Chapter 9

Operations on States

We have seen in the preceding Chapter that the appropriate mathematical language for describing the states of a physical system is that of vectors belonging to a complex vector space. But the state space by itself is insufficient to fully describe the properties of a physical system. Describing such basic physical processes as how the state of a system evolves in time, or how to represent fundamental physical quantities such as position, momentum or energy, requires making use of further developments in the mathematics of vector spaces. These developments involve introducing a new mathematical entity known as an operator whose role it is to 'operate' on state vectors and map them into other state vectors. In fact, the earliest version of modern quantum mechanics, that put forward by Heisenberg, was formulated by him in terms of operators represented by matrices, at that time a not particularly well known (even to Heisenberg) development in pure mathematics. It was Born who recognized, and pointed out to Heisenberg, that he was using matrices in his work – another example of a purely mathematical construct that has proven to be of immediate value in describing the physical world.

Operators play many different roles in quantum mechanics. They can be used to represent physical processes that result in the change of state of the system, such as the evolution of the state of a system in time, or the creation or destruction of particles such as occurs, for instance in the emission or absorption of photons – particles of light – by matter. But operators have a further role, the one recognized by Heisenberg, which is to represent the physical properties of a system that can be, in principle, experimentally measured, such as energy, momentum, position and so on, so-called observable properties of a system. It is the aim of this Chapter to introduce the mathematical concept of an operator, and to show what the physical significance is of operators in quantum mechanics.

In general, in what follows, results explicitly making use of basis states will be presented only for the case in which these basis states are discrete, though in most cases, the same results hold true if a continuous set of basis states, as would arise for state spaces of infinite dimension, were used. The modifications that are needed when this is not the case, are considered towards the end of the Chapter.

9.1 Definition and Properties of Operators

9.1.1 Definition of an Operator

The need to introduce the mathematical notion of an operator into quantum mechanics can be motivated by the recognition of a fairly obvious characteristic of a physical system: it will evolve in time i.e. its state will be time dependent. But a change in state could also come about because some action is performed on the system, such as the system being displaced or rotated in space. The state of a multiparticle system can change as the consequence of particles making up a system being created or destroyed. Changes in state can also be brought about by processes which have a less physically direct meaning. Whatever the example, the fact that the state of a system in quantum mechanics is represented by a vector immediately suggests the possibility of describing a physical process by an exhaustive list of all the changes that the physical process induces on *every* state of the system, i.e. a table of all possible before and after states. A mathematical device can then be constructed which represents this list of before and after states. This mathematical device is, of course, known as an operator.

Thus the operator representing the above physical process is a mathematical object that acts on the state of a system and maps this state into some other state in accordance with the exhaustive list proposed above. If we represent an operator by a symbol $\hat{A}$ – note the presence of the ˆ – and suppose that the system is in a state $|\psi\rangle$, then the outcome of $\hat{A}$ acting on $|\psi\rangle$, written $\hat{A}|\psi\rangle$, (*not* $|\psi\rangle\hat{A}$, which is not a combination of symbols that has been assigned any meaning) defines a new state $|\phi\rangle$ say, so that

$$\hat{A}|\psi\rangle = |\phi\rangle. \tag{9.1}$$

As has already been mentioned, to fully characterize an operator, the effects of the operator acting on *every* possible state of the system must be specified. The extreme case, indicated above, requires, in effect, a complete tabulation of the each state of the system, and the corresponding result of the operator acting on each state. More typically this formidable (impossible?) task would be unneccessary – all that would be needed is some sort of rule that enables $|\phi\rangle$ to be determined for any given $|\psi\rangle$.

Ex 9.1 Consider the operator $\hat{A}$ acting on the states of a spin half system, and suppose, for the arbitrary state $|S\rangle = a|+\rangle + b|-\rangle$, that the action of the operator $\hat{A}$ is such that $\hat{A}|S\rangle = b|+\rangle + a|-\rangle$, i.e. the action of the operator $\hat{A}$ on the state $|S\rangle$ is to exchange the coefficients $\langle\pm|S\rangle \leftrightarrow \langle\mp|S\rangle$. This rule is then enough to define the result of $\hat{A}$ acting on any state of the system, i.e. the operator is fully specified.

Ex 9.2 A slightly more complicated example is one for which the action of an operator $\hat{N}$ on the state $|S\rangle$ is given by $\hat{N}|S\rangle = a^2|+\rangle + b^2|-\rangle$. As we shall see below, the latter operator is of a kind not usually encountered in quantum mechanics, that is, it is non-linear, see Section 9.1.2 below.

While the above discussion provides motivation on physical grounds for introducing the idea of an operator as something that acts to change the state of a quantum system, it is in fact the case that many operators that arise in quantum mechanics, whilst mathematically they can act on a state vector to map it into another state vector, do not represent a physical process acting on the system. In fact, the operators that represent such physical processes as the evolution of the system in time, are but one kind of operator important in quantum mechanics known as unitary operators. Another very important kind of operator is that which represents the physically observable properties of a system, such as monemtum or energy. Each such observable property, or *observable*, is represented by a particular kind of operator known as a Hermitean operator. Mathematically, such an operator can be made to act on the state of a system, thereby yielding a new state, but the interpretation of this as representing an actual physical process is much less direct. Instead, a Hermitean operator acts, in a sense, as a repository of all the possible results that can be obtained when performing a measurement of the physical observable that the operator represents. Curiously enough, it nevertheless turns out that Hermitean operators representing observables of a system, and unitary operators representing possible actions performed on a system are very closely related in a way that will be examined in Chapter **??**.

In quantum mechanics, the task of fully characterizing an operator is actually made much simpler through the fact that most operators in quantum mechanics have a very important property: they are linear, or, at worst, anti-linear.

9.1.2 Linear and Antilinear Operators

There is essentially no limit to the way in which operators could be defined, but of particular importance are operators that have the following property. If $\hat{A}$ is an operator such that for any arbitrary pair of states $|\psi_1\rangle$ and $|\psi_2\rangle$ and for any complex numbers c_1 and c_2:

$$\hat{A}[c_1|\psi_1\rangle + c_2|\psi_2\rangle] = c_1\hat{A}|\psi_1\rangle + c_2\hat{A}|\psi_2\rangle, \tag{9.2}$$

then $\hat{A}$ is said to be a *linear* operator. In quantum mechanics, operators are, with one exception, linear. The exception is the time reversal operator $\hat{T}$ which has the property

$$\hat{T}[c_1|\psi_1\rangle + c_2|\psi_2\rangle] = c_1^*\hat{T}|\psi_1\rangle + c_2^*\hat{T}|\psi_2\rangle \tag{9.3}$$

and is said to be *anti-linear*. We will not have any need to concern ourselves with the time reversal operator, so any operator that we will be encountering here will be tacitly assumed to be linear.

Ex 9.3 Consider the operator $\hat{A}$ acting on the states of a spin half system, such that for the arbitrary state $|S\rangle = a|+\rangle + b|-\rangle$, $\hat{A}|S\rangle = b|+\rangle + a|-\rangle$. Show that this operator is linear. Introduce another state $|S'\rangle = a'|+\rangle + b'|-\rangle$ and consider

$$\begin{aligned}\hat{A}[\alpha|S\rangle + \beta|S'\rangle] &= \hat{A}[(\alpha a + \beta a')|+\rangle + (\alpha b + \beta b')|-\rangle]\\ &= (\alpha b + \beta b')|+\rangle + (\alpha a + \beta a')|-\rangle\\ &= \alpha(b|+\rangle + a|-\rangle) + \beta(b'|+\rangle + a'|-\rangle)\\ &= \alpha\hat{A}|S\rangle + \beta\hat{A}|S'\rangle.\end{aligned}$$

and hence the operator $\hat{A}$ is linear.

Ex 9.4 Consider the operator $\hat{N}$ defined such that if $|S\rangle = a|+\rangle + b|-\rangle$ then $\hat{N}|S\rangle = a^2|+\rangle + b^2|-\rangle$. Show that this operator is non-linear. If we have another state $|S'\rangle = a'|+\rangle + b'|-\rangle$, then

$$\hat{N}[|S\rangle + |S'\rangle] = \hat{N}[(a+a')|+\rangle + (b+b')|-\rangle] = (a+a')^2|+\rangle + (b+b')^2|-\rangle.$$

But

$$\hat{N}|S\rangle + \hat{N}|S'\rangle = (a^2 + a'^2)|+\rangle + (b^2 + b'^2)|-\rangle$$

which is certainly not equal to $\hat{N}[|S\rangle + |S'\rangle]$. Thus the operator $\hat{N}$ is non-linear.

The importance of linearity lies in the fact that since any state vector $|\psi\rangle$ can be written as a linear combination of a complete set of basis states, $\{|\varphi_n\rangle, n = 1, 2, \ldots\}$:

$$|\psi\rangle = \sum_n |\varphi_n\rangle\langle\varphi_n|\psi\rangle$$

then

$$\hat{A}|\psi\rangle = \hat{A}\sum_n |\varphi_n\rangle\langle\varphi_n|\psi\rangle = \sum_n \hat{A}|\varphi_n\rangle\langle\varphi_n|\psi\rangle \tag{9.4}$$

so that provided we know what an operator $\hat{A}$ does to each basis state, we can determine what $\hat{A}$ does to any vector belonging to the state space.

Ex 9.5 Consider the spin states $|+\rangle$ and $|-\rangle$, basis states for a spin half system, and suppose an operator $\hat{A}$ has the properties

$$\hat{A}|+\rangle = \tfrac{1}{2}i\hbar|-\rangle$$
$$\hat{A}|-\rangle = -\tfrac{1}{2}i\hbar|+\rangle.$$

Then if a spin half system is in the state

$$|S\rangle = \frac{1}{\sqrt{2}}[|+\rangle + |-\rangle]$$

then

$$\begin{aligned}\hat{A}|S\rangle &= \frac{1}{\sqrt{2}}\hat{A}|+\rangle + \frac{1}{\sqrt{2}}\hat{A}|-\rangle \\ &= \tfrac{1}{2}\hbar\frac{i}{\sqrt{2}}[|-\rangle - |+\rangle] \\ &= -\tfrac{1}{2}i\hbar\frac{1}{\sqrt{2}}[|+\rangle - |-\rangle].\end{aligned}$$

So the state vector $|S\rangle = \frac{1}{\sqrt{2}}[|+\rangle + |-\rangle]$ is mapped into the state vector $-\frac{1}{2}i\hbar\frac{1}{\sqrt{2}}[|+\rangle - |-\rangle]$, which represents a different physical state, and one which, incidentally, is not normalized to unity.

Ex 9.6 Suppose an operator $\hat{B}$ is defined so that

$$\hat{B}|+\rangle = \tfrac{1}{2}\hbar|-\rangle$$
$$\hat{B}|-\rangle = \tfrac{1}{2}\hbar|+\rangle.$$

If we let $\hat{B}$ act on the state $|S\rangle = [|+\rangle + |-\rangle]/\sqrt{2}$ then we find that

$$\hat{B}|S\rangle = \tfrac{1}{2}\hbar|S\rangle \tag{9.5}$$

i.e. in this case, we regain the same state vector $|S\rangle$, though multiplied by a factor $\frac{1}{2}\hbar$. This last equation is an example of an eigenvalue equation: $|S\rangle$ is said to be an eigenvector of the operator $\hat{B}$, and $\frac{1}{2}\hbar$ is its eigenvalue. The concept of an eigenvalue and eigenvector is very important in quantum mechanics, and much more will be said about it later.

