

## Chapter 4. Operators

§ 1 *Introduction.* In this chapter we show that the expression for the average value of an observable leads us to introduce a new object: the operator  $\hat{A}$  representing the observable  $A$ . This object is so central to quantum mechanics that no progress can be made without thoroughly understanding its properties. We define the adjoint and the inverse of an operator and examine unitary and hermitian operators, which have special properties and play an essential role in quantum mechanics.

In addition, we introduce the eigenvalue problem for operators. We show that the pure states of an observable  $A$  are the eigenstates of the operator  $\hat{A}$  and that the spectrum of  $A$  consists of the eigenvalues of  $\hat{A}$ .

§ 2 *Average values.* If we perform an experiment that puts a system in a state  $|\psi\rangle$  and measure the value of an observable  $A$ , we can only know the probability that the measurement gives a particular result. In such a situation, it is common to describe the measurement by determining the average value of  $A$ . It is therefore necessary to examine how such average values are calculated in quantum mechanics. I start by reminding you how the average of a quantity is defined in probability theory.

If we measure a quantity  $X$ , which can take the values  $x_1, x_2, x_3, \dots$  with the probabilities  $p_1, p_2, p_3, \dots$ , then the average value  $\langle X \rangle$  of  $X$  is defined by

$$\langle X \rangle \equiv \sum_{i=1}^{\infty} x_i p_i \quad (1)$$

If  $X$  takes continuous values  $x \in D$  with the probability distribution  $p(x)$  then

$$\langle X \rangle \equiv \int_{x \in D} x p(x) dx \quad (2)$$

These formulae are the ones you use to calculate the average grade in a course.

Let us use this for quantum mechanics. If a measurement of  $A$  can give one of the values  $a_1, a_2, \dots$  or  $\alpha \in D$  and the system is in state  $|\psi\rangle$ , the average value of  $A$  is

$$\langle A \rangle_{\psi} \equiv \sum_{n=1}^{\infty} P_{\psi}(a_n) a_n + \int_{\alpha \in D} d\alpha P_{\psi}(\alpha) \alpha \quad (3)$$

Here  $P_\psi(a_n)$  is the probability that a measurement of  $A$ , when the system is in state  $|\psi\rangle$ , yields  $a_n$ , and  $P_\psi(\alpha) d\alpha$  is the probability that the measurement of  $A$  yields a value between  $\alpha$  and  $\alpha + d\alpha$ . The notation  $\langle A \rangle_\psi$  is often used in quantum mechanics for the average value of  $A$  when the system is in state  $|\psi\rangle$ . When it is clear which state the system is in, we use  $\langle A \rangle$ .

§ 3 *The operator corresponding to an observable.* We can rewrite Eq. 3 in a different form. The probability  $P_\psi(a_n)$  is

$$P_\psi(a_n) = |\langle a_n | \psi \rangle|^2 = \langle a_n | \psi \rangle^* \langle a_n | \psi \rangle = \langle \psi | a_n \rangle \langle a_n | \psi \rangle \quad (4)$$

Similarly,

$$P_\psi(\alpha) = \langle \psi | \alpha \rangle \langle \alpha | \psi \rangle \quad (5)$$

Using Eqs. 4 and 5 allows us to write Eq. 3 as

$$\langle A \rangle_\psi = \sum_{n=1}^{\infty} \langle \psi | a_n \rangle a_n \langle a_n | \psi \rangle + \int_{\alpha \in D} d\alpha \langle \psi | \alpha \rangle \alpha \langle \alpha | \psi \rangle \quad (6)$$

As was explained in Chapter 3, Dirac wrote this expression as

$$\langle A \rangle_\psi = (\langle \psi |) \left( \sum_{n=1}^{\infty} |a_n\rangle a_n \langle a_n| + \int_{\alpha \in D} d\alpha |\alpha\rangle \alpha \langle \alpha| \right) (|\psi\rangle) \quad (7)$$

and considered it to consist of the symbol

$$\hat{A} \equiv \sum_{n=1}^{\infty} |a_n\rangle a_n \langle a_n| + \int_{\alpha \in D} d\alpha |\alpha\rangle \alpha \langle \alpha| \quad (8)$$

sandwiched between the bra  $\langle \psi |$  and the ket  $|\psi\rangle$ .  $\hat{A}$  is an operator, which acts on an arbitrary ket  $|\psi\rangle$  to give another ket

$$\hat{A}|\psi\rangle = \sum_{n=1}^{\infty} |a_n\rangle a_n \langle a_n | \psi \rangle + \int_{\alpha \in D} d\alpha |\alpha\rangle \alpha \langle \alpha | \psi \rangle \quad (9)$$

Here  $\langle a_n | \psi \rangle$  and  $\langle \alpha | \psi \rangle$  are complex numbers.  $a_n$  and  $\alpha$  are real numbers, since they are possible results of a measurement of  $A$ .  $|a_n\rangle a_n \langle a_n | \psi \rangle$  is a ket because multiplying the ket  $|a_n\rangle$  by the number  $a_n \langle a_n | \psi \rangle$  gives a ket. The sum in Eq. 9 is a ket because a sum of kets is a ket. The integral in Eq. 9 also gives a ket, because an integral is a glorified sum.

However, we might have a problem when we state that  $\hat{A}|\psi\rangle$  is a ket, because the sum has an infinite number of terms and without establishing convergence we are not sure that the sum is meaningful. We leave such subtleties for the mathematicians, mainly because in almost all practical problems we need only a finite number of terms in the sum, and convergence is no longer an issue.

Note that we have a good reason to use the notation  $\hat{A}$  in Eq. 8: the right-hand side of that equation contains only quantities tied to the observable  $A$ ; in particular,  $\hat{A}$  does not depend on the state  $|\psi\rangle$  of the system on which we perform the measurement.

If the observable is energy, and it has the spectrum  $E_1, E_2, \dots$  and  $e \in [0, \infty]$ , we can associate with it the operator

$$\hat{H} = \sum_{n=1}^{\infty} |E_n\rangle E_n \langle E_n| + \int_0^{\infty} de |e\rangle e \langle e| \quad (10)$$

For example, the pure states of the hydrogen atom are  $|n, \ell, m\rangle$  with  $n = 1, 2, \dots$ ,  $\ell = 0, 1, \dots, n-1$ , and  $m = -\ell, -\ell+1, \dots, \ell-1, \ell$ . The energy of a bound state  $|E_n\rangle$  is<sup>1</sup>

$$E_n = -\frac{\mu e^4}{2(4\pi\epsilon_0)^2 \hbar^2} \frac{1}{n^2}, \quad n = 1, 2, \dots$$

The energy of the dissociated electron-proton pair is

$$e_k = \frac{\hbar^2 k^2}{2\mu}, \quad k \in [0, \infty],$$

where  $\hbar k$  is the relative momentum and  $\mu$  is the reduced mass; the corresponding ket is denoted by  $|k\rangle$ .

