function $\psi(x)$ in configuration space, $|\psi(x)|^2$ is the probability density for making measurements of position on an ensemble of systems. The wave function $\psi(x)$ is connected physically with position measurements. (See Sec. B.3 for the definition of “configuration space.”)

Given $\psi(x)$ in configuration space, one can compute the corresponding wave function $\phi(p)$ in momentum space, by a version of the Fourier transform:

$$\phi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int dx e^{-ipx/\hbar} \psi(x). \quad (2)$$

The Fourier transform is linear and invertible (via the inverse Fourier transform), so provides a one-to-one, linear map between the spaces of wave function $\psi(x)$ and that of $\phi(p)$. Moreover, according to the Parseval identity, the norm of ψ in configuration space and the norm of ϕ in momentum space are equal,

$$\int dx |\psi(x)|^2 = \int dp |\phi(p)|^2, \quad (3)$$

so if we consider the Hilbert space of normalizable wave functions $\{\psi(x)\}$ in configuration space, then the set $\{\phi(p)\}$ of corresponding momentum space wave functions also forms a Hilbert space. (The “norm” of a wave function is the square root of the quantity shown in Eq. (3); it is like the length of a vector.)

The momentum space wave function $\phi(p)$ is connected with the measurement of momentum, in particular, if $\phi(p)$ is normalized then $|\phi(p)|^2$ is the probability density for momentum measurements

on an ensemble of systems. The Parseval identity expresses the equality of the total probability for measurements of either position or momentum.

Likewise, let $\{u_n(x), n = 1, 2, \dots\}$ be an orthonormal set of energy eigenfunctions, $Hu_n(x) = E_n u_n(x)$ for some Hamiltonian H . An arbitrary wave function $\psi(x)$ can be expanded as a linear combination of the functions $u_n(x)$, that is, we can write

$$\psi(x) = \sum_n c_n u_n(x), \quad (4)$$

where the expansion coefficients c_n are given by

$$c_n = \int dx u_n(x)^* \psi(x). \quad (5)$$

Equation (4) can be regarded as a linear transformation taking us from the sequence of coefficient (c_1, c_2, \dots) to the wave function $\psi(x)$, while Eq. (5) goes the other way.

The sequence of coefficients (c_1, c_2, \dots) uniquely characterizes the state of the system, and we can think of it as the “wave function in energy space.” Notice that if the sequence is normalized, $\sum_n |c_n|^2 = 1$, then the probability of making an measurement $E = E_n$ of energy is $|c_n|^2$. Also, the norm of the original wave function $\psi(x)$ can be expressed in terms of the expansion coefficients by

$$\int dx |\psi(x)|^2 = \sum_n |c_n|^2, \quad (6)$$

which is finite if $\psi(x)$ belongs to Hilbert space. This is another expression of conservation of probability. Mathematically speaking, the set of sequences $\{(c_1, c_2, \dots)\}$ of complex numbers of finite norm is yet another example of a Hilbert space.

Thus, we have three Hilbert spaces, the two spaces of wave functions, $\{\psi(x)\}$ and $\{\phi(p)\}$, and the space of sequences $\{(c_1, c_2, \dots)\}$, all assumed to be normalizable. These Hilbert spaces are isomorphic, in the sense that the vectors of each space are related to one another by invertible, norm-preserving linear transformations, so they all contain the same information. (See Sec. 28 for some mathematical points regarding Hilbert spaces of wave functions.) Therefore none of these spaces can be taken as more fundamental than the other; any calculation that can be carried out in one space (or in one *representation*, as we will say), can be carried out in another.

Psychologically, there is a tendency to think of wave functions $\psi(x)$ on configuration space as fundamental, probably because we are used to thinking of fields such as electric fields defined over physical space. This is also the bias imposed on us by our first courses in quantum mechanics, as well as the historical development of the subject. Also, we live in physical space, not momentum space. But the Schrödinger wave function is not really defined over physical space, rather it is defined over *configuration space*, which is the same only in the case of a single particle (see Sec. B.3). Furthermore, completely apart from the mathematical equivalence of the different Hilbert spaces discussed above, the physical postulates of quantum mechanics show that the different wave functions discussed above correspond to the measurement of different sets of physical observables. They also show that there

is a kind of democracy among the different physical observables one can choose to measure, as long as the set of observables is compatible and complete in a sense we will discuss later.

Finally, even if we do have a single particle moving in three dimensions so the wave function $\psi(\mathbf{r})$ can be thought of as a field over physical space, the physical measurement of $\psi(\mathbf{r})$ is very different from the measurement of a classical field such as the electric field or the pressure of a fluid. We have thermometers that can measure the temperature at a point, we have pressure gauges, and it is possible to measure electric and magnetic fields at points of space. But there is no such thing as a “ ψ -meter” for measuring the wave function $\psi(\mathbf{r})$. Later we will discuss how one can obtain $\psi(\mathbf{r})$ from measurements, but the process involves several subtleties. (In particular, it requires one to understand the difference between a pure and a mixed state; only pure states have wave functions.) It is a very different business than the measurement of a classical field.

In summary, it is probably best to think of the different wave functions as equivalent descriptions of the same physical reality, none of which is more fundamental than the other. This leads to the concept of the state space of a quantum mechanical system as an abstract vector space in which the state vector lies, while the various wave functions are the expansion coefficients of this state vector with respect to some basis. Different bases give different wave functions, but the state vector is the same.

This was not the point of view in the early days of quantum mechanics, when $\psi(\mathbf{r})$ was seen as a complex-valued field on physical space, and before it was completely clear what the wave function of multiparticle systems should be. We will return to this earlier point of view when we take up the Dirac equation and field theory, but that reconsideration will not change the basic concept of the state space for the quantum mechanical system, a complex vector space in which the state vector lies.

Thus, the mathematical formalism of quantum mechanics exhibits a kind of form invariance or *covariance* with respect to the set of observables one chooses to represent the states of a system. This view of quantum mechanics, and the transformation theory connected with changes of representation, was worked out in the early days of the quantum theory, notably by Dirac, Born, Jordan, von Neumann and others.

3. Dirac Ket Notation

Since none of the three Hilbert spaces has any claim over the other, it is preferable to use a notation that does not prejudice the choice among them. This is what the Dirac notation does. In the Dirac notation, a state of a system is represented by a so-called *ket vector*, denoted by $|\ \rangle$, which stands for any of the three wave functions discussed above, $\psi(x)$, $\phi(p)$, or $\{c_n\}$, assuming that they are related by the transformations (2) and (5). It is customary to put a symbol into the ket to identify the state, to write, for example, $|\psi\rangle$ (although the choice of the letter ψ seems to point to something special about the configuration space wave function $\psi(x)$, which there is not). Some people think of a ket $|\psi\rangle$ as simply another notation for a wave function on configuration space, and

I have even seen some authors write equations like

$$\psi(x) = |\psi\rangle. \quad (7)$$

But our point of view will be that the space of kets is yet another Hilbert space, an abstract space that is isomorphic to but distinct from any of the concrete Hilbert spaces of wave functions discussed above. Thus, we will never use equations like (7). I think this approach is appropriate in view of the physical postulates of quantum mechanics, which show how one can construct a Hilbert space out of the results of physical measurements. There is no need to start with wave functions. Therefore we will now put wave functions aside, and discuss the mathematical properties of ket spaces. Later, in Notes 4, we will return to wave functions, in effect deriving them from the ket formalism. In this set of notes, we use wave functions only for the purpose of illustrating some general features of Hilbert spaces, calling on your experience with them in a previous course in quantum mechanics.

4. Kets, Ket Spaces and Rays

Kets and ket spaces arise out of the physical postulates of quantum mechanics, which we will discuss in detail later. For now we will just say enough to get started.

A basic postulate of quantum mechanics is that a given physical system is associated with a certain complex vector space, which we will denote by \mathcal{E} . The physical principles by which this vector space is constructed will be discussed later, but for now we simply note that the properties of this space such as its dimensionality are determined by the physics of the system under consideration. For some systems, \mathcal{E} is finite-dimensional (notably spin systems), while for others it is infinite-dimensional. The vectors of the space \mathcal{E} are called *kets*, and they are denoted in the Dirac notation by $|\psi\rangle$, as discussed above. Since \mathcal{E} is a complex vector space, the multiplication of kets by complex numbers and the addition of kets is defined. The product of a complex number c and a ket $|\psi\rangle$ can be written in either order,

$$c|\psi\rangle = |\psi\rangle c. \quad (8)$$

In addition, there exist physical principles that allow us to associate a definite (pure) state of a physical system with a ray in \mathcal{E} . We will say more about this association later, when we will define what we mean by a pure state of a physical system, but for now we simply note that a ray in \mathcal{E} is a 1-dimensional vector subspace of \mathcal{E} , that is, it consists of kets $|\psi\rangle$ related to some nonzero ket $|\psi_0\rangle$ by a complex multiplicative factor, that is,

$$|\psi\rangle = c|\psi_0\rangle, \quad (9)$$

for some complex number c . The kets $|\psi\rangle$ in the ray differ from one another only by normalization and overall phase, which have no physical significance; thus, we say that a (pure) physical state corresponds to the ray (not to any particular ket in the ray).

5. The Dual Space and Bras

You are probably familiar with the fact that kets correspond to wave functions $\psi(x)$, while bras correspond to complex conjugated wave functions $\psi(x)^*$. But if we are doing without wave functions until we can properly derive them, how do we describe bras and the complex conjugation of kets? The answer is the dual correspondence.

Also, it would appear that $\psi(x)^*$ is just another wave function, so it should belong to the same space of wave functions as $\psi(x)$. But it turns out that there is great advantage in regarding kets and bras as belonging to two different spaces. Kets belong to the space \mathcal{E} , while bras belong to the so-called *dual space*, denoted \mathcal{E}^* .

If \mathcal{E} is any (complex) vector space, then the *dual space*, denoted \mathcal{E}^* , is the space of complex-valued, linear operators that act on \mathcal{E} . In the Dirac notation, such operators are called *bras*, and are denoted $\langle \cdot |$. (In mathematics, they are called *forms*, *covectors* or *dual vectors*.) As with kets, it is customary to insert some symbol to identify a bra and distinguish it from other bras, such as $\langle \alpha |$. In mathematical language, we write

$$\langle \alpha | : \mathcal{E} \rightarrow \mathbb{C}, \quad (10)$$

which means that $\langle \alpha |$ maps a ket into a complex number. To be a bra, the map must be linear. If $|\psi\rangle$ is a ket, then the complex number that results by letting $\langle \alpha |$ act on it can be written $\langle \alpha | (|\psi\rangle)$. Usually we drop the parentheses write the complex number simply as $\langle \alpha | \psi \rangle$, but if we wish to emphasize the interpretation of the bra as an operator acting on kets we will keep the parentheses.

Since the map is required to be linear, we have

$$\langle \alpha | (c_1|\psi_1\rangle + c_2|\psi_2\rangle) = c_1\langle \alpha | (|\psi_1\rangle) + c_2\langle \alpha | (|\psi_2\rangle). \quad (11)$$

A bra, as a linear operator acting on kets, is not to be confused with the familiar linear operators used in quantum mechanics, such as the Hamiltonian operator, which also act on kets. The distinction is that a bra is a complex-valued operator, whereas the Hamiltonian and other such operators are ket-valued operators.