In the following Sections we look in more detail some of the more important properties of operators in quantum mechanics.

9.1.3 Properties of Operators

In this discussion, a general perspective is adopted, but the properties will be encountered again and in a more concrete fashion when we look at representations of operators by matrices.

Equality of Operators

If two operators, $\hat{A}$ and $\hat{B}$ say, are such that

$$\hat{A}|\psi\rangle = \hat{B}|\psi\rangle \tag{9.6}$$

for all state vectors $|\psi\rangle$ belonging to the state space of the system then the two operators are said to be equal, written

$$\hat{A} = \hat{B}. \tag{9.7}$$

Linearity also makes it possible to set up a direct way of proving the equality of two operators. Above it was stated that two operators, $\hat{A}$ and $\hat{B}$ say, will be equal if $\hat{A}|\psi\rangle = \hat{B}|\psi\rangle$ for all states $|\psi\rangle$. However, it is sufficient to note that if for all the basis vectors $\{|\varphi_n\rangle, n = 1, 2, \dots\}$

$$\hat{A}|\varphi_n\rangle = \hat{B}|\varphi_n\rangle \tag{9.8}$$

then we immediately have, for any arbitrary state $|\psi\rangle$ that

$$\begin{aligned}\hat{A}|\psi\rangle &= \hat{A}\sum_n |\varphi_n\rangle\langle\varphi_n|\psi\rangle \\ &= \sum_n \hat{A}|\varphi_n\rangle\langle\varphi_n|\psi\rangle \\ &= \sum_n \hat{B}|\varphi_n\rangle\langle\varphi_n|\psi\rangle \\ &= \hat{B}\sum_n |\varphi_n\rangle\langle\varphi_n|\psi\rangle \\ &= \hat{B}|\psi\rangle\end{aligned} \tag{9.9}$$

so that $\hat{A} = \hat{B}$. Thus, to prove the equality of two operators, it is sufficient to show that the action of the operators on each member of a basis set gives the same result.

The Unit Operator and the Zero Operator

Of all the operators that can be defined, there are two whose properties are particularly simple – the unit operator $\hat{1}$ and the zero operator $\hat{0}$. The unit operator is the operator such that

$$\hat{1}|\psi\rangle = |\psi\rangle \tag{9.10}$$

for all states $|\psi\rangle$, and the zero operator is such that

$$\hat{0}|\psi\rangle = 0 \tag{9.11}$$

for all kets $|\psi\rangle$.

Addition of Operators

The sum of two operators $\hat{A}$ and $\hat{B}$, written $\hat{A} + \hat{B}$ is defined in the obvious way, that is

$$(\hat{A} + \hat{B})|\psi\rangle = \hat{A}|\psi\rangle + \hat{B}|\psi\rangle \tag{9.12}$$

for all vectors $|\psi\rangle$. The sum of two operators is, of course, another operator, $\hat{S}$ say, written $\hat{S} = \hat{A} + \hat{B}$, such that

$$\hat{S}|\psi\rangle = (\hat{A} + \hat{B})|\psi\rangle = \hat{A}|\psi\rangle + \hat{B}|\psi\rangle \tag{9.13}$$

for all states $|\psi\rangle$.

Ex 9.7 Consider the two operators $\hat{A}$ and $\hat{B}$ defined by

$$\begin{aligned} \hat{A}|+\rangle &= \tfrac{1}{2}i\hbar|-\rangle & \hat{B}|+\rangle &= \tfrac{1}{2}\hbar|-\rangle \\ \hat{A}|-\rangle &= -\tfrac{1}{2}i\hbar|+\rangle & \hat{B}|-\rangle &= \tfrac{1}{2}\hbar|+\rangle. \end{aligned} \tag{9.14}$$

Their sum $\hat{S}$ will then be such that

$$\begin{aligned} \hat{S}|+\rangle &= \tfrac{1}{2}(1 + i)\hbar|-\rangle \\ \hat{S}|-\rangle &= \tfrac{1}{2}(1 - i)\hbar|+\rangle. \end{aligned} \tag{9.15}$$

Multiplication of an Operator by a Complex Number

This too is defined in the obvious way. Thus, if $\hat{A}|\psi\rangle = |\phi\rangle$ then we can define the operator $\lambda\hat{A}$ where λ is a complex number to be such that

$$(\lambda\hat{A})|\psi\rangle = \lambda(\hat{A}|\psi\rangle) = \lambda|\phi\rangle. \tag{9.16}$$

Combining this with the previous definition of the sum of two operators, we can then make say that in general

$$(\lambda\hat{A} + \mu\hat{B})|\psi\rangle = \lambda(\hat{A}|\psi\rangle) + \mu(\hat{B}|\psi\rangle) \tag{9.17}$$

where λ and μ are both complex numbers.

Multiplication of Operators

Given that an operator $\hat{A}$ say, acting on a ket vector $|\psi\rangle$ maps it into another ket vector $|\phi\rangle$, then it is possible to allow a second operator, $\hat{B}$ say, to act on $|\phi\rangle$, producing yet another ket vector $|\xi\rangle$ say. This we can write as

$$\hat{B}\{\hat{A}|\psi\rangle\} = \hat{B}|\phi\rangle = |\xi\rangle. \tag{9.18}$$

This can be written

$$\hat{B}\{\hat{A}|\psi\rangle\} = \hat{B}\hat{A}|\psi\rangle \tag{9.19}$$

i.e. without the braces $\{\ldots\}$, with the understanding that the term on the right hand side is to be interpreted as meaning that first $\hat{A}$ acts on the state to its right, and then $\hat{B}$, in the sense specified in Eq. (9.18). The combination $\hat{B}\hat{A}$ is said to be the product of the two operators $\hat{A}$ and $\hat{B}$. The product of two operators is, of course, another operator. Thus we can write $\hat{C} = \hat{B}\hat{A}$ where the operator $\hat{C}$ is such that

$$\hat{C}|\psi\rangle = \hat{B}\hat{A}|\psi\rangle \tag{9.20}$$

for all states $|\psi\rangle$.

Ex 9.8 Consider the products of the two operators defined in Eq. (9.14). First $\hat{C} = \hat{B}\hat{A}$:

$$\begin{aligned}\hat{C}|+\rangle &= \hat{B}\hat{A}|+\rangle = \hat{B}(\tfrac{1}{2}i\hbar|-\rangle) = \tfrac{1}{4}i\hbar^2|+\rangle\\ \hat{C}|-\rangle &= \hat{B}\hat{A}|-\rangle = \hat{B}(-\tfrac{1}{2}i\hbar|+\rangle) = -\tfrac{1}{4}i\hbar^2|-\rangle,\end{aligned} \tag{9.21}$$

and next $\hat{D} = \hat{A}\hat{B}$:

$$\begin{aligned}\hat{D}|+\rangle &= \hat{A}\hat{B}|+\rangle = \hat{A}(\tfrac{1}{2}\hbar|-\rangle) = -\tfrac{1}{4}i\hbar^2|+\rangle\\ \hat{D}|-\rangle &= \hat{A}\hat{B}|-\rangle = \hat{A}(-\tfrac{1}{2}\hbar|+\rangle) = \tfrac{1}{4}i\hbar^2|-\rangle.\end{aligned} \tag{9.22}$$

Apart from illustrating how to implement the definition of the product of two operators, this example also shows a further important result, namely that, in general, $\hat{A}\hat{B} \neq \hat{B}\hat{A}$. In other words, the order in which two operators are multiplied is important. The difference between the two, written

$$\hat{A}\hat{B} - \hat{B}\hat{A} = [\hat{A}, \hat{B}] \tag{9.23}$$

is known as the commutator of $\hat{A}$ and $\hat{B}$. If the commutator vanishes, the operators are said to commute. The commutator plays a fundamental role in the physical interpretation of quantum mechanics, being both a bridge between the classical description of a physical system and its quantum description, and important in describing the consequences of sequences of measurements performed on a quantum system.

Projection Operators

An operator $\hat{P}$ that has the property

$$\hat{P}^2 = \hat{P} \tag{9.24}$$

is said to be a *projection operator*. An important example of a projection operator is the operator $\hat{P}_n$ defined, for a given set of orthonormal basis states $\{|\varphi_n\rangle; n = 1, 2, 3 \ldots\}$ by

$$\hat{P}_n|\varphi_m\rangle = \delta_{nm}|\varphi_n\rangle. \tag{9.25}$$

That this operator is a projection operator can be readily confirmed:

$$\hat{P}_n^2|\varphi_m\rangle = \hat{P}_n\{\hat{P}_n|\phi_m\rangle\} = \delta_{nm}\hat{P}_n|\varphi_m\rangle = \delta_{nm}^2|\phi_m\rangle. \tag{9.26}$$

But since $\delta_{nm}^2 = \delta_{nm}$, (recall that the Kronecker delta δ_{nm} is either unity for $n = m$ or zero for $n \neq m$) we immediately have that

$$\hat{P}_n^2|\varphi_m\rangle = \delta_{nm}|\varphi_m\rangle = \hat{P}_n|\varphi_m\rangle \tag{9.27}$$

from which we conclude that $\hat{P}_n^2 = \hat{P}_n$. The importance of this operator lies in the fact that if we let it act on an arbitrary vector $|\psi\rangle$, then we see that

$$\hat{P}_n|\psi\rangle = \hat{P}_n \sum_m |\varphi_m\rangle\langle\varphi_m|\psi\rangle = \sum_m \hat{P}_n|\varphi_m\rangle\langle\varphi_m|\psi\rangle = \sum_m \delta_{nm}|\varphi_m\rangle\langle\varphi_m|\psi\rangle = |\varphi_n\rangle\langle\varphi_n|\psi\rangle \tag{9.28}$$

i.e. it 'projects' out the component of $|\psi\rangle$ in the direction of the basis state $|\varphi_n\rangle$.

Ex 9.9 Consider a spin half particle in the state $|\psi\rangle = a|+\rangle + b|-\rangle$ where a and b are both real numbers, and define the operators $\hat{P}_\pm$ such that $\hat{P}_-|-\rangle = |-\rangle$, $\hat{P}_-|+\rangle = 0$, $P_+|-\rangle = 0$, and $\hat{P}_+|+\rangle = |+\rangle$. Show that $\hat{P}_\pm$ are projection operators, and evaluate $\hat{P}_\pm|\psi\rangle$.

To show that the $\hat{P}_\pm$ are projection operators, it is necessary to show that

$$\hat{P}_\pm^2 = \hat{P}_\pm.$$

It is sufficient to let $\hat{P}_-^2$ to act on the basis states $|\pm\rangle$:

$$\hat{P}_-^2|-\rangle = \hat{P}_-\{\hat{P}_-|-\rangle\} = \hat{P}_-|-\rangle = |-\rangle$$

and

$$\hat{P}_-^2|+\rangle = \hat{P}_-\{\hat{P}_-|+\rangle\} = 0 = \hat{P}_-|+\rangle.$$

Thus we have shown that $\hat{P}_-^2|\pm\rangle = \hat{P}_-|\pm\rangle$, so that, since the states $|\pm\rangle$ are a pair of basis states, we can conclude that $\hat{P}_-^2 = \hat{P}_-$, so that $\hat{P}_-$ is a projection operator. A similar argument can be followed through for $\hat{P}_+$.