According to the theory we have just developed, the operator corresponding to the energy of the hydrogen atom is

$$\begin{aligned} \hat{H} = & - \sum_{n=1}^{\infty} \sum_{\ell=0}^{n-1} \sum_{m=-\ell}^{\ell} |n, \ell, m\rangle \frac{\mu e^4}{2(4\pi\epsilon_0)^2 \hbar^2} \frac{1}{n^2} \langle n, \ell, m| \\ & + \int_{-\infty}^{+\infty} dk_x dk_y dk_z |k_x, k_y, k_z\rangle \frac{\hbar^2 k^2}{2\mu} \langle k_x, k_y, k_z| \end{aligned} \quad (11)$$

---

<sup>1</sup> $\mu$  is the reduced mass,  $e$  is the charge of the electron, and  $\epsilon_0$  is the permittivity of vacuum

We use the notation  $\hat{H}$  because the operator corresponding to the energy is called the Hamiltonian. Eq. 11 is a correct representation of the Hamiltonian of the hydrogen atom, even though it looks extremely different from the widely used formula:

$$\hat{H} = -\frac{\hbar^2}{2\mu} \left[ \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right] - \frac{e^2}{4\pi\epsilon_0 r} \quad (12)$$

where  $r = \sqrt{x^2 + y^2 + z^2}$  is the distance between the electron and the proton.

**Exercise 1** Write a formula similar to Eq. 11 for the operator corresponding to the angular momentum squared. The pure states of angular momentum squared are  $|\ell, m\rangle$  with  $\ell = 0, 1, 2, \dots$  and  $m = -\ell, -\ell + 1, \dots, \ell$ .

**§ 4** *The pure states of an observable  $A$  are eigenstates of the operator  $\hat{A}$  and the spectrum of  $A$  consists of all eigenvalues of  $\hat{A}$ .* The result of acting with the operator  $\hat{A}$  on a pure state  $|a_n\rangle$  of  $A$  is particularly simple because of the orthonormality conditions

$$\langle a_n | a_m \rangle = \delta_{nm} \quad (13)$$

and

$$\langle a_n | \alpha \rangle = 0 \quad (14)$$

Indeed, we have

$$\begin{aligned} \hat{A}|a_n\rangle &= \sum_{m=1}^{\infty} |a_m\rangle a_m \langle a_m | a_n \rangle + \int_D d\alpha |\alpha\rangle \alpha \langle \alpha | a_n \rangle \\ &= \sum_{m=1}^{\infty} |a_m\rangle a_m \delta_{mn} = a_n |a_n\rangle \end{aligned} \quad (15)$$

and, similarly,

$$\hat{A}|\alpha\rangle = \alpha|\alpha\rangle \quad (16)$$

for  $\alpha$  in the continuous spectrum of  $A$ .

In operator theory, the equation

$$\hat{A}|\psi\rangle = \lambda|\psi\rangle, \quad (17)$$

where  $|\psi\rangle$  is an *unknown ket* and  $\lambda$  is an *unknown number*, is called the eigenvalue equation of the operator  $\hat{A}$ . The eigenvalue problem for operators is so important to quantum mechanics that I will dedicate a separate chapter to its mathematics.

Any ket  $|\psi\rangle$  that satisfies Eq. 17 is called an *eigenket* or an *eigenvector* of  $\hat{A}$ ; any  $\lambda$  that satisfies Eq. 17 is called an *eigenvalue* of  $\hat{A}$ . Eigenvalues and eigenkets come in pairs. If  $\hat{A}|\psi_1\rangle = \lambda_1|\psi_1\rangle$ , we say that  $|\psi_1\rangle$  is the eigenket associated with the eigenvalue  $\lambda_1$ , or that  $\lambda_1$  is the eigenvalue associated with the eigenket  $|\psi_1\rangle$ .

The formulae in Eqs. 15 and 16 tell us that the *pure states*  $|a_n\rangle$  and  $|\alpha\rangle$  are *eigenfunctions of the operator  $\hat{A}$  and that the spectrum of  $\hat{A}$  consists of the eigenvalues of  $\hat{A}$ .*

People who use quantum mechanics spend most of their time solving various eigenvalue problems. At this point, we are not able to discuss how this is done, because we have a “chicken and egg” problem. We defined operators in terms of pure states and then showed that we can calculate the pure states by solving the eigenvalue problem for the operator; we are running in a circle!. To make progress, we must add to our postulates dynamical laws that allow us to obtain *explicit expressions* for the operators representing observables. Once this is done, we can solve the eigenvalue problem for the operator to obtain the eigenstates and the eigenvalues. The eigenvalues give us the spectrum of the observable, which are the values that the observable can have in a measurement. The eigenstates, which are the pure states of the observable, can be used to calculate the probability that a measurement yields a specific value from the spectrum.

§ 5 *The eigenvalue equation for  $f(\hat{A})$ .* Since we define

$$f(\hat{A}) \equiv \sum_{n=1}^{\infty} |a_n\rangle f(a_n) \langle a_n| + \int_D d\alpha |\alpha\rangle f(\alpha) \langle \alpha| \quad (18)$$

it is easy to show that

$$f(\hat{A})|a_n\rangle = f(a_n)|a_n\rangle \quad (19)$$

and

$$f(\hat{A})|\alpha\rangle = f(\alpha)|\alpha\rangle \quad (20)$$

Thus we see that  $f(\hat{A})$  has the same eigenfunctions as  $\hat{A}$  and its eigenvalues are  $f(a_n)$ ,  $f(\alpha)$ .

§ 6 *The operators representing observables are linear.* You can easily verify that the operator

$$\hat{A} = \sum_{n=1}^{\infty} |a_n\rangle a_n \langle a_n| + \int_D d\alpha |\alpha\rangle \alpha \langle \alpha| \quad (21)$$

has the property:

$$\hat{A}(|\phi\rangle + \eta|\psi\rangle) = \hat{A}|\phi\rangle + \eta\hat{A}|\psi\rangle \quad (22)$$

where  $\eta$  is a complex number and  $|\phi\rangle$  and  $|\psi\rangle$  are two arbitrary kets in the space in which  $\hat{A}$  is defined. An operator for which Eq. 22 is valid is called *linear*.

§ 7 *Other linear operators.* I have pointed out that there is a one-to-one correspondence between kets and elements of the linear spaces  $L^2$ , or  $\ell^2$ , or  $C^N$ . It will turn out that a similar correspondence exists between operators defined in the space of kets and operators defined on  $L^2$ , or  $\ell^2$ , or  $C^N$ . This is why I give here examples of linear operators in  $L^2$ ,  $\ell^2$ , and  $C^N$ .

*Operators in  $L^2$ .* The set of differentiable functions  $f(x)$  satisfying

$$\int_{-\infty}^{+\infty} f(x)^* f(x) dx < \infty \quad (23)$$

forms the linear space  $L^2$ . The derivative  $\frac{d}{dx}$  is an operator in this space if  $f' \equiv df/dx$  also satisfies  $\int_{-\infty}^{+\infty} f'(x)^* f'(x) dx < \infty$ . The derivative is a linear operator because if  $f(x)$  and  $g(x)$  belong to  $L^2$  then

$$\frac{d}{dx} (af(x) + g(x)) = a \frac{df(x)}{dx} + \frac{dg(x)}{dx}$$

The operator  $\hat{O}$  defined by the expression

$$\hat{O}f = \int K(x-y)f(y) dy \equiv \lambda(x) \quad (24)$$

is an operator on functions  $f \in L^2$  if the function  $\lambda(x)$  satisfies the condition in Eq. 23. It is easy to see that  $\hat{O}$  is a linear operator.

*Operators in  $C^3$ .* In the space of three-dimensional vectors  $\vec{x} = \{x_1, x_2, x_3\}$  with complex components  $x_i$ ,  $i = 1, 2, 3$ , a matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad (25)$$

acts on  $\vec{x}$  to give another vector, having the components

$$A\vec{x} \equiv \{a_{11}x_1 + a_{12}x_2 + a_{13}x_3, a_{21}x_1 + a_{22}x_2 + a_{23}x_3, a_{31}x_1 + a_{32}x_2 + a_{33}x_3\} \quad (26)$$

It is easy to see from this definition that  $A(\vec{x} + \alpha\vec{y}) = A\vec{x} + \alpha A\vec{y}$ . A matrix is therefore a linear operator in the  $\mathbb{C}^3$  space.