It is easy to see how to define the product of a bra with a complex scalar, and the sum of two bras; we simply set

$$(c\langle \alpha |)(|\psi\rangle) = c\langle \alpha | (|\psi\rangle), \quad (12)$$

and

$$(\langle \alpha_1 | + \langle \alpha_2 |)(|\psi\rangle) = \langle \alpha_1 | (|\psi\rangle) + \langle \alpha_2 | (|\psi\rangle). \quad (13)$$

Therefore the set of all bras acting on a given ket space forms a vector space in its own right, which by definition is the dual space \mathcal{E}^* . It is easy to show that if \mathcal{E} is finite-dimensional, then \mathcal{E}^* is also, and has the same dimension. If \mathcal{E} is infinite-dimensional, then the dual space \mathcal{E}^* is also infinite-dimensional.

What we have said so far allows us to talk about bras, but not to convert a ket into a bra (the analog of the complex conjugation of wave functions). To do that we need to introduce the metric

or scalar product on our ket space. In real, Euclidean vector spaces the metric is what allows us to measure distances, and from that, angles. This is an analogy to keep in mind while thinking about the metric in (complex) Hilbert spaces, because it provides useful geometrical interpretations.

6. The Scalar Product and Dual Correspondence

We postulate the existence of a metric or scalar product g on \mathcal{E} , which is a function

$$g : \mathcal{E} \times \mathcal{E} \rightarrow \mathbb{C} \quad (14)$$

with certain properties. This notation means that g is a function that takes two kets and produces a complex number; for example, we will write $g(|\psi\rangle, |\phi\rangle)$ for the complex number that results upon letting g act on kets $|\psi\rangle$ and $|\phi\rangle$. The postulated properties of g are the following. First, g is linear in its second operand and antilinear in its first operand. Explicitly, this means

$$g(|\psi\rangle, c_1|\phi_1\rangle + c_2|\phi_2\rangle) = c_1g(|\psi\rangle, |\phi_1\rangle) + c_2g(|\psi\rangle, |\phi_2\rangle), \quad (15a)$$

$$g(c_1|\psi_1\rangle + c_2|\psi_2\rangle, |\phi\rangle) = c_1^*g(|\psi_1\rangle, |\phi\rangle) + c_2^*g(|\psi_2\rangle, |\phi\rangle). \quad (15b)$$

(An antilinear function requires us to take the complex conjugate of the coefficients when evaluating it on linear combinations of vectors.) Next, g is symmetric or Hermitian, in the following sense:

$$g(|\psi\rangle, |\phi\rangle) = g(|\phi\rangle, |\psi\rangle)^*. \quad (16)$$

Third, g is positive definite, which means

$$g(|\psi\rangle, |\psi\rangle) \geq 0, \quad (17)$$

for all $|\psi\rangle$, with equality holding if and only if $|\psi\rangle = 0$. Note that by property (16), the left hand side of Eq. (17) is necessarily real.

Next, given the metric g , we can define the *dual correspondence*, which is a mapping that converts kets into bras. It is the map which in wave function language amounts to taking the complex conjugate of a wave function. We will denote the dual correspondence by DC; it is the map

$$\text{DC} : \mathcal{E} \rightarrow \mathcal{E}^*. \quad (18)$$

If $|\psi\rangle$ is a ket, then we will denote the bra corresponding to it under the dual correspondence by $\langle\psi|$, with the same identifying symbol. We define the bra $\langle\psi|$ in terms of its action on an arbitrary ket $|\phi\rangle$; the definition is

$$\langle\psi|(|\phi\rangle) = g(|\psi\rangle, |\phi\rangle). \quad (19)$$

The idea is that in the scalar product, we regard the first argument as fixed and the second as variable, which gives us a complex valued function of kets, which according to Eq. (15a) is linear. Thus, the function is a bra. By way of notation, we will denote the bra resulting from a given ket under the dual correspondence by a dagger, so that if

$$\text{DC} : |\psi\rangle \mapsto \langle\psi|, \quad (20)$$

then we will write

$$\langle\psi| = (|\psi\rangle)^\dagger. \quad (21)$$

We will say that $\langle\psi|$ is the *Hermitian conjugate* of $|\psi\rangle$. We note that since the scalar product is antilinear in its first operand, the dual correspondence is antilinear, that is,

$$(c_1|\psi_1\rangle + c_2|\psi_2\rangle)^\dagger = c_1^* \langle\psi_1| + c_2^* \langle\psi_2|. \quad (22)$$

If \mathcal{E} is finite-dimensional, then it is easy to prove that the dual correspondence is one-to-one and onto, so that every bra is in fact the Hermitian conjugate of some ket. This follows from the postulated properties of the metric. In infinite-dimensional spaces, this is true only if certain restrictions are placed on the bra space. We will ignore such technicalities, and proceed as if the dual correspondence is always one-to-one and onto (which, with the right understandings, it is for a Hilbert space). This means that every bra is the dagger of some unique ket, so the dual correspondence has an inverse. We will denote the inverse with the same dagger notation, so that

$$|\psi\rangle = (\langle\psi|)^\dagger, \quad (23)$$

and so that

$$(|\psi\rangle)^\dagger{}^\dagger = |\psi\rangle. \quad (24)$$

We now simplify the notation for the scalar product by dropping the parentheses on the left hand side of Eq. (19). That is, we write

$$\langle\psi|(|\phi\rangle) = g(|\psi\rangle, |\phi\rangle) = \langle\psi|\phi\rangle. \quad (25)$$

With this notational change, we can rewrite properties (16) and (17) of the metric as

$$\langle\psi|\phi\rangle = \langle\phi|\psi\rangle^*, \quad (26)$$

and

$$\langle\psi|\psi\rangle \geq 0, \quad (27)$$

for all $|\psi\rangle$, with equality if and only if $|\psi\rangle = 0$. The complex number $\langle\psi|\phi\rangle$ is sometimes called the *inner product* or *scalar product* of $|\psi\rangle$ and $|\phi\rangle$. Using the notation shown in Eq. (25), we can henceforth dispense with the g -notation when talking about scalar products.

7. The Schwarz Inequality

We can now prove an important theorem, namely the *Schwarz inequality*. We are interested in this inequality for complex vector spaces, but it is also true in real vector spaces with a positive definite inner product, that is, in Euclidean spaces. There it is equivalent to the geometrical fact that the shortest distance between two points is a straight line. The Schwarz inequality is important in quantum mechanics because it is used to prove the Heisenberg uncertainty relations.

The Schwarz inequality says that

$$|\langle\psi|\phi\rangle|^2 \leq \langle\psi|\psi\rangle\langle\phi|\phi\rangle, \quad (28)$$

for all $|\psi\rangle$ and $|\phi\rangle$, with equality if and only if $|\psi\rangle$ and $|\phi\rangle$ are linearly dependent, that is, if they lie in the same ray. To prove the theorem, we set

$$|\alpha\rangle = |\psi\rangle + \lambda|\phi\rangle, \quad (29)$$

where λ is a complex number, and we use Eq. (27) to write,

$$\langle\alpha|\alpha\rangle = \langle\psi|\psi\rangle + \lambda\langle\psi|\phi\rangle + \lambda^*\langle\phi|\psi\rangle + |\lambda|^2\langle\phi|\phi\rangle \geq 0. \quad (30)$$

This is true for all kets $|\psi\rangle$ and $|\phi\rangle$, and all complex numbers λ . Since the Schwarz inequality is obviously true if $|\phi\rangle = 0$, we consider the case $|\phi\rangle \neq 0$, and we set

$$\lambda = -\frac{\langle\phi|\psi\rangle}{\langle\phi|\phi\rangle}, \quad (31)$$

so that Eq. (30) becomes

$$\langle\psi|\psi\rangle - \frac{|\langle\psi|\phi\rangle|^2}{\langle\phi|\phi\rangle} \geq 0, \quad (32)$$

which is equivalent to Eq. (28). Finally, it is easy to show that Eq. (32) becomes an equality if $|\psi\rangle = c|\phi\rangle$ for any complex number c , and conversely, if Eq. (32) is an equality, then $\langle\alpha|\alpha\rangle = 0$, which implies $|\alpha\rangle = 0$, or $|\psi\rangle = -\lambda|\phi\rangle$. Therefore the equality holds in Eq. (28) if and only if $|\psi\rangle$ and $|\phi\rangle$ are linearly dependent.

8. Operators

Next we discuss *operators*. Usually in quantum mechanics we use only linear or antilinear operators. A linear operator L and an antilinear operator A are both mappings from the ket space into itself,

$$L : \mathcal{E} \rightarrow \mathcal{E}, \quad A : \mathcal{E} \rightarrow \mathcal{E}, \quad (33)$$

but they have different properties when acting on linear combinations of kets. In particular, we have

$$L(c_1|\psi_1\rangle + c_2|\psi_2\rangle) = c_1L|\psi_1\rangle + c_2L|\psi_2\rangle, \quad (34a)$$

$$A(c_1|\psi_1\rangle + c_2|\psi_2\rangle) = c_1^*A|\psi_1\rangle + c_2^*A|\psi_2\rangle. \quad (34b)$$

The only antilinear operator of interest in nonrelativistic quantum mechanics is the time-reversal operator, which we will discuss later. For now we ignore antilinear operators, and concentrate exclusively on linear operators.

Linear operators themselves can be multiplied by complex numbers and added, so they form a complex vector space in their own right. Linear operators can also be multiplied with one another; the product AB means, apply B first, then A (to some ket). The multiplication is associative,

$$A(BC) = (AB)C, \quad (35)$$

for any linear operators A, B, C , but it is not commutative, that is

$$AB \neq BA \quad (36)$$

in general. The lack of commutativity of two linear operators A and B is measured by their *commutator*, defined by

$$[A, B] = AB - BA. \quad (37)$$

For reference, we also define the *anticommutator*,

$$\{A, B\} = AB + BA. \quad (38)$$

We use the same curly bracket notation for the anticommutator as for the Poisson bracket in classical mechanics (see Sec. B.21). It will usually be clear from context which is intended.

A linear operator A may possess an inverse, A^{-1} ; if it exists, it satisfies

$$AA^{-1} = A^{-1}A = 1, \quad (39)$$

where the identity operator is simply denoted by 1. In finite-dimensional spaces, if A^{-1} exists, then it is both a right- and a left-inverse; but in infinite dimensional spaces, some operators may possess a right-inverse but not a left-inverse, or vice versa. When all indicated inverses exist, we have

$$(AB)^{-1} = B^{-1}A^{-1}. \quad (40)$$

9. Rules for Commutators

The commutator $[,]$ obeys the following properties, which are trivial consequences of the definition (37). In the following, capital letters are linear operators and lower case letters are complex numbers. First, the commutator is linear in both operands,

$$\begin{aligned} [c_1A_1 + c_2A_2, B] &= c_1[A_1, B] + c_2[A_2, B], \\ [A, c_1B_1 + c_2B_2] &= c_1[A, B_1] + c_2[A, B_2], \end{aligned} \quad (41)$$

it is antisymmetric,

$$[A, B] = -[B, A], \quad (42)$$

and it obeys the *Jacobi identity*,

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0. \quad (43)$$

Any bracket operation $[,]$ defined on any vector space (not just spaces of linear operators) that satisfies properties (41)–(43) qualifies that vector space as a *Lie algebra*.