By the properties of $\hat{P}_-$ it follows that

$$\hat{P}_-|\psi\rangle = \hat{P}_-\,[a|+\rangle + b|-\rangle] = b|-\rangle.$$

and similarly

$$\hat{P}_+|\psi\rangle = \hat{P}_+\,[a|+\rangle + b|-\rangle] = a|+\rangle.$$

This result is illustrated in Fig. (9.1) where the projections of $|\psi\rangle$ on to the $|+\rangle$ and $|-\rangle$ basis states are depicted.

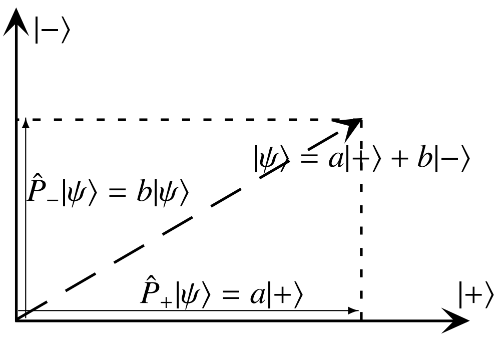

Figure 9.1: An illustration of the action of projection operators on a state $|\psi\rangle = a|+\rangle + b|-\rangle$ of a spin half system where a and b are real.

Functions of Operators

Having defined what is meant by adding and multiplying operators, we can now define the idea of a function of an operator. If we have a function $f(x)$ which we can expand as a power series in x:

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots = \sum_{n=0}^{\infty} a_n x^n \tag{9.29}$$

then we define $f(\hat{A})$, a function of the operator $\hat{A}$, to be also given by the same power series, i.e.

$$f(\hat{A}) = a_0 + a_1 \hat{A} + a_2 \hat{A}^2 + \cdots = \sum_{n=0}^{\infty} a_n \hat{A}^n. \tag{9.30}$$

Questions such as the convergence of such a series (if it is an infinite series) will not be addressed here.

Ex 9.10 The most important example of a function of an operator that we will have to deal with here is the exponential function:

$$f(\hat{A}) = e^{\hat{A}} = 1 + \hat{A} + \frac{1}{2!}\hat{A}^2 + \dots. \tag{9.31}$$

Many important operators encountered in quantum mechanics, in particular the time evolution operator which specifies how the state of a system evolves in time, is given as an exponential function of an operator.

It is important to note that in general, the usual rules for manipulating exponential functions do not apply for exponentiated operators. In particular, it should be noted that in general

$$e^{\hat{A}} e^{\hat{B}} \neq e^{\hat{A}+\hat{B}} \tag{9.32}$$

unless $\hat{A}$ commutes with $\hat{B}$.

The Inverse of an Operator

If, for some operator $\hat{A}$ there exists another operator $\hat{B}$ with the property that

$$\hat{A}\hat{B} = \hat{B}\hat{A} = \hat{1} \tag{9.33}$$

then $\hat{B}$ is said to be the inverse of $\hat{A}$ and is written

$$\hat{B} = \hat{A}^{-1}. \tag{9.34}$$

Ex 9.11 An important example is the inverse of the operator $\exp(\hat{A})$ defined by the power series in Eq. (9.31) above. The inverse of this operator is readily seen to be just $\exp(-\hat{A})$.

Ex 9.12 Another useful result is the inverse of the product of two operators i.e.$(\hat{A}\hat{B})^{-1}$, that is

$$(\hat{A}\hat{B})^{-1} = \hat{B}^{-1}\hat{A}^{-1}. \tag{9.35}$$

provided, of course, that both $\hat{A}$ and $\hat{B}$ have inverses. This result can be readily shown by noting that, by the definition of the inverse of an operator

$$(\hat{A}\hat{B})^{-1}(\hat{A}\hat{B}) = \hat{1}. \tag{9.36}$$

Multiplying this on the right, first by $\hat{B}^{-1}$, and then by $\hat{A}^{-1}$ then gives the required result.

9.2 Action of Operators on Bra Vectors

Given that an operator maps a ket vector into another ket, as summarized in the defining equation $\hat{A}|\psi\rangle = |\phi\rangle$, we can then take the inner product of $|\phi\rangle$ with any other state vector $|\xi\rangle$ say to yield the complex number $\langle\xi|\phi\rangle$. This we can obviously also write as

$$\langle\xi|\phi\rangle = \langle\xi|(\hat{A}|\psi\rangle). \tag{9.37}$$

This then raises the interesting question, since a bra vector is juxtaposed with an operator in Eq. (9.37), whether we could give a meaning to an operator acting on a bra vector. In other words, we can give a meaning to $\langle\xi|\hat{A}$? Presumably, the outcome of $\hat{A}$ acting on a bra vector is to produce another bra vector, i.e. we can write $\langle\xi|\hat{A} = \langle\chi|$, though as yet we have not specified how to determine what the bra vector $\langle\chi|$ might be. But since operators were originally defined above in terms of their action on ket vectors, it makes sense to define the action of an operator on a bra in a way that makes use of what we know about the action of an operator on any ket vector. So, we define $\langle\xi|\hat{A}$ such that

$$(\langle\xi|\hat{A})|\psi\rangle = \langle\xi|(\hat{A}|\psi\rangle) \quad \text{for all ket vectors } |\psi\rangle. \tag{9.38}$$

The value of this definition, apart from the fact that it relates the action of operators on bra vectors back to the action of operators of operators on ket vectors, is that $\langle\xi|(\hat{A}|\psi\rangle)$ will always give the same result as $\langle\xi|(\hat{A}|\psi\rangle)$ i.e. it is immaterial whether we let $\hat{A}$ act on the ket vector first, and then take the inner product with $|\xi\rangle$, or to let $\hat{A}$ act on $\langle\xi|$ first, and then take the inner product with $|\psi\rangle$. Thus the brackets are not needed, and we can write:

$$(\langle\xi|\hat{A})|\psi\rangle = \langle\xi|(\hat{A}|\psi\rangle) = \langle\xi|\hat{A}|\psi\rangle. \tag{9.39}$$

This way of defining the action of an operator on a bra vector, Eq. (9.39), is rather back-handed, so it is important to see that it does in fact do the job! To see that the definition actually works, we will look at the particular case of the spin half state space again. Suppose we have an operator $\hat{A}$ defined such that

$$\begin{aligned} \hat{A}|+\rangle &= |+\rangle + i|-\rangle \\ \hat{A}|-\rangle &= i|+\rangle + |-\rangle \end{aligned} \tag{9.40}$$

and we want to determine $\langle +|\hat{A}$ using the above definition. Let $\langle\chi|$ be the bra vector we are after, i.e. $\langle +|\hat{A} = \langle\chi|$. We know that we can always write

$$\langle\chi| = \langle\chi|+\rangle\langle+| + \langle\chi|-\rangle\langle-| \tag{9.41}$$

so the problem becomes evaluating $\langle\chi|\pm\rangle$. It is at this point that we make use of the defining condition above. Thus, we write

$$\langle\chi|\pm\rangle = \{\langle+|\hat{A}\}|\pm\rangle = \langle+|\{\hat{A}|\pm\rangle\}. \tag{9.42}$$

Using Eq. (9.40) this gives

$$\langle\chi|+\rangle = \langle+|\{\hat{A}|+\rangle\} = 1 \quad \text{and} \quad \langle\chi|-\rangle = \langle+|\{\hat{A}|-\rangle\} = i \tag{9.43}$$

and hence

$$\langle\chi| = \langle+| + i\langle-|. \tag{9.44}$$

Consequently, we conclude that

$$\langle+|\hat{A} = \langle+| + i\langle-|. \tag{9.45}$$

If we note that $\hat{A}|+\rangle = |+\rangle + i|-\rangle$ we can see that $\langle+|\hat{A} \neq \langle+| - i\langle-|$. This example illustrates the result that if $\hat{A}|\psi\rangle = |\phi\rangle$ then, in general, $\langle\psi|\hat{A} \neq \langle\phi|$. This example shows that the above 'indirect' definition of the action of an operator on a bra vector in terms of the action of the operator on ket vectors does indeed give us the result of the operator acting on a bra vector. The general method used in this example can be extended to the general case. So suppose we have a state space for some system spanned by a complete orthonormal set of basis states $\{|\varphi_n\rangle; n = 1, 2, \dots\}$, and assume that we know the action of an operator $\hat{A}$ on an arbitrary basis state $|\varphi_n\rangle$:

$$\hat{A}|\varphi_n\rangle = \sum_m |\varphi_m\rangle A_{mn} \tag{9.46}$$

where the A_{mn} are complex numbers. This equation is analogous to Eq. (9.40) in the example above. Now suppose we allow $\hat{A}$ to act on an arbitrary bra vector $\langle\xi|$:

$$\langle\xi|\hat{A} = \langle\chi| \tag{9.47}$$

We can express $\langle\chi|$ in terms of the basis states introduced above:

$$\langle\chi| = \sum_n \langle\chi|\varphi_n\rangle\langle\varphi_n|. \tag{9.48}$$

Thus, the problem reduces to showing that we can indeed calculate the coefficients $\langle\chi|\varphi_n\rangle$. These coefficients are given by

$$\langle\chi|\varphi_n\rangle = \{\langle\xi|\hat{A}\}|\varphi_n\rangle = \langle\xi|\{\hat{A}|\varphi_n\rangle\} \tag{9.49}$$

where we have used the defining condition Eq. (9.38) to allow $\hat{A}$ to act on the basis state $|\varphi_n\rangle$. Using Eq. (9.46) we can write

$$\begin{aligned}\langle\chi|\varphi_n\rangle &= \langle\xi|\{\hat{A}|\varphi_n\rangle\}\\ &= \langle\xi|\Big[\sum_m |\varphi_m\rangle A_{mn}\Big]\\ &= \sum_m A_{mn}\langle\xi|\varphi_m\rangle.\end{aligned} \tag{9.50}$$

If we substitute this into the expression Eq. (9.48) we find that

$$\langle\chi| = \langle\xi|\hat{A} = \sum_n \Big[\sum_m \langle\xi|\varphi_m\rangle A_{mn}\Big]\langle\varphi_n|. \tag{9.51}$$

The quantity within the brackets is a complex number which we can always evaluate since we know the A_{mn} and can evaluate the inner product $\langle\xi|\varphi_m\rangle$. Thus, by use of the defining condition Eq. (9.38), we are able to calculate the result of an operator acting on a bra vector. Of particular interest is the case in which $\langle\xi| = \langle\varphi_k|$ for which

$$\langle\varphi_k|\hat{A} = \sum_n \Big[\sum_m \langle\varphi_k|\varphi_m\rangle A_{mn}\Big]\langle\varphi_n|. \tag{9.52}$$

Since the basis states are orthonormal, i.e. $\langle\varphi_k|\varphi_m\rangle = \delta_{km}$, then

$$\begin{aligned}\langle\varphi_k|\hat{A} &= \sum_n \Big[\sum_m \delta_{km} A_{mn}\Big]\langle\varphi_n| \\ &= \sum_n A_{kn}\langle\varphi_n|.\end{aligned} \tag{9.53}$$

It is useful to compare this result with Eq. (9.46):

$$\begin{aligned}\hat{A}|\varphi_n\rangle &= \sum_m |\varphi_m\rangle A_{mn} \\ \langle\varphi_n|\hat{A} &= \sum_m A_{nm}\langle\varphi_m|.\end{aligned} \tag{9.54}$$

Either of these expressions lead to the result

$$A_{mn} = \langle\varphi_m|\hat{A}|\varphi_n\rangle. \tag{9.55}$$

For reasons which will become clearer later, the quantities A_{mn} are known as the matrix elements of the operator $\hat{A}$ with respect to the set of basis states $\{|\varphi_n\rangle; n = 1, 2, \ldots\}$.