*Operators in  $\ell^2$ .* The  $\ell^2$  space consists of the lists

$$|f\rangle \equiv \{f_1, f_2, f_3, \dots\} \quad (27)$$

of complex numbers  $f_i$  that satisfy the condition

$$\sum_{i=1}^{\infty} f_i^* f_i < \infty \quad (28)$$

We can define an operator  $\hat{O}$  in this space through

$$|\eta\rangle \equiv \hat{O}|f\rangle = \{\eta_1, \eta_2, \eta_3, \dots\} \quad (29)$$

where the components in the list  $|\eta\rangle$  are defined by

$$\eta_i \equiv \sum_{j=1}^{\infty} O_{ij} f_j \quad (30)$$

$\hat{O}$  is a legitimate operator if

$$\sum_{i=1}^{\infty} \eta_i^* \eta_i < \infty \quad (31)$$

because the definition of the operator requires that  $|\eta\rangle$  be an element of  $\ell^2$ . You can verify easily that  $\hat{O}$  is a linear operator.

Quantum-mechanical equations can be written in different, *equivalent* forms involving kets, or elements of  $\mathbb{L}^2$ , or elements of  $\ell^2$ , or elements of  $\mathbb{C}^N$ . The operators representing observables can be defined in the space of kets, or in  $\mathbb{L}^2$ , or in  $\ell^2$ , or in  $\mathbb{C}^N$ . *They look very different but they give the same results when used to calculate quantities that can be measured.*

This ability of the theory to take different forms, while giving the same results when observables are calculated, is not peculiar to quantum mechanics. Classical mechanics can use Newton equation, Lagrange equation, the

Hamilton-Jacobi formulation, or a variational principle. If you use the same forces in any one of these “representations”, you get the same results for the quantities that can be measured. In statistical mechanics you can use a canonical ensemble, a microcanonical ensemble, or a grand canonical ensemble; these theories look different but give the same results for all observable quantities.

§ 8 In what follows we look at those mathematical properties of operators that are essential to quantum mechanics. We avoid all rigor in proofs because extensive experience with applications shows that the results are correct and that no ambiguity arises even though we lack the precision so dear to mathematicians.

---

**Exercise 2** In the space  $\mathbb{R}^2$  consisting of vectors of the form  $\vec{x} = \{x_1, x_2\}$ , the matrix

$$\hat{A} = \begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix} \quad (32)$$

is an operator. The eigenvalue problem is

$$\hat{A}\vec{x} = a\vec{x} \quad (33)$$

- (a) Solve the eigenvalue problem Eq. 33.
- (b) Eq. 33 has an infinite number of eigenstates for each eigenvalue. The eigenstates of  $\hat{A}$  that are of interest to physics must be pure states of the observable  $A$ . Therefore they must be orthonormal. Pick from among the solutions of Eq. 33 the set that satisfies this condition.

---

**Exercise 3** (a) Solve the eigenvalue problem

$$i\frac{df(x)}{dx} = \lambda f(x) \quad (34)$$

where  $\lambda$  is a real number and  $x \in [-\infty, +\infty]$ .

- (b) Show that the operator  $i\frac{d}{dx}$  has a continuous spectrum. (That is,  $\lambda$  can take any real value and still satisfy Eq. 34.)

- (c) Denote the eigenfunction corresponding to  $\lambda$  by  $|\lambda\rangle$ . Determine the eigenfunctions that satisfy the normalization condition

$$\langle \lambda | \lambda' \rangle = \delta(\lambda - \lambda')$$

Hint: use the known formula

$$\int_{-\infty}^{+\infty} e^{i(\lambda - \lambda')x} dx = 2\pi \delta(\lambda - \lambda')$$

**Exercise 4** Solve the eigenvalue equation

$$-\frac{d^2\psi}{dx^2} = k^2\psi, \quad x \in [-\infty, +\infty]$$

and perform the same analysis as in the previous exercise.

**Exercise 5** Solve the eigenvalue equation

$$-\frac{d^2\psi}{dx^2} = k^2\psi$$

with the conditions

$$x \in [0, L]$$

and

$$\psi(0) = \psi(L)$$

Perform the same analysis as in the two previous exercises.

**§ 9 Operations with operators.** An operator  $\hat{C}$  is the sum of the operators  $\hat{A}$  and  $\hat{B}$  if  $\hat{C}$  operating on any ket  $|\psi\rangle$  is equal to

$$\hat{C}|\psi\rangle = \hat{A}|\psi\rangle + \hat{B}|\psi\rangle \quad (35)$$

Since  $\hat{A}|\psi\rangle$  is a ket and  $\hat{B}|\psi\rangle$  is a ket, the sum  $\hat{A}|\psi\rangle + \hat{B}|\psi\rangle$  is also a ket.

An operator  $\hat{C}$  is the product  $\hat{A}\hat{B}$  of the operators  $\hat{A}$  and  $\hat{B}$  if

$$\hat{C}|\psi\rangle = \hat{A}(\hat{B}|\psi\rangle) \quad (36)$$

for any ket  $|\psi\rangle$ . The right-hand side of Eq. 36 indicates that first you act with  $\hat{B}$  on  $|\psi\rangle$  and obtain the ket  $|\eta\rangle \equiv \hat{B}|\psi\rangle$ , and then you act with  $\hat{A}$  on  $|\eta\rangle$ .

Now that we have defined the product of two operators, we can define how to raise an operator to a power.  $\hat{A}^2$  is

$$\hat{A}^2|\psi\rangle \equiv \hat{A}\hat{A}|\psi\rangle$$

and it is calculated as follows: act with  $\hat{A}$  on  $|\psi\rangle$  to obtain the ket  $|\eta\rangle \equiv \hat{A}|\psi\rangle$  and then act with  $\hat{A}$  on  $|\eta\rangle$  to obtain  $\hat{A}^2|\psi\rangle \equiv \hat{A}|\eta\rangle$ . The meaning of  $\hat{A}^n$ ,  $n \geq 2$ , should now be clear, and of course  $\hat{A}^1 = \hat{A}$ . By convention,  $\hat{A}^0 \equiv \hat{I}$  is the unit operator (the “do nothing” operator), which is defined by  $\hat{I}|\psi\rangle = |\psi\rangle$  for every ket  $|\psi\rangle$ .

The product of two operators *is not a commutative operation*:

$$\hat{A}\hat{B}|\psi\rangle \neq \hat{B}\hat{A}|\psi\rangle \quad \text{in general} \quad (37)$$

I am not saying that all operators must satisfy Eq. 37. Some operators do commute, that is, they satisfy

$$\hat{A}\hat{B}|\psi\rangle = \hat{B}\hat{A}|\psi\rangle$$

for all  $|\psi\rangle$ .

The expression  $\hat{A}\hat{B} - \hat{B}\hat{A}$  is called a *commutator* and is denoted by

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} \quad (38)$$

When two operators commute, their commutator is identically zero.

**§ 10 Examples.** Quantum mechanics can be formulated in a variety of spaces (a space of kets,  $L^2$ ,  $\ell^2$ ,  $C^N$ ) and because of this I give examples of operations with operators in these spaces.