In addition, the commutator satisfies the following properties, sometimes referred to as the derivation property or Leibnitz rule:

$$\begin{aligned} [AB, C] &= A[B, C] + [A, C]B, \\ [A, BC] &= B[A, C] + [A, B]C. \end{aligned} \quad (44)$$

Calculations in quantum mechanics often require one to reduce complicated commutators into simpler ones that are known. Rules (44) are especially useful for this purpose. It is interesting that properties (41)–(44) are also valid for the Poisson bracket in classical mechanics (except that the ordering of the factors in Eq. (44), which must be respected in quantum mechanics, is immaterial in classical mechanics). See Sec. B.21.

10. The Action of Operators on Bras

The operators we have introduced begin life as operators that act on kets, mapping them into other kets; but the definition of these operators is easily extended to allow them to act on bras, whereupon they produce other bras. If $\langle\psi|$ is a bra and A an operator, we denote the action of A on $\langle\psi|$ by $\langle\psi|A$, with the bra to the left. That is, we think of A as acting on kets to the right, but on bras to the left. The definition of $\langle\psi|A$ is as follows. Since $\langle\psi|A$ is supposed to be a new bra, it is specified by its action on an arbitrary ket, say, $|\phi\rangle$; the definition is given by writing

$$(\langle\psi|A)(|\phi\rangle) = \langle\psi|(A|\phi\rangle). \quad (45)$$

Since by this definition the ordering of the parentheses is immaterial, it is customary to drop them, and to write simply

$$(\langle\psi|A)(|\phi\rangle) = \langle\psi|A|\phi\rangle. \quad (46)$$

Thus, we can think of A as acting either to the left or the right in the “matrix element” $\langle\psi|A|\phi\rangle$.

11. The Outer Product

If $|\alpha\rangle$ and $|\beta\rangle$ are kets, then we can define a linear operator denoted by $|\alpha\rangle\langle\beta|$, whose action on an arbitrary ket $|\psi\rangle$ is given by

$$(|\alpha\rangle\langle\beta|)|\psi\rangle = |\alpha\rangle\langle\beta|\psi\rangle, \quad (47)$$

where the right-hand side is the product of the complex number $\langle\beta|\psi\rangle$ times the ket $|\alpha\rangle$. The action of $|\alpha\rangle\langle\beta|$ on bras is given by

$$\langle\psi|(|\alpha\rangle\langle\beta|) = \langle\psi|\alpha\rangle\langle\beta|. \quad (48)$$

This is not the definition of $|\alpha\rangle\langle\beta|$, which is given by Eq. (47), rather it is a simple consequence of that definition. Its proof will be left as an exercise. The operator $|\alpha\rangle\langle\beta|$ can be viewed as the tensor product of the ket $|\alpha\rangle$ with the bra $\langle\beta|$; this kind of product is similar to the dyadic product used in ordinary vector or tensor analysis. The operator $|\alpha\rangle\langle\beta|$ is called the *outer product* of $|\alpha\rangle$ and $|\beta\rangle$.

12. Bases and Orthonormal Bases; Resolutions of the Identity

As in any vector space, a *basis* is a set of linearly independent vectors that span the entire space, and the number of such vectors is the dimension of the space. This is a familiar concept in finite dimensions. In an infinite-dimensional space, we obviously need an infinite number of basis vectors,

but is the basis countably or uncountably infinite? That is, can we label the basis vectors with a discrete index, or do we need continuous indices as well? As it turns out, it is one of the mathematical properties of Hilbert spaces that they always possess a countable basis, that is, a set of linearly independent, normalizable vectors that span the space and can be indexed $n = 1, 2, 3, \dots$, or by some other such discrete indexing scheme. This does not mean that every basis we encounter will be labeled by a discrete index, only that such bases always exist. In fact, in practice in quantum mechanics we frequently encounter “bases” with continuous indices as well, but the basis vectors of the continuum are not normalizable and do not belong to Hilbert space. We will say more about these kinds of bases later, and for now proceed with the case of discrete bases, consisting of normalizable vectors. As an example, you may think of the harmonic oscillator wave functions $u_n(x)$ in one-dimensional quantum mechanics.

Let us denote a discrete basis in a ket space by $\{|n\rangle, n = 1, \dots\}$. We say the basis is *orthonormal* if

$$\langle n|m\rangle = \delta_{mn}. \quad (49)$$

We usually use orthonormal bases in quantum mechanics, but non-orthonormal bases are useful sometimes as well (if much more difficult to work with). In any basis, an arbitrary ket $|\psi\rangle$ can be represented as a linear combination of the basis vectors,

$$|\psi\rangle = \sum_n c_n |n\rangle = \sum_n |n\rangle c_n, \quad (50)$$

and if the basis is orthonormal, the expansion coefficients are given by

$$c_n = \langle n|\psi\rangle. \quad (51)$$

Substituting this into Eq. (50), we have

$$|\psi\rangle = \sum_n |n\rangle \langle n|\psi\rangle = \left(\sum_n |n\rangle \langle n| \right) |\psi\rangle, \quad (52)$$

or, since $|\psi\rangle$ is arbitrary,

$$1 = \sum_n |n\rangle \langle n|, \quad (53)$$

where 1 is the identity operator. This is the *resolution of the identity* associated with the orthonormal basis $\{|n\rangle\}$.

Resolutions of the identity associated with bases with continuous indices will be discussed below.

13. Hermitian Conjugation

We have already defined Hermitian conjugation (denoted by the dagger \dagger) of bras and kets. We now define Hermitian conjugation of operators. If $A : \mathcal{E} \rightarrow \mathcal{E}$ is a linear operator, then $A^\dagger : \mathcal{E} \rightarrow \mathcal{E}$ is another linear operator, defined by its action on an arbitrary ket $|\psi\rangle$ by

$$A^\dagger |\psi\rangle = (\langle \psi|A)^\dagger. \quad (54)$$

We should note that $\langle\psi|$ is an antilinear function of $|\psi\rangle$ (because of the dual correspondence), but when we do the (second) Hermitian conjugation indicated on the right hand side of Eq. (54), the result is a linear function of $|\psi\rangle$. Thus, A^\dagger is a linear operator.

There are a number of immediate consequences of this definition, which we list here. The proofs are left as simple exercises.

$$\langle\phi|A^\dagger|\psi\rangle = \langle\psi|A|\phi\rangle^*, \quad (55)$$

$$(A^\dagger)^\dagger = A, \quad (56)$$

$$(c_1A_1 + c_2A_2)^\dagger = c_1^*A_1^\dagger + c_2^*A_2^\dagger, \quad (57)$$

$$(AB)^\dagger = B^\dagger A^\dagger, \quad (58)$$

$$(|\alpha\rangle\langle\beta|)^\dagger = |\beta\rangle\langle\alpha|. \quad (59)$$

We have now defined the action of the Hermitian conjugate operation † on kets, bras, and operators. It is convenient also to define the action of the Hermitian conjugate on complex numbers as being identical with ordinary complex conjugation. With this understanding, we can state a general rule, which is illustrated in several various ways in the results above, which says that the Hermitian conjugate of a product of objects (kets, bras, operators, complex numbers) is given by reversing the order of the objects and replacing each by its Hermitian conjugate. The products involved can be either ordinary multiplication by scalars, operator products, or inner or outer (tensor) products. We only require that the original product be meaningful.

14. Hermitian, Anti-Hermitian and Unitary Operators

We say that an operator A is *Hermitian* if it is equal to its own Hermitian conjugate,

$$A^\dagger = A. \quad (60)$$

An equivalent definition is

$$\langle\psi|A|\phi\rangle = \langle\phi|A|\psi\rangle^*, \quad (61)$$

for all kets $|\psi\rangle$ and $|\phi\rangle$. Similarly, an operator is said to be *anti-Hermitian* if it satisfies

$$A^\dagger = -A. \quad (62)$$

An arbitrary operator is in general neither Hermitian nor anti-Hermitian, but can always be uniquely decomposed into a sum of a Hermitian and an anti-Hermitian operator:

$$A = \frac{A + A^\dagger}{2} + \frac{A - A^\dagger}{2}. \quad (63)$$

If A is Hermitian, then the matrix element $\langle\psi|A|\psi\rangle$ is necessarily real, for all choices of $|\psi\rangle$, in accordance with Eq. (55). We say that a Hermitian operator is *positive definite* if it satisfies

$$\langle\psi|A|\psi\rangle > 0, \quad (64)$$

for all nonzero kets $|\psi\rangle$. Similarly, we say that A is *nonnegative definite* if

$$\langle\psi|A|\psi\rangle \geq 0, \quad (65)$$

for all kets $|\psi\rangle$.

We have not discussed eigenvalues yet, but a positive definite, Hermitian operator is one whose eigenvalues are all positive, while a nonnegative definite, Hermitian operator is one whose eigenvalues are all nonnegative.

A simple but important theorem is the following:

Theorem 1. *The product of two Hermitian operators is Hermitian if and only if they commute.*

The (trivial) proof is left as an exercise.

An operator U is *unitary* if

$$UU^\dagger = U^\dagger U = 1, \quad (66)$$

so that

$$U^{-1} = U^\dagger. \quad (67)$$

The product of two unitary operators is always unitary.

15. Eigenvalues and Eigenkets; the Spectrum of an Operator

If A is an operator acting on \mathcal{E} and there exists a nonzero ket $|u\rangle$ and complex number a such that

$$A|u\rangle = a|u\rangle, \quad (68)$$

then we say that $|u\rangle$ is an *eigenket* of A and a is a (right) *eigenvalue*. Similarly, if there exists a nonzero bra $\langle v|$ and a complex number b such that

$$\langle v|A = b\langle v|, \quad (69)$$

then we say that $\langle v|$ is an *eigenbra* of A and b is a (left) eigenvalue. In finite dimensions, every right eigenvalue is also a left eigenvalue, but in infinite dimensions this need not be true.

We define the *spectrum* of an operator as the set of its eigenvalues, seen as a subset of the complex plane. (If necessary, we can distinguish between a left and right spectrum.) In finite dimensions, this is a set of discrete points in the complex plane.