Ex 9.13 With respect to a pair of orthonormal vectors $|\varphi_1\rangle$ and $|\varphi_2\rangle$ that span the Hilbert space $\mathcal{H}$ of a certain system, the operator $\hat{A}$ is defined by its action on these base states as follows:

$$\begin{aligned}\hat{A}|\varphi_1\rangle &= 3|\varphi_1\rangle - 4i|\varphi_2\rangle \\ \hat{A}|\varphi_2\rangle &= -4i|\varphi_1\rangle - 3|\varphi_2\rangle.\end{aligned}$$

Evaluate $\langle\varphi_1|\hat{A}$ and $\langle\varphi_2|\hat{A}$.

We proceed by considering the product $\{\langle\varphi_1|\hat{A}\}|\psi\rangle$ where $|\psi\rangle$ is an arbitrary state vector which we can expand with respect to the pair of basis states $\{|\varphi_1\rangle, |\varphi_2\rangle\}$ as $|\psi\rangle = a|\varphi_1\rangle + b|\varphi_2\rangle$.

Using the defining condition Eq. (9.38) we have

$$\{\langle\varphi_1|\hat{A}\}|\psi\rangle = \langle\varphi_1|\{\hat{A}|\psi\rangle\} = \langle\varphi_1|\{a\hat{A}|\varphi_1\rangle + b\hat{A}|\varphi_2\rangle\}$$

Using the properties of $\hat{A}$ as given, and using the fact that $a = \langle\varphi_1|\psi\rangle$ and $b = \langle\varphi_2|\psi\rangle$ we get

$$\{\langle\varphi_1|\hat{A}\}|\psi\rangle = 3a - 4ib = 3\langle\varphi_1|\psi\rangle - 4i\langle\varphi_2|\psi\rangle.$$

Extracting the common factor $|\psi\rangle$ yields

$$\{\langle\varphi_1|\hat{A}\}|\psi\rangle = \{3\langle\varphi_1| - 4i\langle\varphi_2|\}|\psi\rangle$$

from which we conclude, since $|\psi\rangle$ is arbitrary, that

$$\langle\varphi_1|\hat{A} = 3\langle\varphi_1| - 4i\langle\varphi_2|.$$

In a similar way, we find that

$$\langle\varphi_2|\hat{A} = -4i\langle\varphi_1| - 3\langle\varphi_2|.$$

Ex 9.14 A useful an important physical system with which to illustrate some of these ideas is that of the electromagnetic field inside a cavity designed to support a single mode of the field, see p95. In this case, the basis states of the electromagnetic field are the so-called number states $\{|n\rangle, n = 0, 1, 2, \ldots\}$ where the state $|n\rangle$ is the state of the field in which there are n photons present. We can now introduce an operator $\hat{a}$ defined such that

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle, \qquad \hat{a}|0\rangle = 0. \tag{9.56}$$

Understandably, this operator is known as the photon annihilation operator as it transforms a state with n photons into one in which there are $n-1$ photons. The prefactor $\sqrt{n}$ is there for later purposes i.e. it is possible to define an operator $\hat{a}'$ such that $\hat{a}'|n\rangle = |n-1\rangle$, but this operator turns out not to arise in as natural a way as $\hat{a}$, and is not as useful in practice. We can ask the question: what is $\langle n|\hat{a}$?

To address this, we must make use of the manner in which we define the action of an operator on a bra, that is, it must be such that

$$\{\langle n|\hat{a}\}|\psi\rangle = \langle n|\{\hat{a}|\psi\rangle\}$$

holds true for all states $|\psi\rangle$.

If we expand $|\psi\rangle$ in terms of the basis states $\{|n\rangle, n = 0, 1, 2, \ldots\}$ we have

$$|\psi\rangle = \sum_{m=0}^{\infty} |m\rangle\langle m|\psi\rangle$$

where we have used a summation index (m) to distinguish it from the label n

used on the bra vector $\langle n|$. From this we have

$$\begin{aligned}
\{\langle n|\hat{a}\}|\psi\rangle &= \{\langle n|\hat{a}\}\sum_{m=0}^{\infty}|m\rangle\langle m|\psi\rangle \\
&= \sum_{m=0}^{\infty}\{\langle n|\hat{a}\}|m\rangle\langle m|\psi\rangle \\
&= \sum_{m=0}^{\infty}\langle n|\{\hat{a}|m\rangle\langle m|\psi\rangle\} \\
&= \sum_{m=1}^{\infty}\sqrt{m}\langle n|m-1\rangle\langle m|\psi\rangle\}
\end{aligned}$$

where the sum now begins at $m = 1$ as the $m = 0$ term will vanish. This sum can be written as

$$\begin{aligned}
\{\langle n|\hat{a}\}|\psi\rangle &= \sum_{m=0}^{\infty}\sqrt{m+1}\langle n|m\rangle\langle m+1|\psi\rangle\} \\
&= \sum_{m=0}^{\infty}\sqrt{m+1}\delta_{nm}\langle m+1|\psi\rangle \\
&= \left\{\sqrt{n+1}\langle n+1|\right\}|\psi\rangle
\end{aligned}$$

By comparing the left and right hand sides of this last equation, and recognizing that $|\psi\rangle$ is arbitrary, we conclude that

$$\langle n|\hat{a} = \sqrt{n+1}\langle n+1| \tag{9.57}$$

A further consequence of the above definition of the action of operators on bra vectors, which is actually implicit in the derivation of the result Eq. (9.51) is the fact that an operator $\hat{A}$ that is linear with respect to ket vectors, is also linear with respect to bra vectors i.e.

$$[\lambda\langle\psi_1| + \mu\langle\psi_2|]\hat{A} = \lambda\langle\psi_1|\hat{A} + \mu\langle\psi_2|\hat{A} \tag{9.58}$$

which further emphasizes the symmetry between the action of operators on bras and kets.

Much of what has been presented above is recast in terms of matrices and column and row vectors in a later Section.

9.3 The Hermitean Adjoint of an Operator

We have seen above that if $\hat{A}|\psi\rangle = |\phi\rangle$ then $\langle\psi|\hat{A} \neq \langle\phi|$. This then suggests the possibility of introducing an operator related to $\hat{A}$, which we will write $\hat{A}^\dagger$ which is such that

$$\text{if}\quad \hat{A}|\psi\rangle = |\phi\rangle \quad\text{then}\quad \langle\psi|\hat{A}^\dagger = \langle\phi|. \tag{9.59}$$

The operator $\hat{A}^\dagger$ so defined is known as the *Hermitean adjoint* of $\hat{A}$. There are issues concerning the definition of the Hermitean adjoint that require careful consideration if the state space is of infinite dimension. We will not be concerning ourselves with these matters here. Thus we see we have introduced a new operator which has been defined in terms of its actions on bra vectors. In keeping with our point of view that operators should be defined in terms of their action on ket vectors, it should be the case that this above definition should unambiguously tell us what the action of $\hat{A}^\dagger$ will be on any ket vector. In other words, the task at hand is to show that we can evaluate $\hat{A}^\dagger|\psi\rangle$ for any arbitrary ket vector $|\psi\rangle$. In order to do this, we need a useful property of the Hermitean adjoint which can be readily derived from the above definition. Thus, consider $\langle\xi|\hat{A}|\psi\rangle$, which we recognize is simply a complex number given by

$$\langle\xi|\hat{A}|\psi\rangle = \langle\xi|(\hat{A}|\psi\rangle) = \langle\xi|\phi\rangle \tag{9.60}$$

where $\hat{A}|\psi\rangle = |\phi\rangle$. Thus, if we take the complex conjugate. we have

$$\langle\xi|\hat{A}|\psi\rangle^* = \langle\xi|\phi\rangle^* = \langle\phi|\xi\rangle. \tag{9.61}$$

But, since $\hat{A}|\psi\rangle = |\phi\rangle$ then $\langle\psi|\hat{A}^\dagger = \langle\phi|$ so we have

$$\langle\xi|\hat{A}|\psi\rangle^* = \langle\phi|\xi\rangle = (\langle\psi|\hat{A}^\dagger)|\xi\rangle = \langle\psi|\hat{A}^\dagger|\xi\rangle \tag{9.62}$$

where in the last step the brackets have been dropped since it does not matter whether an operator acts on the ket or the bra vector. Thus, taking the complex conjugate of $\langle\xi|\hat{A}|\psi\rangle$ amounts to reversing the order of the factors, and replacing the operator by its Hermitean conjugate. Using this it is then possible to determine the action of $\hat{A}^\dagger$ on a ket vector. The situation here is analogous to that which was encountered in Section 9.2. But before considering the general case, we will look at an example.

Ex 9.15 Suppose an operator $\hat{B}$ is defined, for the two orthonormal states $|\varphi_1\rangle$ and $|\varphi_2\rangle$, by

$$\hat{B}|\varphi_1\rangle = 2|\varphi_2\rangle \quad \text{and} \quad \hat{B}|\varphi_2\rangle = i|\varphi_1\rangle.$$

What are $\hat{B}^\dagger|\varphi_1\rangle$ and $\hat{B}^\dagger|\varphi_2\rangle$?

First consider $\hat{B}^\dagger|\varphi_1\rangle$. We begin by looking at $\langle\chi|\hat{B}^\dagger|\varphi_1\rangle$ where $|\chi\rangle = C_1|\varphi_1\rangle + C_2|\varphi_2\rangle$ is an arbitrary state vector. We then have, by the property proven above:

$$\begin{aligned}\langle\chi|\hat{B}^\dagger|\varphi_1\rangle^* =&\langle\varphi_1|\hat{B}|\chi\rangle\\ =&\langle\varphi_1|\hat{B}\left[C_1|\varphi_1\rangle + C_2|\varphi_2\rangle\right]\\ =&\langle\varphi_1|\left[C_1\hat{B}|\varphi_1\rangle + C_2\hat{B}|\varphi_2\rangle\right]\\ =&\langle\varphi_1|\left[2C_1|\varphi_2\rangle + iC_2|1\rangle\right]\\ =&iC_2\\ =&i\langle\varphi_2|\chi\rangle.\end{aligned}$$

Thus we have shown that

$$\langle\chi|\hat{B}^\dagger|\varphi_1\rangle^* = i\langle\varphi_2|\chi\rangle$$

which becomes, on taking the complex conjugate

$$\langle\chi|\hat{B}^\dagger|\varphi_1\rangle = -i\langle\varphi_2|\chi\rangle^* = -i\langle\chi|\varphi_2\rangle.$$

Since $|\chi\rangle$ is arbitrary, we must conclude that

$$\hat{B}^\dagger|\varphi_1\rangle = -i|\varphi_2\rangle.$$

More generally, suppose we are dealing with a state space spanned by a complete orthonormal set of basis states $\{|\varphi_n\rangle; n = 1, 2, \dots\}$, and suppose we know that action of an operator $\hat{A}$ on each of the basis states:

$$\hat{A}|\varphi_n\rangle = \sum_m |\varphi_m\rangle A_{mn} \tag{9.63}$$

and we want to determine $\hat{A}^\dagger|\psi\rangle$ where $|\psi\rangle$ is an arbitrary ket vector. If we let $\hat{A}^\dagger|\psi\rangle = |\zeta\rangle$, then we can, as usual, make the expansion:

$$|\zeta\rangle = \sum_n |\varphi_n\rangle\langle\varphi_n|\zeta\rangle. \tag{9.64}$$