In the space  $L^2$  define

$$\hat{A}f(x) \equiv \frac{d}{dx}f(x) \quad (39)$$

and

$$\hat{B}f(x) \equiv V(x)f(x) \quad (40)$$

where  $V(x)$  is some function such that  $V(x)f(x)$  is always an element of  $L^2$  when  $f(x)$  is.

We calculate that

$$\begin{aligned}\hat{A}\hat{B}f(x) &= \hat{A}(V(x)f(x)) = \frac{d}{dx}(V(x)f(x)) \\ &= \frac{dV(x)}{dx}f(x) + V(x)\frac{df(x)}{dx}\end{aligned}\quad (41)$$

We can test whether  $\hat{A}$  and  $\hat{B}$  commute, by calculating

$$\hat{B}\hat{A}f(x) = \hat{B}\left(\frac{df(x)}{dx}\right) = V(x)\frac{df(x)}{dx}\quad (42)$$

Clearly  $\hat{A}\hat{B}f(x) \neq \hat{B}\hat{A}f(x)$ . We can write the commutator:

$$[\hat{A}, \hat{B}]f(x) = (\hat{A}\hat{B} - \hat{B}\hat{A})f(x) = \hat{A}\hat{B}f(x) - \hat{B}\hat{A}f(x) = \frac{dV(x)}{dx}f(x)\quad (43)$$

The  $n$ -th power of the derivative operator is

$$\left(\frac{d}{dx}\right)^n f(x) = \frac{d}{dx} \frac{d}{dx} \cdots \frac{d}{dx} f(x) = \frac{d^n f(x)}{dx^n}\quad (44)$$

Consider now the operators  $\hat{A}$  and  $\hat{B}$  on  $\ell^2$  defined by

$$\hat{A}\{\eta_1, \eta_2, \dots\} = \left\{ \sum_{j=1}^{\infty} A_{1j}\eta_j, \sum_{j=1}^{\infty} A_{2j}\eta_j, \dots \right\}\quad (45)$$

$$\hat{B}\{\eta_1, \eta_2, \dots\} = \left\{ \sum_{j=1}^{\infty} B_{1j}\eta_j, \sum_{j=1}^{\infty} B_{2j}\eta_j, \dots \right\}\quad (46)$$

Here  $|\eta\rangle = \{\eta_1, \eta_2, \dots\}$  is an arbitrary element of  $\ell^2$  and the complex numbers  $A_{ij}$  and  $B_{ij}$  are given (they define the operators  $\hat{A}$  and  $\hat{B}$ ). It is easy to show that

$$\hat{B}\hat{A}|\eta\rangle = \left\{ \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} B_{1j}A_{ji}\eta_i, \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} B_{2j}A_{ji}\eta_i, \dots \right\}\quad (47)$$

**Exercise 6** Suppose that  $\{|a_n\rangle\}_{n=1}^{\infty}$  are pure states of an observable  $A$ . Show the following.

1. If  $\hat{P} = |a_n\rangle\langle a_n|$  then  $\hat{P}^2 = \hat{P}$ .

2. If  $\hat{P} = |a_n\rangle\langle a_n|$  and  $\hat{Q} = |a_m\rangle\langle a_m|$ , then  $\hat{P}\hat{Q} = 0$  when  $m \neq n$  and  $[\hat{P}, \hat{Q}] = 0$ .

3. If

$$\hat{A} = \sum_{i=1}^{\infty} |a_n\rangle a_n \langle a_n|$$

then

$$\hat{A}^m = \sum_{i=1}^{\infty} |a_n\rangle a_n^m \langle a_n|$$

**Exercise 7** If  $\{|a_n\rangle\}$ ,  $n = 1, 2, \dots$  and  $\{|b_n\rangle\}$ ,  $n = 1, 2, \dots$  are pure states of two different observables A and B, and if  $\hat{A}$  and  $\hat{B}$  are the corresponding operators, find a condition that ensures that  $[\hat{A}, \hat{B}] = 0$ .

**Exercise 8** Two operators  $\hat{A}$  and  $\hat{B}$  are defined by

$$\hat{A}f = \frac{\partial}{\partial x}f(x)$$

and

$$\hat{B}f = \int_{-\infty}^{+\infty} e^{-(x-y)^2} f(x) dx$$

Do these operators commute?

**§ 11 Functions of operators.** We have shown that to each observable A, having the spectrum  $a_1, a_2, \dots, a_n, \dots$ , and  $\alpha \in D$ , we can associate an operator

$$\hat{A} = \sum_{n=1}^{\infty} |a_n\rangle a_n \langle a_n| + \int_{\alpha \in D} |\alpha\rangle \alpha \langle \alpha| d\alpha \quad (48)$$

In what follows, we will need to define a variety of functions of an operator. If you do not pay attention, you would think that this is a trivial matter. If the function is  $f(x)$  and you want to know  $f(\hat{A})$ , you might replace  $x$  with  $\hat{A}$  and define  $f(\hat{A})$ . Let's see how this works if  $f(x) = \sqrt{x} \sin x$  and you assume  $f(\hat{A}) \equiv \sqrt{\hat{A}} \sin \hat{A}$ . This seems fine, but suppose the operator  $\hat{A}$  is  $\frac{\partial}{\partial x}$ : what is the meaning of  $\sqrt{\frac{\partial}{\partial x}}$  or of  $\sin \frac{\partial}{\partial x}$ ? This is not a trivial question.

A problem of a different kind appears if I want to define functions of two operators  $\hat{A}$  and  $\hat{B}$ . Take  $f(x, y) = xy^2$ . If  $x$  and  $y$  are numerical variables, then  $xy^2 = yxy = y^2x$ . However, if  $\hat{A}$  and  $\hat{B}$  do not commute,  $\hat{A}\hat{B}^2 \neq \hat{B}\hat{A}\hat{B} \neq \hat{B}^2\hat{A}$ . How do I define  $f(\hat{A}, \hat{B})$ ? Is it  $\hat{A}\hat{B}^2$ , or  $\hat{B}\hat{A}\hat{B}$ , or  $\hat{B}^2\hat{A}$ ?

Defining functions of operators is not straightforward. There is, however, a simple solution using Eq. 48.

For  $\hat{A}$  defined by Eq. 48, we define  $f(\hat{A})$  to be

$$f(\hat{A}) \equiv \sum_{n=1}^{\infty} |a_n\rangle f(a_n) \langle a_n| + \int_{\alpha \in D} |\alpha\rangle f(\alpha) \langle \alpha| d\alpha \quad (49)$$

In the right-hand side,  $f$  acts on the numbers  $a_n$  and  $\alpha$  so there is no ambiguity regarding the meaning of  $f(a_n)$  and  $f(\alpha)$ ; the result is a well-defined operator.

We know that  $|a_n\rangle$  and  $|\alpha\rangle$  are eigenvectors (eigenkets) of  $\hat{A}$ , and  $a_n$  and  $\alpha$  are the corresponding eigenvalues. If you want to calculate  $f(\hat{A})$ , you solve the eigenvalue problem for  $\hat{A}$  and use the result to write down the right-hand side of Eq. 49.