Some examples of the spectra of familiar operators in quantum mechanics are illustrated in Fig. 1. In part (a) of the figure we have the spectrum of the harmonic oscillator, whose eigenvalues $E_n = (n + \frac{1}{2})\hbar\omega$ are shown as spots on the real energy axis. This is an example of a *discrete spectrum*. Only real energies are physically meaningful, but we show the whole complex energy plane because in general eigenvalues of operators are complex. (We will see complex energies play a role in scattering and resonance theory later in the course.) In part (b) we have the spectrum of

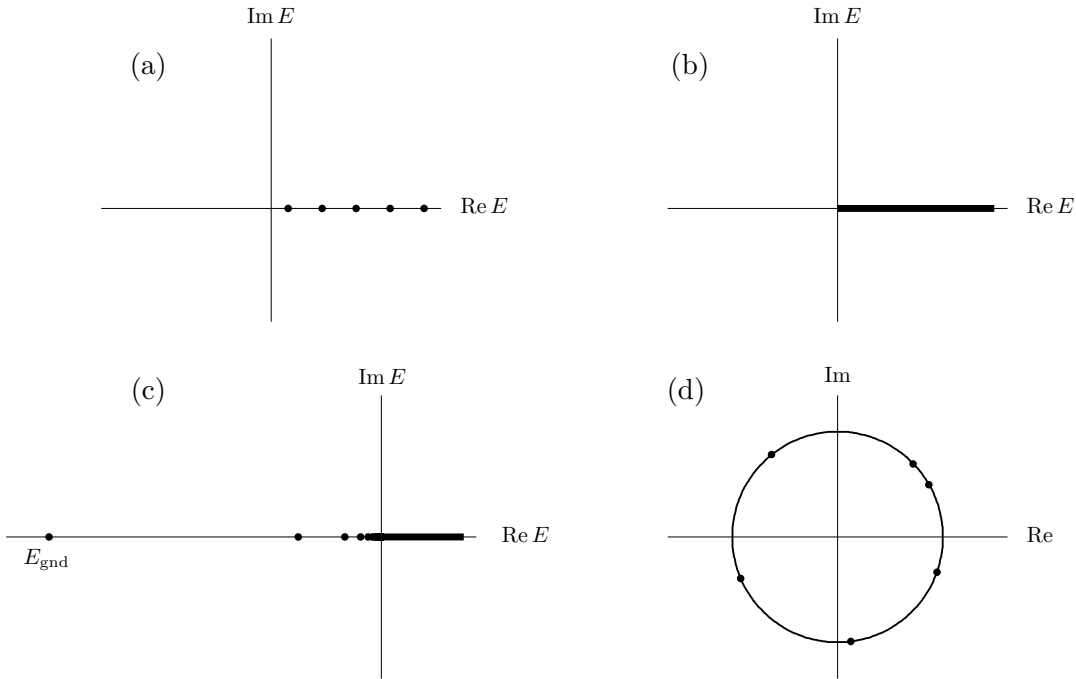


Fig. 1. Examples of spectra of operators. Part (a), the harmonic oscillator; part (b), the free particle; part (c), the hydrogen atom; part (d), a finite-dimensional unitary operator.

the free particle Hamiltonian $H = p^2/2m$. It consists of all real energies $E \geq 0$, indicated by the thick line on the positive real axis. This is an example of the *continuous spectrum*. In part (c) we have the spectrum of the hydrogen atom. It consists of a discrete set of spots at negative energy, representing the bound states, an infinite number of which accumulate as we approach $E = 0$. As with the free particle, there is a continuous spectrum above $E = 0$. Altogether, the hydrogen atom has a *mixed spectrum* (discrete and continuous). Finally, in part (d) we have the spectrum of a unitary operator on a finite-dimensional space. Its eigenvalues are phase factors that lie on the unit circle in the complex plane.

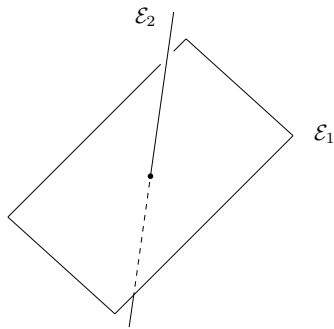


Fig. 2. An operator A acting on ket space \mathcal{E} has eigenspaces \mathcal{E}_n , associated with its discrete eigenvalues a_n .

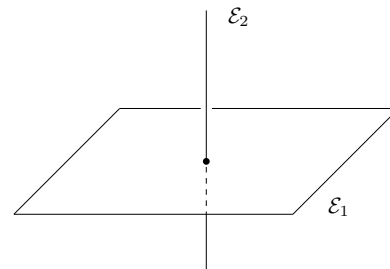


Fig. 3. In the case of a Hermitian operator, the eigenspaces are orthogonal.

In the case of the discrete spectrum, the set of kets $|u\rangle$ that satisfy Eq. (68) for a given eigenvalue a forms a vector subspace of the ket space. This subspace is at least 1-dimensional, but the dimensionality may be higher (even infinite, on infinite-dimensional vector spaces). We will call this subspace the *eigenspace* of the operator corresponding to the given eigenvalue. If the eigenspace is 1-dimensional, then we say the eigenvalue is *nondegenerate*, otherwise, that it is *degenerate*. We refer to the dimensionality of the eigenspace as the *order* of the degeneracy. This is the same as the number of linearly independent eigenvectors of a given eigenvalue. We can imagine the eigenspaces of the discrete spectrum as illustrated in Fig. 2, which shows the eigenspaces for a two-fold degenerate eigenvalue a_1 and a nondegenerate eigenvalue a_2 ($\dim \mathcal{E}_1 = 2$, $\dim \mathcal{E}_2 = 1$).

16. The Case of Hermitian Operators

In general, there is no simple relation between the eigenkets and eigenbras of an operator, even in finite dimensions. But if an operator A is Hermitian, then there are a number of simplifications. We speak for the time being of the discrete spectrum.

First, the eigenvalue a is real, so the spectrum in the complex eigenvalue plane is a subset of the real axis. To prove this we take $A|u\rangle = a|u\rangle$ and multiply on the left by $\langle u|$, obtaining

$$\langle u|A|u\rangle = a\langle u|u\rangle. \quad (70)$$

Now taking the complex conjugate of this, the left-hand side goes into itself, while the right becomes $a^*\langle u|u\rangle$. Thus we have

$$(a - a^*)\langle u|u\rangle = 0. \quad (71)$$

Since $|u\rangle \neq 0$, this implies $a = a^*$.

Second, if $A|u\rangle = a|u\rangle$, then $\langle u|A = a\langle u|$, so that the eigenbras are the Hermitian conjugates of the eigenkets, and the left and right eigenvalues are equal.

Third, any two eigenkets corresponding to distinct eigenvalues are orthogonal. We state this final property as a theorem:

Theorem 2. *The eigenspaces of a Hermitian operator corresponding to distinct eigenvalues are orthogonal. This means that any pair of vectors chosen from the two eigenspaces are orthogonal.*

To prove this, we let $A|u\rangle = a|u\rangle$ and $A|u'\rangle = a'|u'\rangle$ for two eigenkets $|u\rangle$ and $|u'\rangle$ and two eigenvalues a and a' . Multiplying the first equation by $\langle u'|$ and the second by $\langle u|$, we obtain

$$\begin{aligned} \langle u'|A|u\rangle &= a\langle u'|u\rangle, \\ \langle u|A|u'\rangle &= a'\langle u|u'\rangle. \end{aligned} \quad (72)$$

Now taking the Hermitian conjugate of the second equation and subtracting from the first, we find

$$(a - a')\langle u|u'\rangle = 0, \quad (73)$$

showing that either $a = a'$ or else $\langle u|u' \rangle = 0$.

In view of this theorem, Fig. 3 gives a more accurate picture of the eigenspaces of a Hermitian operator than does Fig. 2.

17. Completeness

In looking at Figs. 2 and 3 we may imagine that the space in which the eigenspaces lie is three dimensional, so the two eigenspaces (of dimension 2 and 1) “fill up” the whole space. Do the eigenspaces of an operator always “fill up” the space upon which the operator acts? An operator that does this is said to be *complete*. The answer is no, not even in finite dimensions. For example, the matrix

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad (74)$$

possesses only one eigenvector, namely,

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad (75)$$

so it has only a single one-dimensional eigenspace. The matrix (74) is not Hermitian.

On the other hand, in finite dimensions every Hermitian operator is complete. Thus, the sum of the dimensionalities of the eigenspaces is the dimension of the whole space. Since the eigenspaces are orthogonal, by choosing an orthonormal basis in each eigenspace we obtain an orthonormal eigenbasis for the whole space, that is, a set of orthonormal basis vectors that are also eigenvectors of the operator. This is a consequence of the standard methods taught in linear algebra courses for diagonalizing Hermitian matrices, which explicitly produce the orthonormal eigenbasis.

Notice that the orthonormal eigenbasis is not unique. First, if there is any degeneracy (any eigenspace with dimensionality ≥ 2), then obviously there are an infinite number of ways of choosing an orthonormal basis in that eigenspace. But even in a nondegenerate (one-dimensional) eigenspace, a normalized eigenvector is only determined to within a phase factor.

These nice properties of Hermitian operators are shared by a larger class of *normal* operators, which will be studied in a homework exercise. Occasionally in quantum mechanics we must deal with the eigenvectors of non-Hermitian operators and worry about their completeness properties. This happens, for example, in the theory of the Dirac equation.

18. The Direct Sum; Orthonormal Eigenbases

Suppose we have some vector space \mathcal{E} , and suppose \mathcal{E}_1 and \mathcal{E}_2 are vector subspaces of \mathcal{E} that have no element in common except 0, such as the spaces we see in Figs. 2 and 3. Then we say that the *direct sum* of \mathcal{E}_1 and \mathcal{E}_2 , denoted $\mathcal{E}_1 \oplus \mathcal{E}_2$, is the set of vectors that can be formed by taking linear combinations of vectors in the two subspaces. That is,

$$\mathcal{E}_1 \oplus \mathcal{E}_2 = \{|\psi_1\rangle + |\psi_2\rangle \text{ such that } |\psi_1\rangle \in \mathcal{E}_1, |\psi_2\rangle \in \mathcal{E}_2\}. \quad (76)$$

The direct sum is itself a vector subspace of \mathcal{E} . One can also think of the direct sum as the span of the set of vectors formed when we throw together a basis in \mathcal{E}_1 with a basis in \mathcal{E}_2 . Thus, we have

$$\dim(\mathcal{E}_1 \oplus \mathcal{E}_2) = \dim \mathcal{E}_1 + \dim \mathcal{E}_2. \quad (77)$$

This is not the most abstract definition of the direct sum, but it the one we will use.

We can now express the completeness property in finite dimensions in terms of the direct sum. Let A be an operator acting on a finite-dimensional space \mathcal{E} . Let the eigenvalues of A be a_1, \dots, a_N , and let \mathcal{E}_n be the n -th eigenspace, that is,

$$\mathcal{E}_n = \{|u\rangle \text{ such that } A|u\rangle = a_n|u\rangle\}. \quad (78)$$

Then A is complete if

$$\mathcal{E}_1 \oplus \mathcal{E}_2 \oplus \dots \oplus \mathcal{E}_N = \mathcal{E}. \quad (79)$$

In finite dimensions, any Hermitian A is complete, and moreover it decomposes the original Hilbert space \mathcal{E} into a collection of orthogonal subspaces, the eigenspaces of the operator.