The coefficients $\langle\varphi_n|\zeta\rangle$ can then be written:

$$\begin{aligned}\langle\varphi_n|\zeta\rangle &= \langle\varphi_n|\hat{A}^\dagger|\psi\rangle \\ &= \langle\psi|\hat{A}|\varphi_n\rangle^* \\ &= \left(\langle\psi|\Big[\sum_m |\varphi_m\rangle A_{mn}\Big]\right)^* \\ &= \sum_m \langle\psi|\varphi_m\rangle^* A^*_{mn}\end{aligned} \tag{9.65}$$

so that

$$\begin{aligned}|\zeta\rangle &= \hat{A}^\dagger|\psi\rangle \\ &= \sum_n\sum_m |\varphi_n\rangle\langle\psi|\varphi_m\rangle^* A^*_{mn} \\ &= \sum_n |n\rangle\Big[\sum_m A^*_{mn}\langle\varphi_m|\psi\rangle\Big].\end{aligned} \tag{9.66}$$

The quantity within the brackets is a complex number which we can always evaluate since we know the A_{mn} and can evaluate the inner product $\langle\varphi_m|\psi\rangle$. Thus, we have shown that the action of the Hermitean adjoint on a ket vector can be readily calculated. Of particular interest is the case in which $|\psi\rangle = |\varphi_k\rangle$ so that

$$\hat{A}^\dagger|\varphi_k\rangle = \sum_n |\varphi_n\rangle\Big[\sum_m A^*_{mn}\langle\varphi_m|\varphi_k\rangle\Big]. \tag{9.67}$$

Using the orthonormality of the basis states, i.e. $\langle\varphi_m|\varphi_k\rangle = \delta_{mk}$ we have

$$\begin{aligned}\hat{A}^\dagger|\varphi_k\rangle &= \sum_n |n\rangle\Big[\sum_m A^*_{mn}\delta_{mk}\Big] \\ &= \sum_n |\varphi_n\rangle A^*_{kn}\end{aligned} \tag{9.68}$$

It is useful to compare this with Eq. (9.46):

$$\begin{aligned}\hat{A}|\varphi_n\rangle &= \sum_m |\varphi_m\rangle A_{mn} \\ \hat{A}^\dagger|\varphi_n\rangle &= \sum_m |\varphi_m\rangle A^*_{nm}\end{aligned} \tag{9.69}$$

From these two results, we see that

$$\langle\varphi_m|\hat{A}|\varphi_n\rangle^* = A^*_{mn} = \langle\varphi_n|\hat{A}^\dagger|\varphi_m\rangle. \tag{9.70}$$

Ex 9.16 We can illustrate the results obtained here using the photon annihilation operator defined in Eq. (9.56). There we showed that the operator $\hat{a}$ defined such that $\hat{a}|n\rangle = \sqrt{n}|n\rangle$ where $|n\rangle$ was the state of a collection of identical photons in which there were n photons present. The question to be addressed here is then: what is $\hat{a}^\dagger|n\rangle$?

THis can be determined by considering $\langle\chi|\hat{a}^\dagger|n\rangle$ where $|\chi\rangle$ is an arbitrary state. We then have

$$\langle\chi|\hat{a}^\dagger|n\rangle^* = \langle n|\hat{a}|\chi\rangle = \{\langle n|\hat{a}\}|\chi\rangle.$$

If we now note, from Eq. (9.57), that $\langle n|\hat{a} = \sqrt{n+1}\langle n+1|$, we have

$$\langle\chi|\hat{a}^\dagger|n\rangle^* = \sqrt{n+1}\langle n+1|\chi\rangle$$

and hence

$$\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle. \tag{9.71}$$

As this operator increases the photon number by unity, it is known as a *creation* operator.

Ex 9.17 Show that $(\hat{A}^\dagger)^\dagger = \hat{A}$.

This is shown to be true by forming the quantity $\langle\phi|(\hat{A}^\dagger)^\dagger|\psi\rangle$ where $|\phi\rangle$ and $|\psi\rangle$ are both arbitrary. We then have

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

since, for any complex number z we have $(z^*)^* = z$. The required result then follows by noting that $|\phi\rangle$ and $|\psi\rangle$ are both arbitrary.

Ex 9.18 An important and useful result is that $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$. To prove this, we once again form the quantity $\langle\phi|(\hat{A}\hat{B})^\dagger|\psi\rangle$ where $|\phi\rangle$ and $|\psi\rangle$ are both arbitrary. We then have

$$\langle\phi|(\hat{A}\hat{B})^\dagger|\psi\rangle^* = \langle\psi|\hat{A}\hat{B}|\phi\rangle = \langle\alpha|\beta\rangle$$

where $\langle\psi|\hat{A} = \langle\alpha|$ and $\hat{B}|\phi\rangle = |\beta\rangle$. Taking the complex conjugate of both sides then gives

$$\langle\phi|(\hat{A}\hat{B})^\dagger|\psi\rangle = \langle\alpha|\beta\rangle^* = \langle\beta|\alpha\rangle = \langle\phi|\hat{B}^\dagger\hat{A}^\dagger|\psi\rangle$$

The required result then follows as $|\phi\rangle$ and $|\psi\rangle$ are both arbitrary.

Much of this discussion on Hermitean operators is recast in terms of matrices and column and row vectors in a later Section.

9.3.1 Hermitean and Unitary Operators

These are two special kinds of operators that play very important roles in the physical interpretation of quantum mechanics.

Hermitean Operators If an operator $\hat{A}$ has the property that

$$\hat{A} = \hat{A}^\dagger \tag{9.72}$$

then the operator is said to be Hermitean. If $\hat{A}$ is Hermitean, then

$$\langle\psi|\hat{A}|\phi\rangle^* = \langle\phi|\hat{A}|\psi\rangle \tag{9.73}$$

and, in particular, for states belonging to a complete set of orthonormal basis states $\{|\varphi_n\rangle; n = 1, 2, 3, \ldots\}$ we have

$$A_{mn} = \langle\varphi_m|\hat{A}|\varphi_n\rangle = \langle\varphi_n|\hat{A}^\dagger|\varphi_m\rangle^* = \langle\varphi_n|\hat{A}|\varphi_m\rangle^* = A^*_{nm}. \tag{9.74}$$

Hermitean operators have a number of important mathematical properties that are discussed in detail in Section 9.4.2. It is because of these properties that Hermitean operators place a central role in quantum mechanics in that the observable properties of a physical system such as position, momentum, spin, energy and so on are represented by Hermitean operators. The physical significance of Hermitean operators will be described in the following Chapter.

Unitary Operators If the operator $\hat{U}$ is such that

$$\hat{U}^\dagger = \hat{U}^{-1} \tag{9.75}$$

then the operator is said to be unitary. Unitary operators have the important property that they map normalized states into normalized states. Thus, for instance, suppose the state $|\psi\rangle$ is normalized to unity, $\langle\psi|\psi\rangle = 1$. We then find that the state $|\phi\rangle = \hat{U}|\psi\rangle$ is also normalized to unity:

$$\langle\phi|\phi\rangle = \langle\psi|\hat{U}^\dagger\hat{U}|\psi\rangle = \langle\psi|\hat{1}|\psi\rangle = \langle\psi|\psi\rangle = 1. \tag{9.76}$$

It is because of this last property that unitary operators play a central role in quantum mechanics in that such operators represent performing actions on a system, such as displacing the system in space or time. The time evolution operator with which the evolution in time of the state of a quantum system can be determined is an important example of a unitary operator. If some action is performed on a quantum system, then the probability interpretation of quantum mechanics makes it physically reasonable to expect that the state of the system after the action is performed should be normalized if the state was initally normalized. If this were not the case, it would mean that this physical action in some way results in the system losing or gaining probability.

Ex 9.19 Consider a negatively charged ozone molecule O_3^-. The oxygen atoms, labelled A, B, and C are arranged in an equilateral triangle. The electron can be found on any one of these atoms, the corresponding state vectors being $|A\rangle$, $|B\rangle$, and $|C\rangle$. An operator $\hat{E}$ can be defined with the properties that

$$\hat{E}|A\rangle = |B\rangle,\ \hat{E}|B\rangle = |C\rangle,\ \hat{E}|C\rangle = |A\rangle$$

i.e. it represents the physical process in which the electron 'jumps' from one atom to its neighbour, as in $|A\rangle \to |B\rangle$, $|B\rangle \to |C\rangle$ and $|C\rangle \to |A\rangle$. Show that the operator $\hat{E}$ is unitary.

To show that $\hat{E}$ is unitary requires showing that $\hat{E}^\dagger\hat{E} = \hat{1}$ which amounts to showing that $\langle\psi|\hat{E}^\dagger\hat{E}|\phi\rangle = \langle\psi|\phi\rangle$ for all states $|\psi\rangle$ and $|\phi\rangle$. As the states $\{|A\rangle, |B\rangle, |C\rangle\}$ form a complete orthonormal set of basis states, we can write

$$|\psi\rangle = a|A\rangle + b|B\rangle + c|C\rangle$$

so that

$$\hat{E}|\psi\rangle = a|B\rangle + b|C\rangle + c|A\rangle.$$

Likewise, if we write

$$|\phi\rangle = \alpha|A\rangle + \beta|B\rangle + \gamma|C\rangle$$

then

$$\hat{E}|\phi\rangle = \alpha|B\rangle + \beta|C\rangle + \gamma|A\rangle$$

and hence

$$\langle\phi|\hat{E}^\dagger = \alpha^*\langle B| + \beta^*\langle C| + \gamma^*\langle A|$$

which gives

$$\langle\psi|\hat{E}^\dagger\hat{E}|\phi\rangle = \alpha^* a + \beta^* b + \gamma^* c = \langle\psi|\phi\rangle.$$

Thus $\hat{E}$ is unitary.

An Analogue with Complex Numbers

It can be a useful mneumonic to note the following analogues between Hermitean and unitary operators and real and unimodular complex numbers respectively. Thus we find that Hermitean operators are the operator analogues of real numbers:

$$\hat{A} = \hat{A}^\dagger \quad \leftrightarrow \quad z = z^* \tag{9.77}$$

while unitary operators are the analogue of complex numbers of unit modulus, i.e. of the form $\exp(i\theta)$ where θ is real:

$$\hat{U}^\dagger = \hat{U}^{-1} \quad \leftrightarrow \quad (e^{i\theta})^* = e^{-i\theta} = (e^{i\theta})^{-1}. \tag{9.78}$$

The analogue goes further that. It turns out that a unitary operator $\hat{U}$ can be written in the form

$$\hat{U} = e^{-i\hat{A}}$$

where $\hat{A}$ is Hermitean. Results of this form will be seen to arise in the cases of unitary operators representing time translation, space displacement, and rotation.

9.4 Eigenvalues and Eigenvectors

It can happen that, for some operator $\hat{A}$, there exists a state vector $|\phi\rangle$ that has the property

$$\hat{A}|\phi\rangle = a_\phi|\phi\rangle \tag{9.79}$$

where a_ϕ is, in general, a complex number. We have seen an example of such a situation in Eq. (9.5). If a situation such as that presented in Eq. (9.79) occurs, then the state $|\phi\rangle$ is said to be an eigenstate or eigenket of the operator $\hat{A}$ with a_ϕ the associated eigenvalue. Often the notation

$$\hat{A}|a\rangle = a|a\rangle \tag{9.80}$$

is used in which the eigenvector is labelled by its associated eigenvalue. This notation will be used almost exclusively here. Determining the eigenvalues and eigenvectors of a given operator $\hat{A}$, occasionally referred to as solving the eigenvalue problem for the operator, amounts to finding solutions to the eigenvalue equation Eq. (9.79). If the vector space is of finite dimension, then this can be done by matrix methods, while if the state space is of infinite dimension, then solving the eigenvalue problem can require solving a differential equation. Examples of both possibilities will be looked at later. An operator $\hat{A}$ may have

1. no eigenstates;
2. real or complex eigenvalues;
3. a discrete collection of eigenvalues $a_1, a_2, \ldots$ and associated eigenvectors $|a_1\rangle$, $|a_2\rangle$, $\ldots$;
4. a continuous range of eigenvalues and associated eigenvectors;
5. a combination of both discrete and continuous eigenvalues.