This definition may seem arbitrary but it has survived because it is useful. It is also consistent with other definitions in use. One of the most popular applies if  $f(x)$  can be represented by a convergent power-series expansion

$$f(x) = f_0 + f_1x + f_2x^2 + \cdots \quad (50)$$

For such functions,  $f(\hat{A})$  is defined by

$$f(\hat{A}) = f_0 + f_1\hat{A} + f_2\hat{A}^2 + \cdots \quad (51)$$

If we know  $\hat{A}$  — and we are supposed to know it if we want to define the function  $f(\hat{A})$  — we can evaluate the right-hand side of Eq. 51 because it involves only powers of  $\hat{A}$ . When we write such infinite series, we must be concerned with convergence. If the power series in Eq. 50 is convergent, this does not mean that the power series in Eq. 51 is convergent. Nor have we indicated how to define the domain of “values” of  $\hat{A}$  for which Eq. 51 is convergent. A physicist’s favorite method for dealing with convergence is to pray that the series is meaningful. In practice they approximate  $f(\hat{A})$  with a finite sum  $f(\hat{A}) \approx f_0 + f_1\hat{A} + f_2\hat{A}^2 + \cdots + f_N\hat{A}^N$ . By increasing  $N$ , they check that at some point adding a new term causes a negligible change in

$f(\hat{A})$ . When this happens, they are satisfied with their approximation. This is not always a guarantee that the series converges, but it works most of the time.

It is very easy to show that the definition provided by Eq. 49 is consistent with Eq. 51. Just replace  $f(a_n)$  and  $f(\alpha)$  in Eq. 48 with the expansion in Eq. 50.

The definition in Eq. 49 allows us to construct new and interesting operators. You will see that the operator  $\exp[-\hat{H}/k_B T]$  is essential in statistical mechanics and the operator  $\exp[-i\hat{H}t/\hbar]$  gives the time evolution of the state of a system in quantum mechanics.

**§ 12 Adjoint, Hermitian, inverse, and unitary operators: definitions.** We defined the operator  $\hat{A}$  through Eq. 48:

$$\hat{A} = \sum_{n=1}^{\infty} |a_n\rangle a_n \langle a_n| + \int_{\alpha \in D} |\alpha\rangle \alpha \langle \alpha| d\alpha \quad (52)$$

Here  $|a_n\rangle$  are the eigenstates of  $\hat{A}$  and  $a_n$  are its eigenvalues. Starting from this expression we can define new operators, which are all functions of the operator  $\hat{A}$ .

1. The *adjoint* (or *Hermitian conjugate*) of  $\hat{A}$  is

$$\hat{A}^\dagger = \sum_{n=1}^{\infty} |a_n\rangle a_n^* \langle a_n| + \int_{\alpha \in D} |\alpha\rangle \alpha^* \langle \alpha| d\alpha \quad (53)$$

2. An operator is called *Hermitian* (or *self-adjoint*) if all  $a_n$  and  $\alpha \in D$  are real numbers. From Eqs. 52 and 53 it follows that  $\hat{A}$  is Hermitian if and only if

$$\hat{A}^\dagger = \hat{A} \quad (54)$$

3. An operator is *unitary* if

$$|a_n| = 1 \text{ for all } n \text{ and } |\alpha| = 1 \text{ for all } \alpha \in D \quad (55)$$

Here  $|z|$  is the absolute value of the complex number  $z$ . You can verify that Eq. 55 is equivalent to requiring that

$$a_n = e^{ip_n} \text{ for } n \geq 1 \text{ and } \alpha = e^{ip(\alpha)} \text{ for } \alpha \in D \quad (56)$$

where  $p_n$  and  $p(\alpha)$  are real numbers and  $i = \sqrt{-1}$ .

## 4. The operator

$$\hat{A}^{-1} \equiv \sum_{n=1}^{\infty} |a_n\rangle \frac{1}{a_n} \langle a_n| + \int_{\alpha \in D} |\alpha\rangle \frac{1}{\alpha} \langle \alpha| d\alpha \quad (57)$$

is called the *inverse* of  $\hat{A}$ . Note that if some  $a_n$  is zero then this expression makes no sense, and in that case we say that  $\hat{A}$  *does not have an inverse* or that it is *singular*. If the integral in Eq. 57 is not well-defined, the operator does not have an inverse.

You will use these definitions, and the properties that follow from them, over and over and over. Other books use different, but equivalent, definitions. The ones used here lead most efficiently to useful results. The methods used in other books are more general but we do not need this additional generality here.

**§ 13 Notation.** In what follows we derive some important properties that follow from these definitions. Before proceeding, let us establish some notation. If  $|\eta\rangle \equiv \hat{A}|\psi\rangle$  then I either write

$$\langle \phi | \eta \rangle \equiv \langle \phi | \hat{A} \psi \rangle \quad (58)$$

or

$$\langle \phi | \eta \rangle \equiv \langle \phi | \hat{A} | \psi \rangle \quad (59)$$

Mathematicians tend to use Eq. 58 and physicists, Eq. 59.

I also use the notation

$$\langle \eta | \phi \rangle \equiv \langle \hat{A} \psi | \phi \rangle \quad (60)$$

**§ 14 Some properties of adjoint operators.**

*Property 1.* If  $\hat{A}^\dagger$  is the adjoint of  $\hat{A}$ , then for any  $|\phi\rangle$  and  $|\psi\rangle$  we have

$$\langle \phi | \hat{A} \psi \rangle = \langle \hat{A}^\dagger \phi | \psi \rangle \quad (61)$$

and

$$\langle \hat{A} \phi | \psi \rangle = \langle \phi | \hat{A}^\dagger \psi \rangle \quad (62)$$

Here  $\langle \hat{A}^\dagger \phi |$  and  $\langle \hat{A} \phi |$  are the bras corresponding to the kets  $\hat{A}^\dagger |\phi\rangle$  and  $\hat{A} |\phi\rangle$ , respectively. Eqs. 61 and 62 follow from the definition of the adjoint operator and the properties of the scalar product. To simplify our lives, I will discard

the integral in the definitions of the operators. Since the integral is the limit of a sum, what goes for the sum goes for the integral. Start with  $\langle \hat{A}^\dagger \phi | \psi \rangle$  and rewrite it using known relationships, to show that it is the same as  $\langle \phi | \hat{A} | \psi \rangle$ :

$$\begin{aligned}
 \langle \hat{A}^\dagger \phi | \psi \rangle &= \langle \psi | \hat{A}^\dagger \phi \rangle^* && \text{(used } \langle x | y \rangle = \langle y | x \rangle^*) \\
 &= \left( \sum_{n=1}^{\infty} \langle \psi | a_n \rangle a_n^* \langle a_n | \phi \rangle \right)^* && \text{(used Eq. 53)} \\
 &= \sum_{n=1}^{\infty} \langle \psi | a_n \rangle^* a_n \langle a_n | \phi \rangle && \text{(used } (zw)^* = z^* w^*) \\
 &= \sum_{n=1}^{\infty} \langle \phi | a_n \rangle a_n \langle a_n | \psi \rangle && \text{(used } \langle x | y \rangle = \langle y | x \rangle^*) \\
 &= \langle \phi | \hat{A} \psi \rangle && \text{(used Eq. 52)}
 \end{aligned}$$

This proves Eq. 61. For Eq. 62, first note that, from the definition (Eq. 53), it follows that

$$(\hat{A}^\dagger)^\dagger = \hat{A} \quad (63)$$

Therefore

$$\begin{aligned}
 \langle \phi | \hat{A}^\dagger \psi \rangle &= \langle (\hat{A}^\dagger)^\dagger \phi | \psi \rangle && \text{(used Eq. 61)} \\
 &= \langle \hat{A} \phi | \psi \rangle && \text{(used Eq. 63)}
 \end{aligned}$$

Many mathematicians prefer to use Eq. 61 as a definition of an adjoint operator, because it is more general than the one used here; it is valid for operators not covered by the definition in Eq. 52 on which the material developed here is based.