Let us choose an orthonormal basis, say $|nr\rangle$, in each of the subspaces \mathcal{E}_n , where r is an index labeling the orthonormal vectors in the subspace, so that $r = 1, \dots, \dim \mathcal{E}_n$. This can always be done, in many ways. Then, since the various subspaces are themselves orthogonal, the collection of vectors $|nr\rangle$ for all n and r forms an orthonormal basis on the entire space \mathcal{E} , so that

$$\langle nr|n'r'\rangle = \delta_{nn'} \delta_{rr'}. \quad (80)$$

These kets are eigenkets of A ,

$$A|nr\rangle = a_n|nr\rangle, \quad (81)$$

that is, they constitute an orthonormal eigenbasis. Furthermore, we have the resolution of the identity,

$$1 = \sum_{nr} |nr\rangle\langle nr|, \quad (82)$$

corresponding to the chosen orthonormal eigenbasis of the complete Hermitian operator A . Also, the operator itself can be written in terms of this eigenbasis,

$$A = \sum_{nr} a_n |nr\rangle\langle nr|. \quad (83)$$

This follows if we let both sides of this equation act on a basis vector $|n'r'\rangle$, whereupon the answers are the same. By linear superposition, the same applies to any vector. Thus, both sides are the same operator.

19. Spectral Properties in Infinite Dimensions

Here we make some remarks about infinite-dimensional ket spaces and illustrate the kinds of new phenomena that can arise.

In infinite dimensions, there exist Hermitian operators that are not complete. To cover this case, we define a complete Hermitian operator as an *observable*. This terminology is related to the physical interpretation of observables in quantum mechanics, as explained in Notes 2. In practice we seldom deal with incomplete Hermitian operators in quantum mechanics.

We now use examples from one-dimensional wave mechanics to illustrate some phenomena that can occur in infinite dimensions. The Hilbert space of wave functions in one dimension consists of normalizable functions $\psi(x)$ with the scalar product,

$$\langle \psi | \phi \rangle = \int_{-\infty}^{+\infty} dx \psi^*(x) \phi(x). \quad (84)$$

We will present four examples of operators and their spectral properties.

The first is the one-dimensional harmonic oscillator with Hamiltonian,

$$H = \frac{p^2}{2m} + \frac{m\omega^2 x^2}{2}, \quad (85)$$

which is a Hermitian operator. The energy eigenvalues are

$$E_n = (n + \frac{1}{2})\hbar\omega, \quad (86)$$

so that the spectrum consists of a discrete set of isolated points on the positive energy axis, as illustrated in Fig. 1(a). The eigenfunctions $u_n(x)$ are normalizable (they belong to Hilbert space) and orthogonal for different values of the energy. Thus they can be normalized so that

$$\langle n | m \rangle = \int dx u_n^*(x) u_m(x) = \delta_{mn}. \quad (87)$$

These eigenfunctions are nondegenerate and complete, so we have a resolution of the identity,

$$1 = \sum_{n=0}^{\infty} |n\rangle \langle n|. \quad (88)$$

Not all Hermitian operators on an infinite dimensional space have a discrete spectrum. As a second example, consider the momentum operator in one dimension, which is $\hat{p} = -i\hbar d/dx$. (We put a hat on the operator, to distinguish it from the eigenvalue p .) The momentum is a Hermitian operator (see below). The eigenvalue-eigenfunction equation for this operator is

$$\hat{p}u_p(x) = pu_p(x) = -i\hbar \frac{du_p}{dx}, \quad (89)$$

where $u_p(x)$ is the eigenfunction with eigenvalue p . This equation has a solution for all values, real and complex, of the parameter p ,

$$u_p(x) = c e^{ipx/\hbar}, \quad (90)$$

where c is a constant. If p has a nonzero imaginary part, then $u_p(x)$ diverges exponentially as $x \rightarrow \infty$ or $x \rightarrow -\infty$, so the “eigenfunction” is rather badly behaved. It is certainly not normalizable in this case. But even if p is purely real, the function $u_p(x)$ is still not normalizable, so it does not represent a physically allowable state (a vector of Hilbert space). The momentum operator \hat{p} does not have any eigenvectors that belong to Hilbert space, much less a basis.

One way to handle this case, which provides a kind of generalization of the nice situation we find in finite dimensions, is to regard every real number p as an eigenvalue of \hat{p} , which means that the spectrum of p is the entire real axis (in the complex p -plane). See Fig. 1(b) for an illustration of the spectrum of the free particle Hamiltonian, which is the portion $E \geq 0$ of the real axis. Recall that in finite dimensions a Hermitian operator has only real eigenvalues. This is an example of the *continuous spectrum*. Also, the eigenfunctions that result in this case are orthonormal and complete in a certain sense, but as noted they do not belong to Hilbert space. We obtain an orthonormal basis, but it consists of vectors that lie outside Hilbert space.

We adopt a normalization of the eigenfunctions (that is, we choose the constant c in Eq. (90)) so that

$$u_p(x) = \frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}}, \quad (91)$$

because this implies a convenient normalization relation,

$$\langle p_0 | p_1 \rangle = \int dx u_{p_0}^*(x) u_{p_1}(x) = \int \frac{dx}{2\pi\hbar} e^{i(p_1 - p_0)x/\hbar} = \delta(p_0 - p_1), \quad (92)$$

where we write the function $u_p(x)$ in ket language as $|p\rangle$ (giving only the eigenvalue). The eigenfunctions $u_p(x)$ of the continuous spectrum are orthonormal only in the δ -function sense. They are also complete, in the sense that an arbitrary wave function $\psi(x)$ can be expanded as a linear combination of them, but since they are indexed by a continuous variable, the linear combination is an integral, not a sum. That is, given $\psi(x)$, we can find expansion coefficients $\phi(p)$ such that

$$\psi(x) = \int_{-\infty}^{\infty} dp \phi(p) u_p(x) = \int_{-\infty}^{\infty} \frac{dp}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} \phi(p). \quad (93)$$

We know this because Eq. (93) is essentially a Fourier transform, whose inverse is given by

$$\phi(p) = \int \frac{dx}{\sqrt{2\pi\hbar}} e^{-ipx/\hbar} \psi(x) = \int dx u_p^*(x) \psi(x) = \langle p | \psi \rangle. \quad (94)$$

We see that the “momentum space wave function” $\phi(p)$ is otherwise the scalar product $\langle p | \psi \rangle$.

Another example of an operator with a continuous spectrum is the position operator \hat{x} (again using hats to denote an operator). This operator acting on wave functions means, multiply by x . If we denote the eigenfunction of this operator with eigenvalue x_0 by $f_{x_0}(x)$, then the eigenvalue equation is

$$\hat{x} f_{x_0}(x) = x f_{x_0}(x) = x_0 f_{x_0}(x), \quad (95)$$

or,

$$(x - x_0) f_{x_0}(x) = 0. \quad (96)$$

This implies either $x = x_0$ or $f_{x_0}(x) = 0$, so $f_{x_0}(x)$ is a “function” that is zero everywhere except possibly at x_0 . We interpret this situation in terms of the Dirac delta function, by writing the solution as

$$f_{x_0}(x) = \delta(x - x_0). \quad (97)$$

Again, the spectrum is continuous and occupies the entire real axis; again, the “eigenfunctions” are of infinite norm and obey the continuum orthonormality condition,

$$\langle x_0 | x_1 \rangle = \int dx \delta(x - x_0) \delta(x - x_1) = \delta(x_0 - x_1), \quad (98)$$

where we write the function $f_{x_0}(x)$ in ket language as $|x_0\rangle$ (giving only the eigenvalue). In addition, we have the identity,

$$\psi(x_0) = \int dx \delta(x - x_0) \psi(x) = \int dx f_{x_0}^*(x) \psi(x) = \langle x_0 | \psi \rangle. \quad (99)$$

We see that the wave function $\psi(x_0)$ is otherwise the scalar product $\langle x_0 | \psi \rangle$. Substituting this back into Eq. (99) and using Eq. (98), we have

$$\langle x_0 | \psi \rangle = \int dx \langle x_0 | x \rangle \langle x | \psi \rangle, \quad (100)$$

or, since x_0 is arbitrary,

$$|\psi\rangle = \int dx |x\rangle \langle x | \psi \rangle. \quad (101)$$

Finally, since $|\psi\rangle$ is arbitrary, we can strip it off, leaving

$$1 = \int dx |x\rangle \langle x|, \quad (102)$$

the resolution of the identity in the case of the continuous eigenbasis of the position operator.

We can also get the resolution of the identity in the momentum basis. Since $\psi(x) = \langle x | \psi \rangle$ for any ψ , we have $u_p(x) = \langle x | p \rangle$. Substituting these and Eq. (94) into Eq. (93), we have

$$\langle x | \psi \rangle = \int dp \langle x | p \rangle \langle p | \psi \rangle, \quad (103)$$

or, stripping off both $\langle x |$ on the left and $|\psi\rangle$ on the right,

$$1 = \int dp |p\rangle \langle p|. \quad (104)$$

Equations (102) and (104) illustrate resolutions of the identity in the case of the continuous spectrum.

For a final example, let us consider the 1-dimensional potential illustrated in Fig. 4, which has a well followed on the right by an infinitely high barrier. The purpose of the barrier is simply to make the unbound energy eigenfunctions nondegenerate, which simplifies the following discussion. In a potential such as this, we expect there to be some number of bound eigenfunctions $u_n(x)$ in the potential well, which have negative energies $E_n < 0$. These eigenfunctions are normalizable, because

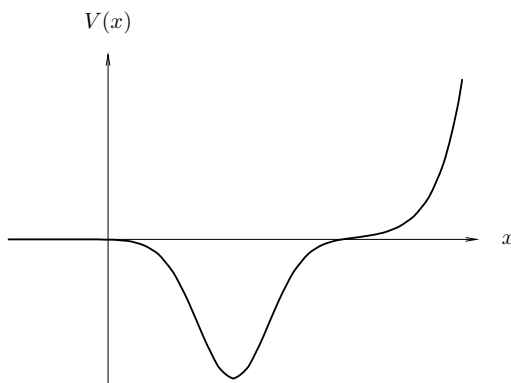


Fig. 4. A potential in 1-dimension with both bound and unbound eigenstates.

the wave functions die off exponentially outside the well, and we can normalize them according to $\langle n|m\rangle = \delta_{nm}$, as with the discrete spectra discussed above. But there will also be unbound eigenfunctions $u_E(x)$ corresponding to any energy $E \geq 0$; these eigenfunctions become free particle wave functions at large negative x , and therefore have infinite norm. We normalize them according to $\langle E|E'\rangle = \delta(E - E')$. Finally, all bound states are orthogonal to all continuum states, $\langle n|E\rangle = 0$. Thus, the completeness relation in this case has the form,

$$1 = \sum_n |n\rangle\langle n| + \int_0^\infty dE |E\rangle\langle E|. \quad (105)$$

In this case we have an example of a *mixed spectrum*, with some discrete and some continuous eigenvalues. The spectrum consists of a discrete set of points E_n on the negative energy axis, plus the entire positive energy axis $E \geq 0$.