The collection of all the eigenvalues of an operator is called the *eigenvalue spectrum* of the operator. Note also that more than one eigenvector can have the same eigenvalue. Such an eigenvalue is said to be *degenerate*.

Ex 9.20 An interesting example of an operator with complex eigenvalues is the annihilation operator $\hat{a}$ introduced in Eq. (9.56). This operator maps the state of a system of identical photons in which there is exactly n photons present, $|n\rangle$, into the state $|n-1\rangle$: $\hat{a}|n\rangle = \sqrt{n}|n-1\rangle$.

The eigenstates of this operator can be found by looking for the solutions to the eigenvalue equation $\hat{a}|\alpha\rangle = \alpha|\alpha\rangle$ where α and $|\alpha\rangle$ are the eigenvalue and associated eigenstate to be determined. Expanding $|\alpha\rangle$ in terms of the number state basis $\{|n\rangle; n = 0, 1, 2, 3, \ldots\}$ gives

$$|\alpha\rangle = \sum_{n=0}^{\infty} c_n|n\rangle$$

where, by using the orthonormality condition $\langle n|m\rangle = \delta_{nm}$, we have $c_n = \langle n|\alpha\rangle$. Thus it follows that

$$\hat{a}|\alpha\rangle = \sum_{n=1}^{\infty} c_n \sqrt{n}|n-1\rangle = \sum_{n=0}^{\infty} c_{n+1} \sqrt{n+1}|n\rangle = \sum_{n=0}^{\infty} \alpha c_n|n\rangle$$

where the last expression is just $\alpha|\alpha\rangle$.

Equating coefficients of $|n\rangle$ gives

$$c_{n+1} = \frac{\alpha c_n}{\sqrt{n+1}}$$

a recurrence relation, from which we can build up each coefficient from the one before, i.e. assuming $c_0 \neq 0$ we have

$$\begin{aligned}
c_1 &= \frac{\alpha}{\sqrt{1}}c_0 \\
c_2 &= \frac{\alpha}{\sqrt{2}}c_1 = \frac{\alpha^2}{\sqrt{2.1}}c_0 \\
c_3 &= \frac{\alpha}{\sqrt{3}}c_2 = \frac{\alpha^3}{\sqrt{3!}}c_0 \\
&\vdots \\
c_n &= \frac{\alpha^n}{\sqrt{n!}}c_0
\end{aligned}$$

and hence

$$|\alpha\rangle = c_0 \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}}|n\rangle.$$

Requiring this state to be normalized to unity gives

$$\langle\alpha|\alpha\rangle = \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}}\langle\alpha|n\rangle c_0.$$

But $c_n = \langle n|\alpha\rangle$, so $\langle\alpha|n\rangle = c_n^*$ and we have

$$\langle\alpha|n\rangle = \frac{\alpha^{*n}}{\sqrt{n!}}c_0^*$$

and hence

$$\langle\alpha|\alpha\rangle = 1 = |c_0|^2 \sum_{n=0}^{\infty} \frac{|\alpha|^{2n}}{n!} = |c_0|^2 e^{|\alpha|^2}.$$

Thus

$$c_0 = e^{-|\alpha|^2/2}$$

where we have set an arbitrary phase factor to unity. Thus we end up with

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}}|n\rangle. \tag{9.81}$$

This then is the required eigenstate of $\hat{a}$. It is known as a *coherent state*, and plays a very important role, amongst other things, as the 'most classical' state possible for a harmonic oscillator, which includes the electromagnetic field.

It should be noted in this derivation that no restriction was needed on the value of α. In other words all the states $|\alpha\rangle$ for any value of the complex number α will be an eigenstate of $\hat{a}$. It can also be shown that the states $|\alpha\rangle$ are not orthogonal for different vaues of α, i.e. $\langle\alpha'|\alpha\rangle \neq 0$ if $\alpha' \neq \alpha$, a fact to be contrasted with what is seen later when the eigenstates of Hermitean operators are considered.

Attempting a similar calculation to the above to try to determine what are the eigenstates of the creation operator $\hat{a}^\dagger$ for which $\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n\rangle$ (see p126) quickly shows that this operator has, in fact, no eigenstates.

Ex 9.21 If $|a\rangle$ is an eigenstate of $\hat{A}$ with eigenvalue a, then a function $f(\hat{A})$ of $\hat{A}$ (where the function can be expanded as a power series) will also have $|a\rangle$ as an eigenstate with eigenvalue $f(a)$. This can be reasily shown by first noting that

$$\hat{A}^n|a\rangle = \hat{A}^{(n-1)}\hat{A}|a\rangle = \hat{A}^{(n-1)}a|a\rangle = a\hat{A}^{(n-1)}|a\rangle.$$

Repeating this a further $n-1$ times then yields

$$\hat{A}^n|a\rangle = a^n|a\rangle.$$

If we then apply this to

$$f(\hat{A})|a\rangle$$

where $f(\hat{A})$ has the power series expansion

$$f(\hat{A}) = c_0 + c_1\hat{A} + c_2\hat{A}^2 + \ldots$$

then

$$\begin{aligned} f(\hat{A})|a\rangle &= (c_0 + c_1\hat{A} + c_2\hat{A}^2 + \ldots)|a\rangle \\ &= c_0|a\rangle + c_1 a|a\rangle + c_2 a^2|a\rangle + \ldots \\ &= (c_0 + c_1 a + c_2 a^2 + \ldots)|a\rangle \\ &= f(a)|a\rangle. \end{aligned} \quad (9.82)$$

This turns out to be a very valuable result as we will often encounter functions of operators when we deal, in particular, with the time evolution operator. The time evolution operator is expressed as an exponential function of another operator (the Hamiltonian) whose eigenvalues and eigenvectors are central to the basic formalism of quantum mechanics.

9.4.1 Eigenkets and Eigenbras

The notion of an eigenstate has been introduced above with respect to ket vectors, in which case the eigenstates could be referred to as *eigenkets*. The same idea of course can be applied to bra vectors, i.e. if for some operator $\hat{A}$ there exists a bra vector $\langle\phi|$ such that

$$\langle\phi|\hat{A} = a_\phi\langle\phi| \quad (9.83)$$

then $\langle\phi|$ is said to be an eigenbra of $\hat{A}$ and a_ϕ the associated eigenvalue. The eigenvalues in this case can have the same array of possible properties as listed above for eigenkets.

It is important to note that an operator may have eigenkets, but have no eigenbras. Thus, we have seen that the annihilation operator $\hat{a}$ has a continuously infinite number of eigenkets, but it has no eigenbras, for the same reason that $\hat{a}^\dagger$ has no eigenkets.

9.4.2 Eigenstates and Eigenvalues of Hermitean Operators

If an operator is Hermitean then its eigenstates and eigenvalues are found to possess a number of mathematical properties that are of substantial significance in quantum mechanics. So, if we suppose that an operator $\hat{A}$ is Hermitean i.e. $\hat{A} = \hat{A}^\dagger$ then the following three properties hold true.

1. The eigenvalues of $\hat{A}$ are all real.

The proof is as follows. Since
$$\hat{A}|a\rangle = a|a\rangle$$
then
$$\langle a|\hat{A}|a\rangle = a\langle a|a\rangle.$$
Taking the complex conjugate then gives
$$\langle a|\hat{A}|a\rangle^* = a^*\langle a|a\rangle.$$
Now, using the facts that $\langle\phi|\hat{A}|\psi\rangle = \langle\psi|\hat{A}|\phi\rangle^*$ (Eq. (9.62)), and that $\langle a|a\rangle$ is real, we have
$$\langle a|\hat{A}^\dagger|a\rangle = a^*\langle a|a\rangle.$$
Since $\hat{A} = \hat{A}^\dagger$ this then gives
$$\langle a|\hat{A}|a\rangle = a^*\langle a|a\rangle = a\langle a|a\rangle$$
and hence
$$(a^* - a)\langle a|a\rangle = 0.$$
And so, finally, since $\langle a|a\rangle \neq 0$,
$$a^* = a.$$

This property is of central importance in the physical interpretation of quantum mechanics in that all physical observable properties of a system are represented by Hermitean operators, with the eigenvalues of the operators representing all the possible values that the physical property can be observed to have.

2. Eigenvectors belonging to different eigenvalues are orthogonal, i.e. if $\hat{A}|a\rangle = a|a\rangle$ and $\hat{A}|a'\rangle = a'|a'\rangle$ where $a \neq a'$, then $\langle a|a'\rangle = 0$.

The proof is as follows. Since
$$\hat{A}|a\rangle = a|a\rangle$$
then
$$\langle a'|\hat{A}|a\rangle = a\langle a'|a\rangle.$$
But
$$\hat{A}|a'\rangle = a'|a'\rangle$$
so that
$$\langle a|\hat{A}|a'\rangle = a'\langle a|a'\rangle$$
and hence on taking the complex conjugate
$$\langle a'|\hat{A}^\dagger|a\rangle = a'^*\langle a'|a\rangle = a'\langle a'|a\rangle$$
where we have used the fact that the eigenvalues of $\hat{A}$ are real, and hence $a' = a'^*$. Overall then,
$$\langle a'|\hat{A}|a\rangle = a'\langle a'|a\rangle = a\langle a'|a\rangle$$

and hence

$$(a' - a)\langle a'|a\rangle = 0$$

so finally, if $a' \neq a$, then

$$\langle a'|a\rangle = 0.$$

The importance of this result lies in the fact that it makes it possible to construct a set of orthonormal states that define a basis for the state space of the system. To do this, we need the next property of Hermitean operators.

3. The eigenstates form a complete set of basis states for the state space of the system.

This can be proven to be always true if the state space is of finite dimension. If the state space is of infinite dimension, then completeness of the eigenstates of a Hermitean operator is not guaranteed. As we will see later, this has some consequences for the physical interpretation of such operators in quantum mechanics.

We can also always assume, *if the eigenvalue spectrum is discrete*, that these eigenstates are normalized to unity. If we were to suppose that they were not so normalized, for instance if the eigenstate $|a\rangle$ of the operator $\hat{A}$ is such that $\langle a|a\rangle \neq 1$, then we simply define a new state vector by

$$\widetilde{|a\rangle} = \frac{|a\rangle}{\sqrt{\langle a|a\rangle}} \tag{9.84}$$

which is normalized to unity. This new state $\widetilde{|a\rangle}$ is still an eigenstate of $\hat{A}$ with eigenvalue a – in fact it represents the same physical state as $|a\rangle$ – so we might as well have assumed from the very start that $|a\rangle$ was normalized to unity. Thus, provided the eigenvalue spectrum is discrete, then as well as the eigenstates forming a complete set of basis states, they also form an orthonormal set. Thus, if the operator $\hat{A}$ is Hermitean, and has a complete set of eigenstates $\{|a_n\rangle; n = 1, 2, 3 \dots\}$, then these eigenstates form an orthonormal basis for the system. This means that any arbitrary state $|\psi\rangle$ can be written as

$$|\psi\rangle = \sum_n |a_n\rangle\langle a_n|\psi\rangle. \tag{9.85}$$

If the eigenvalue spectrum of an operator is continuous, then it is not possible to assume that the eigenstates can be normalized to unity. A different normalization scheme is required, as will be discussed in the next section.