*Property 2.* If

$$\langle \hat{A} \psi | \phi \rangle = \langle \psi | \hat{B} \phi \rangle \quad (64)$$

for any  $|\phi\rangle$  and  $|\psi\rangle$ , then  $\hat{B}$  is the adjoint of  $\hat{A}$ .

Write Eq. 64 as

$$\langle \phi | \hat{A} \psi \rangle^* = \langle \psi | \hat{B} \phi \rangle \quad (65)$$

Since this is valid for any kets  $|\phi\rangle$  and  $|\psi\rangle$ , it is valid for the pure states  $|a_n\rangle$  of A (eigenstates of  $\hat{A}$ ). Therefore Eq. 65 gives (using Eq. 52 and orthonormality)

$$\langle a_n | \hat{B} | a_n \rangle = \langle a_n | \hat{A} | a_n \rangle^* = a_n^* \quad (66)$$

and, for  $m \neq n$ ,

$$\langle a_n | \hat{B} | a_m \rangle = \langle a_n | \hat{A} | a_m \rangle^* = 0 \quad (67)$$

Since the set of pure states is complete,  $\sum_n |a_n\rangle \langle a_n| = \hat{I}$  and so we can write  $\hat{B}$  as

$$\hat{B} = \hat{I} \hat{B} \hat{I} = \sum_n \sum_m |a_n\rangle \langle a_n | \hat{B} | a_m \rangle \langle a_m| = \sum_n |a_n\rangle a_n^* \langle a_n| \quad (68)$$

To obtain the last equality, I used Eqs. 66 and 67. Eq. 68 proves that  $\hat{B} = \hat{A}^\dagger$  (see Eq. 53).

We will often have to deal with products of operators, such as  $\hat{A}\hat{B}$ . What is the adjoint of the product? That is, if we set  $\hat{C} = \hat{A}\hat{B}$ , what is  $\hat{C}^\dagger$ ?

*Property 3.*  $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$

From Property 1, we know that we can write

$$\langle \phi | \hat{C}\psi \rangle = \langle \hat{C}^\dagger\phi | \psi \rangle \quad (69)$$

for any operator  $\hat{C}$ . Let us replace  $\hat{C}$  in the left-hand side of this equation with  $\hat{C} = \hat{A}\hat{B}$  and use Property 1 twice:

$$\langle \phi | \hat{C}\psi \rangle = \langle \phi | \hat{A}(\hat{B}\psi) \rangle = \langle \hat{A}^\dagger\phi | \hat{B}\psi \rangle = \langle \hat{B}^\dagger\hat{A}^\dagger\phi | \psi \rangle \quad (70)$$

This is true for any kets  $|\phi\rangle$  and  $|\psi\rangle$ . This means that (compare Eqs. 69 and 70)  $\hat{C}^\dagger = \hat{B}^\dagger\hat{A}^\dagger$ , which is what we set out to prove.

### § 15 Properties of a self-adjoint (Hermitian) operator.

*Property 4.* The eigenvalues of a Hermitian operator must be real numbers. This property follows from the results in §4, where we showed that  $a_n$  and  $\alpha$  are the eigenvalues of  $\hat{A}$ . If  $\hat{A}$  is Hermitian, then (by the definition)  $a_n$  and  $\alpha$  are real.

Property 4 is very important, for the following reason. Consider an observable  $A$  with the spectrum  $a_n$ ,  $n = 1, 2, \dots$  and  $\alpha \in D$ . These are the values that a measurement of  $A$  can give and they must be real numbers. You have also learned that the spectrum of  $A$  is identical to the set of all eigenvalues of the operator  $\hat{A}$  corresponding to the observable  $A$  (§4). Therefore all eigenvalues of  $\hat{A}$  must be real numbers. According to Property 3, this requirement is automatically fulfilled if  $\hat{A}$  is Hermitian. Therefore, we expect that *all operators corresponding to observables are Hermitian*.

*Property 5.*  $\hat{A}$  is self-adjoint (Hermitian) if and only if  $\hat{A} = \hat{A}^\dagger$ .

The proof is straightforward. By the definition of adjoint, if  $\hat{A}$  is defined by

$$\hat{A} = \sum_{n=1}^{\infty} |a_n\rangle a_n \langle a_n| \quad (71)$$

then

$$\hat{A}^\dagger = \sum_{n=1}^{\infty} |a_n\rangle a_n^* \langle a_n| \quad (72)$$

Obviously if each  $a_n$  is real (i.e.  $\hat{A}$  is Hermitian) then these two expressions are identical.

Conversely, if  $\hat{A} = \hat{A}^\dagger$ , then for any value of  $n$ ,

$$\langle a_n | \hat{A}^\dagger | a_n \rangle = \langle a_n | \hat{A} | a_n \rangle \quad (73)$$

Using Eq. 72 and the fact that  $\langle a_i | a_j \rangle = \delta_{ij}$ , it is easy to see that  $\langle a_n | \hat{A}^\dagger | a_n \rangle = a_n^*$ . Similarly, using Eq. 71,  $\langle a_n | \hat{A} | a_n \rangle = a_n$ . These, together with Eq. 73, imply that  $a_n^* = a_n$  and  $a_n$  must be a real number; therefore if  $\hat{A} = \hat{A}^\dagger$  then  $\hat{A}$  is Hermitian.

*Property 6.* If  $\hat{A}$  is self-adjoint, then  $\langle \hat{A}\phi | \psi \rangle = \langle \phi | \hat{A}\psi \rangle$ .

This follows from Properties 1 and 5.

**Exercise 9** Suppose that  $\hat{A}$  and  $\hat{B}$  are Hermitian. Prove the following.

- (a)  $(\hat{A}\hat{B})^\dagger = \hat{B}\hat{A}$ .
- (b)  $\hat{A} + \hat{B}$  is Hermitian.
- (c)  $[\hat{A}, \hat{B}]^\dagger = \hat{B}\hat{A} - \hat{A}\hat{B} = -[\hat{A}, \hat{B}]$   
where  $[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}$  is the commutator of  $\hat{A}$  with  $\hat{B}$ .
- (d)  $\{\hat{A}, \hat{B}\} \equiv \hat{A}\hat{B} + \hat{B}\hat{A}$  is Hermitian.

**§ 16 Inverse operators.** If  $\hat{A}$  is defined by

$$\hat{A} = \sum_{n=1}^{\infty} |a_n\rangle a_n \langle a_n| \quad (74)$$

then its inverse is

$$\hat{A}^{-1} \equiv \sum_{n=1}^{\infty} |a_n\rangle \frac{1}{a_n} \langle a_n| \quad (75)$$

As noted earlier, if any  $a_n$  in the spectrum of  $\hat{A}$  is equal to zero, then the operator  $\hat{A}$  has no inverse (is singular). We know that each  $a_n$  is an eigenvalue of  $\hat{A}$  and therefore we can state:

*Property 7.* If one of the eigenvalues of  $\hat{A}$  is equal to zero, then  $\hat{A}$  does not have an inverse.

*Property 8.* If  $\hat{A}$  and  $\hat{B}$  satisfy

$$\hat{A}\hat{B} = \hat{B}\hat{A} = \hat{I} \quad (76)$$

and all the eigenvalues of  $\hat{A}$  differ from zero, then  $\hat{B} = \hat{A}^{-1}$ .