More complicated kinds of spectra also occur. It is possible to have discrete states imbedded in the continuum, not in real physical problems but in approximations to them, in which certain interactions are switched off. When all the physics is taken into account, such discrete states typically move off the real axis, and become resonances with a finite lifetime. Examples of this occurs in the autoionizing states of helium, or in the radiative decay of excited atomic states. It is also possible to have a discrete spectrum in which the eigenvalues are not isolated, as in the random potentials that occur in the theory of Anderson localization, an important topic in solid state physics.

20. Generalizations about Spectra, Eigenspaces and Eigenbases

On a finite-dimensional Hilbert space, all Hermitian operators have a discrete spectrum (and they are moreover complete). The continuous and more complicated types of spectra arise only in the infinite dimensional case.

In the infinite dimensional case, each eigenvalue of the discrete spectrum corresponds to a definite eigenspace, a subspace of the Hilbert space consisting of normalizable eigenvectors of the Hermitian operator in question. If such an eigenspace is finite dimensional, then it possesses a

finite, discrete basis of normalizable eigenvectors, as does any finite-dimensional Hilbert space. If an eigenspace of the discrete spectrum is infinite dimensional, then it is an infinite dimensional subspace of a Hilbert space and is therefore an infinite dimensional Hilbert space in its own right. This means that it is possible to choose a discrete basis of normalizable vectors that span this space, but that it may be convenient in practice to choose other types of bases, such as a continuous basis, consisting of unnormalizable vectors outside Hilbert space, or a mixed basis, or other possibilities.

An example of an infinite dimensional eigenspace of an operator with a discrete spectrum is given by the parity operator acting on wave functions $\psi(x)$ in one dimension. The two eigenspaces consist of wave functions that are even and odd under $x \rightarrow -x$, corresponding to the eigenvalues $+1$ and -1 of the parity operator, and both are infinite dimensional.

In the continuous spectrum, there is no subspace of the Hilbert space corresponding to a given eigenvalue. For example, in the case of the free particle in one dimension with Hamiltonian $H = p^2/2m$, for a given $E > 0$ there are two linearly independent eigenfunctions, e^{ikx} and e^{-ikx} , where $E = \hbar^2 k^2/2m$. But these eigenfunctions are not normalizable, and do not belong to Hilbert space. Since there is no linear combination of these eigenfunctions that is normalizable, they do not specify a subspace of Hilbert space. On the other hand, any interval in the continuous spectrum does correspond to a subspace of the Hilbert space. This subspace consists of normalizable vectors that can be represented as linear combinations of the unnormalizable eigenvectors whose eigenvalues lie in the given interval.

For example, let $I = [p_0, p_1]$ be an interval of the momentum axis. Then normalizable wave functions $\psi(x)$ whose Fourier transform $\phi(p)$ vanishes outside the interval I form a subspace of the Hilbert space of normalizable wave functions $\psi(x)$.

Concepts like this are familiar in electronics. A circuit that passes a signal in a certain frequency range $[\omega_0, \omega_1]$ while rejecting frequency components outside that range is analogous to a projection operator in quantum mechanics that projects onto a certain interval of the momentum axis $[p_0, p_1]$. The main difference is that one works in the time domain and the other in the spatial domain (when acting on wave functions $\psi(x)$).

The concept of a subspace corresponding to an interval of the continuous spectrum plays a role in the measurement postulates of quantum mechanics, as we will see in Notes 2.

21. Proving of Hermiticity of the Momentum

To show that the momentum operator \hat{p} is Hermitian on the Hilbert space of wave functions $\psi(x)$ with scalar product (84), we appeal to the definition of Hermiticity in the form (61). Letting $\psi(x)$ and $\phi(x)$ be two wave functions, we have

$$\langle \psi | \hat{p} | \phi \rangle = \int dx \psi^* \left(-i\hbar \frac{\partial \phi}{\partial x} \right) = \int dx \left(-i\hbar \frac{\partial}{\partial x} \right) (\psi^* \phi) + \int dx \left(i\hbar \frac{\partial \psi^*}{\partial x} \right) \phi, \quad (106)$$

by integration by parts. The first integral vanishes if ψ and ϕ go to zero as $x \rightarrow \pm\infty$, something we

normally expect for normalizable wave functions (but see Sec. 28). The final integral is

$$\left[\int dx \phi^* \left(-i\hbar \frac{\partial \psi}{\partial x} \right) \right]^* = \langle \phi | \hat{p} | \psi \rangle^*, \quad (107)$$

which completes the proof. A similar proof can be given for Hamiltonians of the form $H = p^2/2m + V(x)$. More simply, once we know that \hat{x} and \hat{p} are Hermitian, we can use the general rules presented in these notes (for example, Eq. (58) and Sec. 25)) to conclude that the sum of any real function of \hat{x} and any real function of \hat{p} is Hermitian.

22. Orthonormality and Completeness in the General (Degenerate) Case

We return now to the general case of a complete Hermitian operator A (an observable) acting on an infinite-dimensional Hilbert space. In general, the spectrum consists of a discrete part, with eigenvalues a_n , and a continuous part, with eigenvalues $a(\nu)$ that are functions of some continuous variable ν . (We could choose $\nu = a$, but sometimes it is convenient to use another continuous parameter upon which the eigenvalue depends. For example, we may wish to use the momentum p instead of the energy E to parameterize continuum energy eigenvalues and eigenfunctions.) In addition, we must in general introduce another index r to resolve degeneracies in degenerate subspaces; for simplicity we will assume that r is a discrete index, although in some problems this would be continuous, too. Then the orthonormality conditions have the form,

$$\begin{aligned} \langle nr | n' r' \rangle &= \delta_{nn'} \delta_{rr'}, \\ \langle nr | \nu r' \rangle &= 0, \\ \langle \nu r | \nu' r' \rangle &= \delta(\nu - \nu') \delta_{rr'}, \end{aligned} \quad (108)$$

and the completeness relation has the form,

$$1 = \sum_{nr} |nr\rangle \langle nr| + \int d\nu \sum_r |\nu r\rangle \langle \nu r|. \quad (109)$$

As a specific example of these relations, consider the hydrogen atom in the electrostatic model, in which we neglect the electron spin. The bound state wave functions are $\psi_{n\ell m}(\mathbf{r})$ in the usual notation, which we write in ket language as $|n\ell m\rangle$. We normalize these so that $\langle n\ell m | n'\ell' m' \rangle = \delta_{nn'} \delta_{\ell\ell'} \delta_{mm'}$. In addition, there are unbound (and unnormalizable) eigenfunctions of energy $E \geq 0$, which we write in ket language as $|E\ell m\rangle$. We normalize these so that $\langle E\ell m | E'\ell' m' \rangle = \delta(E - E') \delta_{\ell\ell'} \delta_{mm'}$. Also, the bound states are orthogonal to the unbound states, $\langle n\ell m | E'\ell' m' \rangle = 0$. Then the completeness relation has the form,

$$1 = \sum_{n\ell m} |n\ell m\rangle \langle n\ell m| + \int_0^\infty dE \sum_{\ell m} |E\ell m\rangle \langle E\ell m|. \quad (110)$$

The generalized orthonormality and completeness relations we have just presented are examples of an important theorem, which we now quote:

Theorem 3. *An observable always possesses an orthonormal eigenbasis.*

This eigenbasis must include states of infinite norm, if the observable possesses a continuous spectrum.

23. Simultaneous Eigenbases of Commuting Observables

Next we turn to an important theorem, applications of which will occur many times in this course:

Theorem 4. *Two observables possess a simultaneous eigenbasis if and only if they commute. This eigenbasis may be chosen to be orthonormal. By extension, a collection of observables possesses a simultaneous eigenbasis if and only if they commute.*

For simplicity we will work on a finite-dimensional Hilbert space \mathcal{E} . We begin with the case of two observables. In the first part of the proof, we assume we have two Hermitian operators A , B that commute. We wish to prove that they possess a simultaneous eigenbasis.

Consider the n -th eigenspace \mathcal{E}_n of the operator A , and let $|u\rangle$ be a eigenket in this eigenspace, so that

$$A|u\rangle = a_n|u\rangle. \quad (111)$$

Then by multiplying by B and using the fact that $AB = BA$, we easily conclude that

$$A(B|u\rangle) = a_n(B|u\rangle), \quad (112)$$

in which we use parentheses to emphasize the fact that the ket $B|u\rangle$ is also an eigenket of A with the same eigenvalue, a_n . Note the possibility that $B|u\rangle$ could vanish, which does not, however, change any of the following argument. We see that if any ket $|u\rangle$ belongs to the subspace \mathcal{E}_n , then so does the ket $B|u\rangle$.

To complete the proof we need the concept of the *restriction* of an operator to a subspace. In general, it is only meaningful to talk about the restriction of an operator to some subspace of a larger space upon which the operator acts if it should happen that the operator maps any vector of the given subspace into another vector of the same subspace. In this case, we say that the subspace is *invariant* under the action of the operator.

In the present example, both operators A and B leave the subspace \mathcal{E}_n invariant. As for A , this follows from Eq. (111), which is in effect the definition of \mathcal{E}_n . As for B , this follows from Eq. (112). Let us denote the restriction of A and B to the subspace \mathcal{E}_n by \bar{A} and \bar{B} , respectively. Then, according to Eq. (111), \bar{A} is a multiple of the identity, $\bar{A} = a_n\bar{I}$, where \bar{I} stands for the identity operator acting on the subspace. The operator \bar{B} cannot be expressed so simply, but it is a Hermitian operator, since B was Hermitian on the larger space \mathcal{E} . Since we are assuming that \mathcal{E} is finite-dimensional, so is \mathcal{E}_n , and therefore \bar{B} is complete on \mathcal{E}_n . Therefore \bar{B} possesses an orthonormal eigenbasis in \mathcal{E}_n , say,

$$\bar{B}|nr\rangle = b_{nr}|nr\rangle, \quad (113)$$

where $r = 1, \dots, \dim \mathcal{E}_n$. Here n is considered fixed, and if there are any degeneracies of \bar{B} , they are indicated by repetitions of the eigenvalues b_{nr} for different values of r .

If we now construct the bases $|nr\rangle$ in a similar manner for each value of n , then we obtain a simultaneous, orthonormal eigenbasis for both A and B on the whole space \mathcal{E} . Notice that there may be degeneracies of B that cross the eigenspaces of A , that is, there may be repetitions in the values b_{nr} for different values of n .

To prove the converse of Theorem 4, we assume that A and B possess a simultaneous eigenbasis. We sequence the kets of this eigenbasis according to some index s , and write

$$A|s\rangle = a_s|s\rangle, \quad B|s\rangle = b_s|s\rangle. \quad (114)$$

In these formulas, we do not assume that the eigenvalues a_s or b_s are distinct for different values of s ; degeneracies are indicated by repetitions of the values a_s or b_s for different values of s . Now from Eq. (114) it follows immediately that

$$AB|s\rangle = a_s b_s |s\rangle = BA|s\rangle, \quad (115)$$

or,

$$[A, B]|s\rangle = 0. \quad (116)$$

But since the kets $|s\rangle$ form a basis, Eq. (116) implies $[A, B] = 0$. QED. (The extension of Theorem 4 will be left as an exercise.)