9.4.3 Continuous Eigenvalues

Far from being the exception, Hermitean operators with continuous eigenvaues are basic to quantum mechanics, and it is consequently necessary to come to some understanding of the way the continuous case is distinct from the discrete case, and where they are the same. So in the following, consider a Hermitean operator $\hat{A}$ with continuous eigenvalues a lying in some range, between α_1 and α_2 say:

$$\hat{A}|a\rangle = a|a\rangle \qquad \alpha_1 < a < \alpha_2. \tag{9.86}$$

That there is a difficulty in dealing with eigenstates associated with a continuous range of eigenvalues can be seen if we make use of the (assumed) completeness of the eigenstates of a Hermitean operator, Eq. (9.85). It seems reasonable to postulate that in the case of continuous eigenvalues, this completeness relation would become an integral over the continuous range of eigenvalues:

$$|\psi\rangle = \int_{\alpha_1}^{\alpha_2} |a\rangle\langle a|\psi\rangle da. \tag{9.87}$$

We have seen this situation before in the discussion in Section 8.6.3 of the basis states $|x\rangle$ for the position of a particle. There we argued that the above form of the completeness relation can be used, but doing so requires that the inner product $\langle a'|a\rangle$, must be interpreted as a delta function:

$$\langle a'|a\rangle = \delta(a - a'). \tag{9.88}$$

The states $|a\rangle$ are said to be delta function normalized, in contrast to the orthonormal property of discrete eigenstates. As pointed out in Section 8.6.3, the result of this is that states such as $|a\rangle$ are of infinite norm and so cannot be normalized to unity. Such states states cannot represent possible physical states of a system, which is an awkward state of affairs if the state is supposed to represent hat appears to be a physically reasonable state of the system. Fortunately it is possible to think of such states as idealized limits, and to work with them as if they were physically realizable, provided care is taken. Mathematical (and physical) paradoxes can arise otherwise. However, linear combinations of these states can be normalized to unity, as this following example illustrates. If we consider a state $|\psi\rangle$ given by

$$|\psi\rangle = \int_{\alpha_1}^{\alpha_2} |a\rangle\langle a|\psi\rangle da, \tag{9.89}$$

then

$$\langle\psi|\psi\rangle = \int_{\alpha_1}^{\alpha_2} \langle\psi|a\rangle\langle a|\psi\rangle da. \tag{9.90}$$

But $\langle a|\psi\rangle = \psi(a)$ and $\langle\psi|a\rangle = \psi(a)^*$, so that

$$\langle\psi|\psi\rangle = \int_{\alpha_1}^{\alpha_2} |\psi(a)|^2 da. \tag{9.91}$$

Provided $|\psi(a)|^2$ is a well behaved function, this integral will converge to a finite result, so that the state $|\psi\rangle$ can indeed be normalized to unity and thus represent physically realizable states.

9.5 Dirac Notation for Operators

The above discussion of the properties of operators was based on making direct use of the defining properties of an operator, that is, in terms of their actions on ket vectors, in particular the vectors belonging to a set of basis states. All of these properties can be represented in a very succinct way that makes explicit use of the Dirac notation. The essential idea is to give a meaning to the symbol $|\phi\rangle\langle\psi|$, and we can see how this meaning is arrived at by considering the following example.

Suppose we have a spin half system with basis states $\{|+\rangle, |-\rangle\}$ and we have an operator $\hat{A}$ defined such that

$$\left.\begin{aligned}\hat{A}|+\rangle &= a|+\rangle + b|-\rangle\\ \hat{A}|-\rangle &= c|+\rangle + d|-\rangle\end{aligned}\right\} \tag{9.92}$$

and we calculate the quantity $\langle\phi|\hat{A}|\psi\rangle$ where $|\phi\rangle$ and $|\psi\rangle$ are arbitrary states. This is given by

$$\begin{aligned}\langle\phi|\hat{A}|\psi\rangle &= \langle\phi|\Big\{\hat{A}(|+\rangle\langle+|\psi\rangle + |-\rangle\langle-|\psi\rangle)\Big\}\\ &= \langle\phi|\Big[(a|+\rangle + b|-\rangle)\langle+|\psi\rangle + (c|+\rangle + d|-\rangle)\langle-|\psi\rangle\big]\\ &= \langle\phi|\big[a|+\rangle\langle+|\psi\rangle + b|-\rangle\langle+|\psi\rangle + c|+\rangle\langle-|\psi\rangle + d|-\rangle\langle-|\psi\rangle\Big].\end{aligned} \tag{9.93}$$

We note that the term enclosed within the square brackets contains, symbolically at least, a common 'factor' $|\psi\rangle$ which we will move outside the brackets to give

$$\langle\phi|\hat{A}|\psi\rangle = \langle\phi|\Big[a|+\rangle\langle+| + b|-\rangle\langle+| + c|+\rangle\langle-| + d|-\rangle\langle-|\Big]|\psi\rangle \tag{9.94}$$

It is now tempting to make the identification of the operator $\hat{A}$ appearing on the left hand side of this expression with the combination of symbols appearing between the square brackets on the right hand side of the equation, i.e.

$$\hat{A} \leftrightarrow a|+\rangle\langle+| + b|-\rangle\langle+| + c|+\rangle\langle-| + d|-\rangle\langle-|. \tag{9.95}$$

We can do so provided we give appropriate meanings to this combination of ket-bra symbols such that it behaves in exactly the same manner as the operator $\hat{A}$ itself. Thus if we require that the action of this combination on a ket be given by

$$\begin{aligned}&\Big[a|+\rangle\langle+| + b|-\rangle\langle+| + c|+\rangle\langle-| + d|-\rangle\langle-|\Big]|\psi\rangle\\ &\qquad= a|+\rangle\langle+|\psi\rangle + b|-\rangle\langle+|\psi\rangle + c|+\rangle\langle-|\psi\rangle + d|-\rangle\langle-|\psi\rangle\\ &\qquad= |+\rangle\big(a\langle+|\psi\rangle + c\langle-|\psi\rangle\big) + |-\rangle\big(b\langle+|\psi\rangle + d\langle-|\psi\rangle\big)\end{aligned} \tag{9.96}$$

we see that this gives the correct result for $\hat{A}$ acting on the ket $|\psi\rangle$. In particular, if $|\psi\rangle = |\pm\rangle$ we recover the defining equations for $\hat{A}$ given in Eq. (9.92). If we further require that the action of this combination on a bra be given by

$$\begin{aligned}&\langle\phi|\Big[a|+\rangle\langle+| + b|-\rangle\langle+| + c|+\rangle\langle-| + d|-\rangle\langle-|\Big]\\ &\qquad= a\langle\phi|+\rangle\langle+| + b\langle\phi|-\rangle\langle+| + c\langle\phi|+\rangle\langle-| + d\langle\phi|-\rangle\langle-|\\ &\qquad= \big(a\langle\phi|+\rangle + b\langle\phi|-\rangle\big)\langle+| + \big(c\langle\phi|+\rangle + d\langle\phi|-\rangle\big)\langle-|\end{aligned} \tag{9.97}$$

we see that this gives the correct result for $\hat{A}$ acting on the bra $\langle\psi|$. In particular, if $\langle\psi| = \langle\pm|$, this gives

$$\left.\begin{aligned}\langle+|\hat{A} &= a\langle+| + c\langle-|\\ \langle-|\hat{A} &= b\langle+| + d\langle-|\end{aligned}\right\} \tag{9.98}$$

which can be checked, using the defining condition for the action of an operator on a bra vector, Eq. (9.39), to be the correct result.

Thus we see that we can indeed write

$$\hat{A} = a|+\rangle\langle+| + b|-\rangle\langle+| + c|+\rangle\langle-| + d|-\rangle\langle-|. \tag{9.99}$$

as an valid expression for the operator $\hat{A}$ in terms of bra and ket symbols, provided we interpret the symbols in the manner indicated above, and summarized in more detail below.

The interpretation that is given is defined as follows:

$$\begin{aligned}(|\phi\rangle\langle\psi|)|\alpha\rangle &= |\phi\rangle\langle\psi|\alpha\rangle\\ \langle\alpha|(|\phi\rangle\langle\psi|) &= \langle\alpha|\phi\rangle\langle\psi|\end{aligned} \tag{9.100}$$

i.e. it maps kets into kets and bras into bras, exactly as an operator is supposed to.

If we further require $|\phi\rangle\langle\psi|$ to have the linear property

$$\begin{aligned}|\phi\rangle\langle\psi|(c_1|\psi_1\rangle + c_2|\psi_2\rangle) &= c_1(|\phi\rangle\langle\psi|)|\psi_1\rangle + c_2(|\phi\rangle\langle\psi|)|\psi_2\rangle\\ &= |\phi\rangle(c_1\langle\psi|\psi_1\rangle + c_2\langle\psi|\psi_2\rangle)\end{aligned} \tag{9.101}$$

and similarly for the operator acting on bra vectors, we have given the symbol the properties of a *linear operator*.

We can further generalize this to include sums of such bra-ket combinations, e.g.

$$c_1|\phi_1\rangle\langle\psi_1| + c_2|\phi_2\rangle\langle\psi_2|$$

where c_1 and c_2 are complex numbers, is an operator such that

$$(c_1|\phi_1\rangle\langle\psi_1| + c_2|\phi_2\rangle\langle\psi_2|)|\xi\rangle = c_1|\phi_1\rangle\langle\psi_1|\xi\rangle + c_2|\phi_2\rangle\langle\psi_2|\xi\rangle \tag{9.102}$$

and similarly for the action on bra vectors.

Finally, we can define the product of bra-ket combinations in the obvious way, that is

$$(|\phi\rangle\langle\psi|)(|\alpha\rangle\langle\beta|) = |\phi\rangle\langle\psi|\alpha\rangle\langle\beta| = \langle\psi|\alpha\rangle|\phi\rangle\langle\beta|. \tag{9.103}$$

Below we describe a number of examples that illustrate the usefulness of this notation.

Ex 9.22 The three operators (the Pauli spin operators) for a spin half system whose state space is spanned by the usual basis states $\{|+\rangle, |-\rangle\}$ are given, in Dirac notation, by the expressions

$$\begin{aligned}\hat{\sigma}_x &= |-\rangle\langle+| + |+\rangle\langle-|\\ \hat{\sigma}_y &= i|-\rangle\langle+| - i|+\rangle\langle-|\\ \hat{\sigma}_z &= |+\rangle\langle+| - |-\rangle\langle-|.\end{aligned}$$

Determine the action of these operators on the basis states $|\pm\rangle$.

First we consider $\hat{\sigma}_x|+\rangle$ which can be written

$$\hat{\sigma}_x|+\rangle = \Big[|-\rangle\langle+| + |+\rangle\langle-|\Big]|+\rangle = |-\rangle\langle+|+\rangle + |+\rangle\langle-|+\rangle = |-\rangle.$$

Similarly, for instance

$$\hat{\sigma}_y|-\rangle = \Big[i|-\rangle\langle+| - i|+\rangle\langle-|\Big]|-\rangle = i|-\rangle\langle+|-\rangle - i|+\rangle\langle-|-\rangle = -i|+\rangle.$$

In each of the above examples, the orthonormality of the states $|\pm\rangle$ has been used.