Here  $\hat{I}$  is the unit operator. Note that the equality in Eq. 76 means that  $\langle \psi | \hat{A}\hat{B} | \phi \rangle = \langle \psi | \hat{B}\hat{A} | \phi \rangle = \langle \psi | \phi \rangle$  for every  $|\psi\rangle$  and  $|\phi\rangle$ .

Write

$$\hat{A} = \sum_{n=1}^{\infty} |a_n\rangle a_n \langle a_n| \quad (77)$$

I have dropped the integral since that term can be treated by analogy with the sum. From  $\hat{B}\hat{A} = \hat{I}$ , we have, for every  $n$ ,

$$\langle a_n | \hat{B}\hat{A} | a_n \rangle = \langle a_n | \hat{I} | a_n \rangle = 1 \quad (\text{used } \langle a_n | a_n \rangle = 1) \quad (78)$$

But also

$$\langle a_n | \hat{B}\hat{A} | a_n \rangle = a_n \langle a_n | \hat{B} | a_n \rangle \quad (\text{used } \hat{A}|a_n\rangle = a_n|a_n\rangle) \quad (79)$$

Comparing these two values for  $\langle a_n | \hat{B}\hat{A} | a_n \rangle$ , we see that

$$\langle a_n | \hat{B} | a_n \rangle = 1/a_n \quad (80)$$

Starting again from  $\hat{B}\hat{A} = \hat{I}$ , we have

$$\langle a_n | \hat{B}\hat{A} | a_m \rangle = \langle a_n | \hat{I} | a_m \rangle = 0 \quad (\text{used orthogonality}) \quad (81)$$

for every  $n$  and  $m \neq n$ , and also

$$\langle a_n | \hat{B}\hat{A} | a_m \rangle = a_m \langle a_n | \hat{B} | a_m \rangle \quad (\text{used } \hat{A}|a_m\rangle = a_m|a_m\rangle) \quad (82)$$

If  $a_m \neq 0$  for all  $m$ , comparison of Eqs. 81 and 82 leads to

$$\langle a_n | \hat{B} | a_m \rangle = 0, \quad n \neq m \quad (83)$$

We conclude from Eqs. 80 and 83 that when all the eigenvalues of  $\hat{A}$  are non-zero,

$$\langle a_n | \hat{B} | a_m \rangle = \delta_{nm}/a_m \quad (84)$$

Now write

$$\begin{aligned} \hat{B} &= \hat{I} \hat{B} \hat{I} = \sum_n \sum_m |a_n\rangle \langle a_n | \hat{B} | a_m \rangle \langle a_m| \\ &= \sum_n \sum_m |a_n\rangle \frac{\delta_{nm}}{a_m} \langle a_m| \quad (\text{used Eq. 84}) \\ &= \sum_n |a_n\rangle \frac{1}{a_n} \langle a_n| \end{aligned} \quad (85)$$

which, from the definition of inverse operator, means that  $\hat{B} = \hat{A}^{-1}$ .

Inverse operators are handy when we want to solve an equation of the form

$$\hat{A}|\psi\rangle = |\phi\rangle \quad (86)$$

where  $|\psi\rangle$  is unknown and  $\hat{A}$  and  $|\phi\rangle$  are known. Acting with  $\hat{A}^{-1}$  on Eq. 86 and using  $\hat{A}^{-1}\hat{A} = \hat{I}$  and  $\hat{I}|\psi\rangle = |\psi\rangle$  converts Eq. 86 to

$$|\psi\rangle = \hat{A}^{-1}|\phi\rangle \quad (87)$$

To this particular solution we can add any  $|\chi\rangle$  having the property

$$\hat{A}|\chi\rangle = 0 \quad (88)$$

Therefore the general solution of  $\hat{A}|\psi\rangle = |\phi\rangle$  is

$$|\psi\rangle = \hat{A}^{-1}|\phi\rangle + |\chi\rangle \quad (89)$$

since then

$$\hat{A}|\psi\rangle = \hat{A}(\hat{A}^{-1}|\phi\rangle + |\chi\rangle) = \hat{A}\hat{A}^{-1}|\phi\rangle + \hat{A}|\chi\rangle = \hat{I}|\phi\rangle + 0 = |\phi\rangle$$

This seems to be a very efficient way of solving Eq. 86, but the trouble is that in many cases  $\hat{A}^{-1}$  is hard to calculate. Nevertheless, the inverse of an operator is used often in quantum mechanics.

Consider the following equation in  $L^2$ :

$$\frac{d^2\psi(x)}{dx^2} = \phi(x)$$

You can formally write

$$\psi(x) = \left( \frac{d^2}{dx^2} \right)^{-1} \phi(x) + \chi(x)$$

where  $\chi(x)$  satisfies

$$\frac{d^2\chi(x)}{dx^2} = 0$$

But what is  $\left( \frac{d^2\psi(x)}{dx^2} \right)^{-1}$ ? In this case you can calculate it (do so), but, in general, finding the inverse of an operator is not easy.

§ 17 *Unitary operators.* The operator  $\hat{A}$  defined through

$$\hat{A} = \sum_{n=1}^{\infty} |a_n\rangle a_n \langle a_n| \quad (90)$$

where  $|a_n\rangle$  is an orthonormal set of kets, is called *unitary* if

$$a_n^* a_n = 1, \quad n = 1, 2, \dots \quad (91)$$

This is equivalent to requiring that

$$a_n = e^{i\phi_n}, \quad n = 1, 2, \dots \quad (92)$$

where  $\phi_n$  is a real number.<sup>2</sup>

*Property 9.* The kets  $|a_n\rangle$  entering into Eq. 90 for a unitary operator are the eigenkets of that operator and the numbers  $a_n$  are the corresponding eigenvalues. Therefore the eigenkets of a unitary operator are orthonormal and the absolute values of the eigenvalues are equal to 1 (see Eq. 91). The proof of these statements is through straightforward calculation.

---

<sup>2</sup>We have introduced Hermitian operators through Eq. 91 based on physical arguments. We cannot use the same arguments to lead us to the conclusion that unitary operators have the this form. The definition give here assumes that this form is correct and then it defines unitarity. This form implies that the kets  $|a_n\rangle$  are orthonormal and that  $a_n$  are the eigenvalues and  $|a_n\rangle$  are the eigenvectors of  $\hat{A}$ . Most texts define the adjoint operator  $A^\dagger$  through  $\langle \hat{A}^\dagger \psi | \phi \rangle \equiv \langle \psi | \hat{A} \phi \rangle$  and unitary operators through  $\hat{A} \hat{A}^\dagger = \hat{A}^\dagger \hat{A} = \hat{I}$ . Then one can show that an operator is unitary if and only if satisfies the definition given here (see Halmos, *Finite-Dimensional Vector Spaces*, §80, the section on normal transformations).

*Property 10.*  $\hat{A}$  is unitary if and only if it has an inverse and  $\hat{A}^{-1} = \hat{A}^\dagger$ .