An important corollary of Theorem 4 occurs when some eigenvalue a_n of A is nondegenerate, so that \mathcal{E}_n is 1-dimensional. In that case, all kets in \mathcal{E}_n are proportional to any nonzero ket in \mathcal{E}_n . Thus, if we let $|n\rangle$ be some standard, normalized eigenket of A with eigenvalue a_n , then Eq. (112) tells us that $B|n\rangle$ must be proportional to $|n\rangle$, say,

$$B|n\rangle = b|n\rangle, \quad (117)$$

for some number b . We see that b is an eigenvalue of B , and $|n\rangle$ is an eigenket, although the theorem does not tell us the value of b . The number b must be real, since B is Hermitian. We summarize these facts by another theorem.

Theorem 5. *If two observables A and B commute, $[A, B] = 0$, then any nondegenerate eigenket of A is also an eigenket of B . (It may be a degenerate eigenket of B .) By extension, if A_1, \dots, A_n are observables that commute with each other and with another observable B , then any nondegenerate, simultaneous eigenket of A_1, \dots, A_n is also an eigenket of B .*

The proof of the extension will be left as an exercise.

As an example of this theorem, consider a Hamiltonian that commutes with parity, $[H, \pi] = 0$. According to the theorem, every nondegenerate eigenstate of the Hamiltonian is necessarily a state of definite parity (even or odd), that is, it is an eigenstate of parity. Recall that in the one-dimensional harmonic oscillator, the energy eigenstate for $n = \text{even}$ are even under parity, while those for $n = \text{odd}$ are odd.

24. Projection Operators

We turn now to the subject of projection operators. We say that an operator P is a *projection operator* or a *projector* if it is an observable that satisfies

$$P^2 = P. \quad (118)$$

(On account of this equation, we say that P is *idempotent*, from Latin *idem* = ‘same,’ and *potens* = ‘powerful.’) It follows immediately from this definition that the eigenvalues of P are either 0 or 1, for if we write

$$P|p\rangle = p|p\rangle, \quad (119)$$

then by Eq. (118) we have

$$P^2|p\rangle = P|p\rangle = p^2|p\rangle = p|p\rangle, \quad (120)$$

or,

$$(p^2 - p)|p\rangle = 0. \quad (121)$$

Thus, either $p^2 = p$ or $|p\rangle = 0$; we exclude the latter possibility, since eigenkets are supposed to be nonzero, and therefore conclude that either $p = 0$ or $p = 1$. Therefore P breaks the Hilbert space \mathcal{E} into two orthogonal subspaces,

$$\mathcal{E} = \mathcal{E}_0 \oplus \mathcal{E}_1, \quad (122)$$

such that P annihilates any vector in \mathcal{E}_0 , and leaves any vector in \mathcal{E}_1 invariant. We say that \mathcal{E}_1 is the space upon which P *projects*.

Projection operators arise naturally in the eigenvalue problem for an observable A . For simplicity, assume that A has a discrete spectrum, so that we can write

$$A|nr\rangle = a_n|nr\rangle, \quad (123)$$

where $|nr\rangle$ is an orthonormal eigenbasis of A , with an arbitrarily chosen orthonormal basis $r = 1, \dots, \dim \mathcal{E}_n$ in each eigenspace \mathcal{E}_n . We then define

$$P_n = \sum_r |nr\rangle\langle nr|. \quad (124)$$

It is left as an exercise to show that the operator P_n actually is a projector, and that it projects onto eigenspace \mathcal{E}_n of A . The set of projectors P_n for different n also satisfy

$$P_n P_m = P_m P_n = \delta_{nm} P_n. \quad (125)$$

As we say, the projectors P_n form an orthogonal set.

It follows immediately from the definition (124) and the resolution of the identity (82) that

$$\sum_n P_n = 1, \quad (126)$$

which is yet another way of writing the completeness relation for A .

One can also express the operator A itself in terms of its projectors. We simply let A act on both sides of the resolution of the identity, Eq. (82), and use Eqs. (123) and (124). The result is

$$A = \sum_n a_n P_n. \quad (127)$$

In the case of the continuous spectrum, a projector can be associated with an interval $I = [\nu_1, \nu_2]$ of the parameter describing the continuous spectrum. For if we normalize the eigenkets of some operator A as in Eq. (108), then it is easy to show that the operator

$$P_I = \int_{\nu_1}^{\nu_2} d\nu \sum_r |\nu r\rangle \langle \nu r| \quad (128)$$

is a projector, and that the projectors for two intervals I and I' are orthogonal if and only if I and I' do not intersect.

The use of projection operators is generally preferable to the use of eigenvectors, because the latter are not unique. Even when nondegenerate and normalized, an eigenvector is subject to multiplication by an arbitrary phase factor, and in the case of degeneracies, an orthonormal basis can be chosen in the degenerate eigenspace in many ways. However, it is easy to show that the projector defined in Eq. (124) is invariant under all such redefinitions, as should be obvious from the geometrical fact that P_n projects onto \mathcal{E}_n .

25. A Function of an Observable

The concept of a *function* of an observable arises in many places. We begin with the case of the discrete spectrum. If A is an observable with discrete eigenvalues a_1, a_2, \dots and corresponding projectors P_1, P_2, \dots , we define $f(A)$ by

$$f(A) = \sum_n f(a_n) P_n. \quad (129)$$

It is easy to show that when f is a polynomial or power series, $f(A)$ is the same as the obvious polynomial or power series in the operator A ; but Eq. (129) defines a function of an observable in more general cases, such as the important cases $f(x) = 1/(x - x_0)$, or even $f(x) = \delta(x - x_0)$. When A has a continuous spectrum, we can no longer express $f(A)$ in terms of projectors, but we can write it in terms of the complete basis of eigenkets, in the notation of Eq. (109). In this case we have

$$f(A) = \sum_{nr} f(a_n) |nr\rangle \langle nr| + \int d\nu \sum_r f(a(\nu)) |\nu r\rangle \langle \nu r|. \quad (130)$$

26. Operators and Kets versus Matrices and Vectors

We now consider the relation between abstract kets, bras, and operators and the vectors and matrices of numbers that are used to represent those abstract objects. Consider, for example, the eigenvalue problem $A|u\rangle = a_n|u\rangle$ for a Hermitian operator A . To convert this into numerical form, we simply choose an arbitrary orthonormal basis $\{|k\rangle\}$, $\langle k|\ell\rangle = \delta_{k\ell}$, and insert a resolution of the identity $\sum |\ell\rangle\langle\ell|$ before the ket $|u\rangle$ on both sides. We then multiply from the left with bra $\langle k|$, to obtain

$$\sum_{\ell} \langle k|A|\ell\rangle\langle\ell|u\rangle = c_n\langle k|u\rangle. \quad (131)$$

Thus, with

$$A_{k\ell} = \langle k|A|\ell\rangle, \quad u_k = \langle k|u\rangle, \quad (132)$$

we have

$$\sum_{\ell} A_{k\ell} u_{\ell} = a_n u_k. \quad (133)$$

The matrix elements $A_{k\ell}$ form a Hermitian matrix,

$$A_{k\ell} = A_{\ell k}^*, \quad (134)$$

and the eigenvalue problem becomes that of determining the eigenvalues and eigenvectors of a Hermitian matrix. Now the eigenvector u_k , as a column vector of numbers, contains the components of the eigenket $|u\rangle$ with respect to the chosen basis. Of course, if the Hilbert space is infinite-dimensional, then the matrix $A_{k\ell}$ is also infinite-dimensional. In practice, one usually truncates the infinite-dimensional matrix at some finite size, in order to carry out numerical calculations. One can hope that if the answers seem to converge as the size of the truncation is increased, then one will obtain good approximations to the true answers (this is the usual procedure).

The above is for a discrete basis. Sometimes we are also interested in a continuous basis, for example, the eigenbasis $|x\rangle$ of the operator \hat{x} . Again working with the eigenvalue problem $A|u\rangle = a_n|u\rangle$, we can insert a resolution of the identity $\int dx' |x'\rangle\langle x'|$ before $|u\rangle$ on both sides, and multiply from the left by $\langle x|$. Then using the orthogonality relation (98), we obtain

$$\int dx' \langle x|A|x'\rangle\langle x'|u\rangle = a_n\langle x|u\rangle, \quad (135)$$

or,

$$\int dx' A(x, x')u(x') = a_n u(x), \quad (136)$$

where

$$A(x, x') = \langle x|A|x'\rangle, \quad u(x) = \langle x|u\rangle. \quad (137)$$

We see that the “matrix element” of an operator with respect to a continuous basis is nothing but the kernel of the integral transform that represents the action of that operator in the given basis. (The kernel of an integral transform such as (136) is the function $A(x, x')$ that appears under the integral.)

27. Traces

The trace of a matrix is defined as the sum of the diagonal elements of the matrix. Likewise, we define the trace of an operator A acting on a Hilbert space as the sum of the diagonal matrix elements of A with respect to any orthonormal basis. Thus, in a discrete basis $\{|k\rangle\}$ we have

$$\text{tr } A = \sum_k \langle k|A|k\rangle. \quad (138)$$

One can easily show that it does not matter which basis is chosen; the trace is a property of the operator alone. One can also use a continuous basis to compute the trace (replacing sums by integrals). The operator in Eq. (138) need not be Hermitian, but if it is, then the trace is equal to the sum of the eigenvalues, each weighted by the order of the degeneracy,

$$\text{tr } A = \sum_n g_n a_n, \quad (139)$$

where g_n is the order of the degeneracy of a_n . This is easily proven by choosing the basis in Eq. (138) to be the orthonormal eigenbasis of A . But if A has a continuous range of eigenvalues, then the result diverges (or is not defined). Even if the spectrum of A is discrete, the sum (139) need not converge.

The trace of a product of operators is invariant under a cyclic permutation of the product. For example, in the case of three operators, we have

$$\text{tr}(ABC) = \text{tr}(BCA) = \text{tr}(CAB). \quad (140)$$

This property is easily proven. Another simple identity is

$$\text{tr } |\alpha\rangle\langle\beta| = \langle\beta|\alpha\rangle. \quad (141)$$

If A and B are operators, we are often interested in a kind of “scalar product” of A with B , especially when we are expressing some operators as linear combinations of other operators (the Pauli matrices, for example). The most useful definition of a scalar product of two operators is

$$(A, B) = \text{tr}(A^\dagger B) = \sum_{mn} A_{mn}^* B_{mn}, \quad (142)$$

where the final expression involves the matrix elements of A and B with respect to an arbitrary orthonormal basis. Thus, the scalar product of an operator with itself is the sum of the absolute squares of the matrix elements,

$$(A, A) = \sum_{mn} |A_{mn}|^2, \quad (143)$$

which vanishes if and only if $A = 0$.

28. Mathematical Notes

We collect here some remarks of a mathematical nature on the preceding discussion. For a precise treatment of the mathematics involved, see the references.

The set of normalizable wave functions $\{\psi(x)\}$ is infinite-dimensional because given any finite sequence of normalizable wave functions, (ψ_1, \dots, ψ_n) , one can always find a new normalizable wave function that is linearly independent of those given.