Ex 9.23 For the Pauli spin operators defined above, determine the action of these operators on the bra vectors $\langle\pm|$.

We find, for instance

$$\langle-|\hat{\sigma}_z = \langle-|\Big[|+\rangle\langle+| - |-\rangle\langle-|\Big] = \langle-|+\rangle\langle+| - \langle-|-\rangle\langle-| = -\langle-|.$$

Ex 9.24 Calculate the product $\hat{\sigma}_x\hat{\sigma}_y$ and the commutator $[\hat{\sigma}_x, \hat{\sigma}_y]$.

This product is:

$$\begin{aligned}\hat{\sigma}_x\hat{\sigma}_y &= \Big[|-\rangle\langle+| + |+\rangle\langle-|\Big]\Big[i|-\rangle\langle+| - i|+\rangle\langle-|\Big]\\ &= i|-\rangle\langle+|-\rangle\langle+| - i|-\rangle\langle+|+\rangle\langle-| + i|+\rangle\langle-|-\rangle\langle+| - i|+\rangle\langle-|+\rangle\langle-|\\ &= -i|-\rangle\langle-| + i|+\rangle\langle+|\\ &= i\hat{\sigma}_z.\end{aligned}$$

In the same fashion, it can be shown that $\hat{\sigma}_y\hat{\sigma}_x = -i\hat{\sigma}_z$ so that we find that

$$[\hat{\sigma}_x, \hat{\sigma}_y] = 2i\hat{\sigma}_z.$$

There are further important properties of this Dirac notation for operators worth highlighting.

Projection Operators In this notation, a projection operator $\hat{P}$ will be simply given by

$$\hat{P} = |\psi\rangle\langle\psi| \tag{9.104}$$

provided $|\psi\rangle$ is normalized to unity, since we have

$$\hat{P}^2 = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = \hat{P} \tag{9.105}$$

as required for a projection operator.

Completeness Relation This new notation also makes it possible to express the completeness relation in a particularly compact form. Recall that if the set of ket vectors $\{|\varphi_n\rangle; n = 1, 2, 3\dots\}$ is a complete set of orthonormal basis states for the state space of a system, then any state $|\psi\rangle$ can be written

$$|\psi\rangle = \sum_n |\varphi_n\rangle\langle\varphi_n|\psi\rangle \tag{9.106}$$

which in our new notation can be written

$$|\psi\rangle = \Big(\sum_n |\varphi_n\rangle\langle\varphi_n|\Big)|\psi\rangle \tag{9.107}$$

so that we must conclude that

$$\sum_n |\varphi_n\rangle\langle\varphi_n| = \hat{1} \tag{9.108}$$

where $\hat{1}$ is the unit operator. It is often referred to as a *decomposition of unity*.

In the case of continuous eigenvalues, the same argument as above can be followed through. Thus, if we suppose that a Hermitean operator $\hat{A}$ has a set of eigenstates $\{|a\rangle; \alpha_1 < a < \alpha_2\}$, then we can readily show that

$$\int_{\alpha_1}^{\alpha_2} |a\rangle\langle a| da = \hat{1}. \tag{9.109}$$

Note that, in practice, it is often the case that an operator can have both discrete and continuous eigenvalues, in which case the completeness relation can be written

$$\sum_n |\varphi_n\rangle\langle\varphi_n| + \int_{\alpha_1}^{\alpha_2} |a\rangle\langle a| da = \hat{1} \tag{9.110}$$

The completeness relation expressed in this fashion (in both the discrete and continuous cases) is extremely important and has widespread use in calculational work, as illustrated in the following examples.

Ex 9.25 Show that any operator can be expressed in terms of this Dirac notation. We can see this for an operator A by writing

$$\hat{A} = \hat{1}\hat{A}\,\hat{1} \tag{9.111}$$

and using the decomposition of unity twice over to give

$$\begin{aligned}\hat{A} &= \sum_m \sum_n |\varphi_m\rangle\langle\varphi_m|\hat{A}|\varphi_n\rangle\langle\varphi_n| \\ &= \sum_m \sum_n |\varphi_m\rangle\langle\varphi_n| A_{mn}\end{aligned} \tag{9.112}$$

where $A_{mn} = \langle\varphi_m|\hat{A}|\varphi_n\rangle$.

Ex 9.26 Using the decomposition of unity in terms of the basis states $\{|\varphi_n\rangle; n = 1, 2, 3\dots\}$, expand $\hat{A}|\varphi_m\rangle$ in terms of these basis states. This calculation proceeds by inserting the unit operator in a convenient place:

$$\begin{aligned}\hat{A}|\varphi_m\rangle = \hat{1}\hat{A}|\varphi_m\rangle &= \Big(\sum_n |\varphi_n\rangle\langle\varphi_n|\Big)\hat{A}|\varphi_m\rangle \\ &= \sum_n |\varphi_n\rangle\langle\varphi_n|\hat{A}|\varphi_m\rangle \\ &= \sum_n A_{nm}|\varphi_n\rangle\end{aligned} \tag{9.113}$$

where $A_{nm} = \langle\varphi_n|\hat{A}|\varphi_m\rangle$.

Ex 9.27 Using the decomposition of unity, we can insert the unit operator between the two operators in the quantity $\langle\psi|\hat{A}\hat{B}|\phi\rangle$ to give

$$\langle\psi|\hat{A}\hat{B}|\phi\rangle = \langle\psi|\hat{A}\,\hat{1}\hat{B}|\phi\rangle = \sum_n \langle\psi|\hat{A}|\varphi_n\rangle\langle\varphi_n|\hat{B}|\phi\rangle. \tag{9.114}$$

Hermitean conjugate of an operator It is straightforward to write down the Hermitean conjugate of an operator. Thus, for the operator $\hat{A}$ given by

$$\hat{A} = \sum_n c_n |\phi_n\rangle\langle\psi_n| \tag{9.115}$$

we have

$$\langle\phi|\hat{A}|\psi\rangle = \sum_n c_n \langle\phi|\phi_n\rangle\langle\psi_n|\psi\rangle \tag{9.116}$$

so that taking the complex conjugate we get

$$\langle\psi|\hat{A}^\dagger|\phi\rangle = \sum_n c_n^* \langle\psi|\psi_n\rangle\langle\phi_n|\phi\rangle = \langle\psi|\Big(\sum_n c_n^* |\psi_n\rangle\langle\phi_n|\Big)|\phi\rangle. \tag{9.117}$$

We can then extract from this the result

$$\hat{A}^\dagger = \sum_n c_n^* |\psi_n\rangle\langle\phi_n|. \tag{9.118}$$

Spectral decomposition of an operator As a final important result, we can look at the case of expressing an Hermitean operator in terms of projectors onto its basis states. Thus, if we suppose that $\hat{A}$ has the eigenstates $\{|a_n\rangle; n = 1, 2, 3\ldots\}$ and associated eigenvalues $a_n, n = 1, 2, 3\ldots$, so that

$$\hat{A}|a_n\rangle = a_n|a_n\rangle \tag{9.119}$$

then by noting that the eigenstates of $\hat{A}$ form a complete orthonormal set of basis states we can write the decomposition of unity in terms of the eigenstates of $\hat{A}$ as

$$\sum_n |a_n\rangle\langle a_n| = \hat{1}. \tag{9.120}$$

Thus we find that

$$\hat{A} = \hat{A}\,\hat{1} = \hat{A}\sum_n |a_n\rangle\langle a_n| = \sum_n \hat{A}|a_n\rangle\langle a_n| = \sum_n a_n|a_n\rangle\langle a_n|. \tag{9.121}$$

so that

$$\hat{A} = \sum_n a_n|a_n\rangle\langle a_n|. \tag{9.122}$$

The analogous result for continuous eigenstates is then

$$\hat{A} = \int_{\alpha_1}^{\alpha_2} a|a\rangle\langle a|da \tag{9.123}$$

while if the operator has both continuous and discrete eigenvalues, the result is

$$\hat{A} = \sum_n a_n|a_n\rangle\langle a_n| + \int_{\alpha_1}^{\alpha_2} a|a\rangle\langle a|da. \tag{9.124}$$

This is known as the spectral decomposition of the operator $\hat{A}$, the name coming, in part, from the fact that the collection of eigenvalues of an operator is known as its eigenvalue spectrum.

Ex 9.28 Calculate the spectral decomposition of the operator $\hat{A}^2$, where $\hat{A}$ is as given in Eq. (9.124).

We can write $\hat{A}^2$ as

$$\begin{aligned}\hat{A}^2 &= \hat{A}\sum_n a_n|a_n\rangle\langle a_n| + \hat{A}\int_{\alpha_1}^{\alpha_2} a|a\rangle\langle a|da \\ &= \sum_n a_n\hat{A}|a_n\rangle\langle a_n| + \int_{\alpha_1}^{\alpha_2} a\hat{A}|a\rangle\langle a|da \\ &= \sum_n a_n^2|a_n\rangle\langle a_n| + \int_{\alpha_1}^{\alpha_2} a^2|a\rangle\langle a|da.\end{aligned}$$

Ex 9.29 Express $f(\hat{A})$ in Dirac notation, where $\hat{A}$ is a Hermitean operator given by Eq. (9.124) and where $f(x)$ can be expanded as a power series in x.

From the preceding exercise, it is straightforward to show that

$$\hat{A}^k = \sum_n a_n^k|a_n\rangle\langle a_n| + \int_{\alpha_1}^{\alpha_2} a^k|a\rangle\langle a|da.$$

Since $f(x)$ can be expanded as a power series in x, we have

$$f(\hat{A}) = \sum_{k=0}^{\infty} c_k\hat{A}^k$$

so that

$$\begin{aligned}f(\hat{A}) &= \sum_{k=0}^{\infty} c_k\left\{\sum_n a_n^k|a_n\rangle\langle a_n| + \int_{\alpha_1}^{\alpha_2} a^k|a\rangle\langle a|da\right\} \\ &= \sum_n\left\{\sum_{k=0}^{\infty} c_k a_n^k\right\}|a_n\rangle\langle a_n| + \int_{\alpha_1}^{\alpha_2}\left\{\sum_{k=0}^{\infty} c_k a^k\right\}|a\rangle\langle a|da \\ &= \sum_n f(a_n)|a_n\rangle\langle a_n| + \int_{\alpha_1}^{\alpha_2} f(a)|a\rangle\langle a|da\end{aligned}$$

Ex 9.30 Determine $f(\hat{A})|a_n\rangle$ where $\hat{A} = \sum_n a_n|a_n\rangle\langle a_n|$.

In this case, we can use the expansion for $f(\hat{A})$ as obtained in the previous example, that is

$$f(\hat{A}) = \sum_n f(a_n)|a_n\rangle\langle a_n|$$

so that

$$f(\hat{A})|a_k\rangle = \sum_n f(a_n)|a_n\rangle\langle a_n|a_k\rangle = \sum_n f(a_n)|a_n\rangle\delta_{nk} = f(a_k)|a_k\rangle.$$

Since k is a dummy index here, we can write this as $f(\hat{A})|a_n\rangle = f(a_n)|a_n\rangle$. Thus, the effect is simply to replace the operator in $f(\hat{A})$ by its eigenvalue, i.e. $f(a_n)$.

This last result can be readily shown to hold true in the case of continuous eigenvalues. It is a very important result that finds very frequent application in practice.