The inverse of  $\hat{A} = \sum_{n=1}^{\infty} |a_n\rangle a_n \langle a_n|$  is  $\sum_{n=1}^{\infty} |a_n\rangle \frac{1}{a_n} \langle a_n|$ . Since  $\hat{A}$  is unitary,  $a_n = \exp[i\phi_n]$  for some real numbers  $\phi_n$  and

$$\frac{1}{a_n} = \exp[-i\phi_n] = a_n^*$$

Therefore

$$\hat{A}^{-1} = \sum_{n=1}^{\infty} |a_n\rangle \frac{1}{a_n} \langle a_n| \sum_{n=1}^{\infty} |a_n\rangle a_n^* \langle a_n| = \hat{A}^\dagger$$

Next, assume that  $\hat{A}^{-1} = \hat{A}^\dagger$ . This means (using the definitions of  $\hat{A}^{-1}$  and  $\hat{A}^\dagger$ ) that

$$\sum_{n=1}^{\infty} |a_n\rangle \frac{1}{a_n} \langle a_n| = \sum_{n=1}^{\infty} |a_n\rangle a_n^* \langle a_n| \quad (93)$$

Since  $\langle a_m | a_n \rangle = \delta_{nm}$ , Eq. 93 leads to (act with Eq. 93 on  $|a_m\rangle$  and then act with  $\langle a_m|$  on the result)

$$\frac{1}{a_m} = a_m^* \text{ for each } m \quad (94)$$

And this means that  $a_m a_m^* = 1$  for each  $m$  and therefore  $\hat{A}$  is unitary.

*Property 11.* If  $\hat{A}$  is Hermitian then the operator  $\hat{U} \equiv \exp[i\lambda\hat{A}]$ , where  $\lambda$  is a real number, is unitary. Also, if  $\hat{U}$  is a unitary operator, then it can be written as  $\hat{U} = \exp[i\hat{B}]$  where  $\hat{B}$  is Hermitian.

For  $\hat{A} = \sum_{n=1}^{\infty} |a_n\rangle a_n \langle a_n|$ , the operator  $\exp[i\lambda\hat{A}]$  is defined by

$$\exp[i\lambda\hat{A}] = \sum_{n=1}^{\infty} |a_n\rangle e^{i\lambda a_n} \langle a_n| \quad (95)$$

If  $\hat{A}$  is Hermitian then all the  $a_n$  are real numbers and so are all  $\lambda a_n$ ; therefore  $(e^{i\lambda a_n})^* e^{i\lambda a_n} = e^{-i\lambda a_n} e^{i\lambda a_n} = 1$  for all  $n$  and  $\exp[i\lambda\hat{A}]$  is unitary.

Conversely, if an operator  $\hat{U}$  is unitary then it must have the form

$$\hat{U} = \sum_{m=1}^{\infty} |u_m\rangle e^{i\phi_m} \langle u_m| \quad (96)$$

with  $\phi_m$  real.

Consider now the operator  $\hat{B}$  defined to have the spectrum  $\phi_m$  and the pure states  $|u_m\rangle$ , so that

$$\hat{B} = \sum_{m=1}^{\infty} |u_m\rangle \phi_m \langle u_m| \quad (97)$$

$\hat{B}$  is Hermitian because the  $\phi_m$  are real numbers. A function  $f(\hat{B})$  is defined by

$$f(\hat{B}) = \sum_{m=1}^{\infty} |u_m\rangle f(\phi_m) \langle u_m| \quad (98)$$

If  $f$  is the function  $f(x) = e^{i\lambda x}$  then

$$f(\hat{B}) \equiv e^{i\lambda\hat{B}} = \sum_{m=1}^{\infty} |u_m\rangle e^{i\lambda\phi_m} \langle u_m| \quad (99)$$

which means that  $e^{i\lambda\hat{B}}$  is unitary.

*Property 12.* If  $\hat{A}$  is Hermitian and  $\hat{U} \equiv \exp[i\hat{A}]$  then  $\hat{A}$  and  $\hat{U}$  have the same eigenvectors (eigenkets): if  $\hat{A}|a_n\rangle = a_n|a_n\rangle$  then  $\hat{U}|a_n\rangle = u_n|a_n\rangle$  and vice versa. In addition,  $u_n = e^{ia_n}$ .

The proof of this is hidden in the proof of Property 9. You can prove it by direct calculation.

**Exercise 10** Consider a function  $f(x)$  of the real variable  $x$  for which the values  $f(x)$  are complex ( $\exp[ix]$  is an example). Let  $\hat{A}$  be a Hermitian operator with spectrum  $\{a_n\}$  and eigenkets  $\{|a_n\rangle\}$ . Show that if

$$|f(a_n)| = 1 \quad \text{for all } n \quad (100)$$

then  $f(\hat{A})$  is unitary. Also show the converse: if  $f(\hat{A})$  is unitary then (100) must be true.

**Exercise 11** Show that if  $x$  varies along the real axis then  $y \equiv (x-i)/(x+i)$  varies on the unit circle in the complex plane.

**Exercise 12** Show that if  $\hat{A}$  is Hermitian then

$$\hat{B} \equiv (\hat{A} - i\hat{I})(\hat{A} + i\hat{I})^{-1}$$

is unitary (show first that  $\hat{A} + i\hat{I}$  has an inverse). This is called a Cayley transform, and it is the basis of the Crank-Nicholson method for solving differential equations.

---

*Property 13.*  $\hat{U}$  is unitary if and only if

$$\langle \hat{U}\phi | \hat{U}\psi \rangle = \langle \phi | \psi \rangle \text{ for any kets } |\phi\rangle, |\psi\rangle \quad (101)$$

We say that the scalar product is invariant under a unitary transformation.

It is easy to show that Eq. 101 holds for a unitary transformation:

$$\begin{aligned} \langle \hat{U}\phi | \hat{U}\psi \rangle &= \langle \phi | \hat{U}^\dagger \hat{U} \psi \rangle \quad (\text{used Property 1}) \\ &= \langle \phi | \hat{U}^{-1} \hat{U} \psi \rangle \quad (\text{used Property 4}) \\ &= \langle \phi | \psi \rangle \end{aligned}$$

Let us prove the reverse: if Eq. 101 is valid then  $\hat{U}$  is unitary. If  $|u_n\rangle$  are the pure states of  $\hat{U}$  then

$$\hat{U} = \sum_n |u_n\rangle u_n \langle u_n| \quad (102)$$

and

$$\hat{I} = \sum_n |u_n\rangle \langle u_n| \quad (103)$$

We start with Eq. 101:

$$\begin{aligned} \langle \phi | \psi \rangle &= \langle \hat{U}\phi | \hat{U}\psi \rangle \\ &= \langle \phi | \hat{U}^\dagger \hat{U} \psi \rangle \quad (\text{used Property 1}) \\ &= \langle \phi | \sum_m \sum_n |u_n\rangle u_n^* \langle u_n | u_m \rangle u_m \langle u_m | \psi \rangle \quad (\text{used Eq. 102}) \\ &= \langle \phi | \sum_m \sum_n |u_n\rangle u_n^* \delta_{nm} u_m \langle u_m | \psi \rangle \quad (\text{used } \langle u_n | u_m \rangle = \delta_{nm}) \\ &= \langle \phi | \sum_n |u_n\rangle u_n^* u_n \langle u_n | \psi \rangle \end{aligned}$$

This is equal to  $\langle \phi | \psi \rangle$  only if

$$\sum_n |u_n\rangle u_n^* u_n \langle u_n| = \hat{I}$$

Comparing this to Eq. 103 (completeness) implies that  $u_n^* u_n = 1$  for all  $n$ , which means that  $\hat{U}$  is unitary.

Consider two sets of kets  $\{|x_i\rangle\}_{i=1}^N$  and  $\{|y_i\rangle \equiv \hat{U}|x_i\rangle\}_{i=1}^N$  for some operator  $\hat{U}$ . If  $\hat{U}$  is unitary and one of the sets is complete and orthonormal then the other set is also complete and orthonormal. In addition, if both sets are complete and orthonormal then  $\hat{U}$  is unitary. This is a corollary of Property 13.