A Hilbert space is a complex, normed vector space that is *complete*. The latter property means that all Cauchy sequences converge to a point of the space. This property is automatic in finite dimensions. In addition, many authors require a Hilbert space to be *separable*, which means that it possesses a countable basis. Sakurai is mistaken when he says that a Hilbert space has an uncountable basis; he is thinking of the “bases” of the continuous spectrum, such as the set $\{\delta(x - x_0) | x_0 \in \mathbb{R}\}$, but these basis wave functions are not normalizable and do not belong to Hilbert space. In field theory, one must deal with spaces that are not separable (since they possess uncountable bases of normalizable vectors, but not countable ones). Such spaces are called *Fock spaces*.

Properly speaking, functions $\psi(x)$ do not stand in one-to-one correspondence with elements of Hilbert space. The functions must be measurable, and two functions that differ by a subset of measure zero are considered equivalent. Thus, it is really an equivalence class of wave functions $\psi(x)$ that corresponds to a state vector $|\psi\rangle$. All these wave functions, however, give the same results in their physical predictions (such as the probability of finding a particle on some interval). One cannot measure a wave function at a point.

Sometimes Hermitian operators are also called *self-adjoint*; to be precise about this terminology, on a finite-dimensional Hilbert space “Hermitian” and “self-adjoint” are equivalent, but on an infinite-dimensional space this is true only if certain assumptions are made about the properties of the operators and their domains of definition.

The usual Hermitian operators encountered in quantum mechanics (for example, Hamiltonians) are complete, although the mathematical proof is often technical. In fact, the completeness property is intimately tied up with the physical interpretation of the measurement process, since (according to the usual interpretation of measurement) if a measurement were to correspond to an incomplete operator, it would mean that measuring all possible outcomes would result in a total probability less than unity. In this course (and in most of the literature on quantum mechanics) we assume the completeness of all Hermitian operators we encounter without further comment.

Regarding the first integral on the right in Eq. (106), which must vanish in order to show that $\langle \psi | \hat{p} | \phi \rangle = \langle \phi | \hat{p} | \psi \rangle^*$, we note that normalizability does not imply that a wave function $\psi(x)$ approaches zero as $x \rightarrow \pm\infty$. At issue here is the fact that the momentum operator \hat{p} is *unbounded*, that is, it maps some normalizable wave functions into unnormalizable ones (taking them out of Hilbert space). If, however, both $\hat{p}\psi$ and $\hat{p}\phi$ are normalizable, then the integral in question does vanish.

29. References

The standard early reference for physicists on the formalism of quantum mechanics is P. A. M. Dirac, *The Principles of Quantum Mechanics*, 4th ed (Oxford University Press, Oxford, 1974), while for mathematicians it is John von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, 1955) (this is the English translation of the German original). Modern textbooks that devote more than the usual attention to mathematical questions include Albert Messiah, *Quantum Mechanics* vols. I and II (John Wiley, New York, 1966) (translation of French original), and Leslie E. Ballentine, *Quantum Mechanics: A Modern Development* (World Scientific, Singapore, 1998). The standard modern reference on the mathematics of quantum mechanics is Michael Reed and Barry Simon, *Functional Analysis* revised ed (Academic Press, New York, 1980), and, by the same authors, *Fourier Analysis, Self-Adjointness* (Academic Press, New York, 1975). See also *Quantum Mechanics in Hilbert Space* by Eduard Prugovečki (Dover, Mineola, New York, 1981).

Problems

1. Consider the 2×2 matrices:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (144)$$

Matrices σ_x , σ_y and σ_z are called the *Pauli matrices*. We will also denote them σ_1 , σ_2 , σ_3 .

(a) Show that

$$\sigma_i \sigma_j = \delta_{ij} + i \epsilon_{ijk} \sigma_k, \quad (145)$$

where δ_{ij} is understood to be multiplied by I , where we use the summation convention (see Sec. E.3), and where ϵ_{ijk} is the completely antisymmetric Levi-Civita symbol (see Sec. E.15, and ask if you don't understand, since this notation will be used frequently in the course). Notice that Eq. (145) implies

$$[\sigma_i, \sigma_j] = 2i \epsilon_{ijk} \sigma_k, \quad (146)$$

and

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij}, \quad (147)$$

where the curly brackets are the anticommutator.

What we mean by a *vector operator* is a vector of operators, for example, $\mathbf{A} = (A_x, A_y, A_z)$, where A_x , A_y and A_z are operators. Now given any two vector operators \mathbf{A} , \mathbf{B} that commute with $\boldsymbol{\sigma}$ but not necessarily with each other, use Eq. (145) to show that

$$(\boldsymbol{\sigma} \cdot \mathbf{A})(\boldsymbol{\sigma} \cdot \mathbf{B}) = \mathbf{A} \cdot \mathbf{B} + i\boldsymbol{\sigma} \cdot (\mathbf{A} \times \mathbf{B}), \quad (148)$$

where

$$\boldsymbol{\sigma} = \sigma_x \hat{\mathbf{x}} + \sigma_y \hat{\mathbf{y}} + \sigma_z \hat{\mathbf{z}}. \quad (149)$$

We think of $\boldsymbol{\sigma}$ as a “vector” of 2×2 matrices.

Two vector operators are considered to commute if all of their components commute. Note that in general, $\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A}$, and $\mathbf{A} \times \mathbf{B} \neq -\mathbf{B} \times \mathbf{A}$.

(b) Let $\hat{\mathbf{n}}$ be an arbitrary unit vector and θ an arbitrary angle. Show that

$$\exp(-i\theta \boldsymbol{\sigma} \cdot \hat{\mathbf{n}}) = I \cos \theta - i(\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}) \sin \theta. \quad (150)$$

In this course we will often drop the I , it being understood that a scalar like $\cos \theta$ is multiplied by the identity matrix.

2. Lie algebraic techniques for operators.

(a) Consider two operators A, B that do not necessarily commute. Show that

$$e^A B e^{-A} = B + [A, B] + \frac{1}{2!}[A, [A, B]] + \frac{1}{3!}[A, [A, [A, B]]] + \dots \quad (151)$$

Hint: Replace A by λA , where λ is a parameter, and let the left-hand side be $F(\lambda)$. Find a differential equation satisfied by $F(\lambda)$, and solve it. Alternatively, expand $F(\lambda)$ in a Taylor series in λ .

(b) Let $A(t)$ be an operator that depends on time. Derive the following operator identity:

$$\frac{d(e^A)}{dt} e^{-A} = \sum_{n=0}^{\infty} \frac{1}{(n+1)!} L_A^n \left(\frac{dA}{dt} \right), \quad (152)$$

where

$$L_A(X) = [A, X], \quad L_A^2(X) = [A, [A, X]], \quad \dots, \quad (153)$$

where X is an arbitrary operator. Also, $L_A^0(X) = X$. Do not assume that A commutes with dA/dt . We will use this identity when we study the nonrelativistic limit of the Dirac equation.

3. Pauli matrices and traces.

(a) If A and B are square matrices, show that

$$\text{tr}(AB) = \text{tr}(BA). \quad (154)$$

Also prove Eq. (140).

(b) Show that $\text{tr} \sigma_i = 0$. It may be easiest just to use the definitions (144). Then use Eq. (145) to find a simple expression for $\text{tr}(\sigma_i \sigma_j)$.

(c) The set of four 2×2 matrices, $(I, \boldsymbol{\sigma})$ form a basis in the space of 2×2 matrices, that is, an arbitrary 2×2 matrix can be expressed as a linear combination of these four matrices. Thus, if M is an arbitrary 2×2 matrix, it can be expressed as

$$M = aI + \mathbf{b} \cdot \boldsymbol{\sigma}, \quad (155)$$

This follows from the fact that I and the three σ_i are linearly independent, as can be seen just by looking at the four matrices. Find neat expressions for the expansion coefficients, a and $\mathbf{b} = (b_1, b_2, b_3)$. Do this by taking traces, or by multiplying by a Pauli matrix and then taking traces. Also show that M is Hermitian if and only if a, \mathbf{b} are real.

(d) Now suppose that M is nonzero in only the (r, s) slot, where it has a 1. That is, let

$$M_{ij} = \delta_{ir} \delta_{js}. \quad (156)$$

Use this in Eq. (155) to find a nice expression for $(\sigma_m)_{ij}(\sigma_m)_{kl}$ (summation convention everywhere, so sum on m in this expression) in terms of Kronecker deltas.

4. Some easy proofs from the notes.

(a) Show that Eq. (48) follows from Eq. (47).

(b) Prove Eqs. (55), (58), and (59).

(c) Prove that the product of two Hermitian operators is Hermitian if and only if they commute.

(d) Show that

$$\text{tr}(|\alpha\rangle\langle\beta|) = \langle\beta|\alpha\rangle. \quad (157)$$

5. The problem of finding the eigenkets and eigenbras of an arbitrary operator is more complicated and full of exceptions than in the case of Hermitian operators. There are, however, other classes of operators that share many of the nice properties of Hermitian operators. These include anti-Hermitian and unitary operators.

We define an operator to be *normal* if it commutes with its Hermitian conjugate, $[A, A^\dagger] = 0$. Notice that Hermitian, anti-Hermitian, and unitary operators are normal. In the following you may assume that you are working on a finite-dimensional Hilbert space.

(a) Show that if A is normal, and $A|u\rangle = a|u\rangle$ for some nonzero $|u\rangle$, then $A^\dagger|u\rangle = a^*|u\rangle$. Thus, the eigenbras of A are the Hermitian conjugates of the eigenkets, and the left spectrum is identical to the right spectrum. Hint: it is not necessary to introduce orthonormal bases or anything of the kind.

(b) Show that the eigenspaces corresponding to distinct eigenvalues of a normal operator are orthogonal. This is a generalization of the easy and familiar proof for Hermitian operators.

6. Some things that will be needed in subsequent problems. You may assume that you are working on a finite-dimensional Hilbert space.

(a) If A is an observable, show that

$$P_n = \prod_{k \neq n} \frac{A - a_k}{a_n - a_k}, \quad (158)$$

where P_n is the projector onto the n -th eigenspace of A . This shows that the projector P_n is a function of the operator A (see Sec. 25).

(b) Show that if

$$\langle \psi | A | \psi \rangle = 0 \tag{159}$$

for all $|\psi\rangle$, then $A = 0$. (A is not necessarily Hermitian.) Would this be true on a real vector space? (A real vector space is one in which the coefficients allowed in forming linear combinations of basis vectors are real. In a complex vector space, such as Hilbert space, these coefficients are allowed to be complex.)

(c) Let U be a linear operator, and let $|\psi'\rangle = U|\psi\rangle$. Show that $\langle \psi' | \psi' \rangle = \langle \psi | \psi \rangle$ for all kets $|\psi\rangle$ if and only if U is unitary.

7. Let V be a real vector space with a positive definite scalar product. If x and y are vectors in V , we denote their scalar product by (x, y) . This is like the scalar product $\langle x | y \rangle$ in the Hilbert spaces of quantum mechanics, except only real coefficients are allowed when forming linear combinations of vectors. We interpret (x, x) as the square of the length of the vector x . Show that the Schwarz inequality (28) implies that the shortest distance between two points is a straight line.