

Chapter 10

Representations of State Vectors and Operators

In the preceding Chapters, the mathematical ideas underpinning the quantum theory have been developed in a fairly general (though, admittedly, not a mathematically rigorous) fashion. However, much that has been presented, particularly the concept of an operator, can be developed in another way that is, in some respects, less abstract than what has been used so far. This alternate form of presentation involves working with the *components* of the state vectors and operators, leading to their being represented by column and row vectors, and matrices. This development of the theory is completely analogous to the way in which this is done when dealing with the position vector in ordinary three dimensional space. Below, we will look at how the idea of writing the position vectors in two dimensional space can be written in terms of column and row vectors. We will then use the ideas developed there to show how state vectors and operators can be expressed in a similar fashion. This alternate route offers another way of introducing such concepts as adding, multiplying and taking the inverse of operators through their representations as matrices, and further provides another way to introduce the idea of the Hermitean adjoint of an operator, and of a Hermitian operator.

10.1 Representation of Vectors In Euclidean Space as Column and Row Vectors

When writing down a vector, we have so far made explicit the basis vectors when writing an expression such as $\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}}$ for a position vector, or $|S\rangle = a|+\rangle + b|-\rangle$ for the state of a spin half system. But the choice of basis vectors is not unique, i.e. we could use any other pair of orthonormal unit vectors $\hat{\mathbf{i}}'$ and $\hat{\mathbf{j}}'$, and express the vector \mathbf{r} in terms of these new basis vectors, though of course the components of \mathbf{r} will change. The same is true for the spin basis vectors $|\pm\rangle$, i.e. we can express the state $|S\rangle$ in terms of some other basis vectors, such as the states for which the x component of spin has the values $S_x = \frac{1}{2}\hbar$, though once again, the components of $|S\rangle$ will now be different. But it is typically the case that once the choice of basis vectors have been decided on, it should not be necessary to always write them down when writing down a vector, i.e. it would be just as useful to just write down the components of a vector. Doing this leads to a convenient notation in

which vectors are written in terms of column and row vectors. It also provides a direct route to some of the important mathematical entities encountered in quantum mechanics such as bra vectors and operators that are more rigorously introduced in a more abstract way.

10.1.1 Column Vectors

To illustrate the ideas, we will use the example of a position vector in two dimensional space. The point that is to be established here is that a position vector is an independently existing geometrical object ‘suspended’ in space, much as a pencil held in the air with a steady position and orientation has a fixed length and orientation. One end of the pencil, say where the eraser is, can be taken to be the origin O , and the other end (the sharp end) the position of a point P . Then the position and orientation of the pencil defines a position vector \mathbf{r} of P with respect to the origin O . This vector can be represented by a single arrow joining O to P whose length and orientation specify the position of P with respect to O . As such, the vector \mathbf{r} also has an independent existence as a geometrical object sitting in space, and we can work with this vector \mathbf{r} and others like it by, for instance, performing vector additions by using the triangle law of vector addition as illustrated in Fig. (8.1), or performing scalar products by use of the definition Eq. (8.2).

In what was just described, we work only with the whole vector itself. This is in contrast with a very useful alternate way of working with vectors, that is to express any vector as a linear combination of a pair of basis vectors (in two dimensions), which amounts to building around these vectors some sort of ‘scaffolding’, a coordinate system such as a pair of X and Y axes, and describe the vector in terms of its components with respect to these axes. More to the point, what is provided is a pair of basis vectors such as the familiar unit vectors $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$ and write $\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}}$. We see that *any* position vector can be written in this way, i.e. the unit vectors constitute a pair of orthonormal basis vectors, and x and y are known as the components of \mathbf{r} with respect to the basis vectors $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$. We can then work out how to add vectors, calculate scalar products and so on working solely with these components. For instance, if we have two vectors \mathbf{r}_1 and \mathbf{r}_2 given by $\mathbf{r}_1 = x_1\hat{\mathbf{i}} + y_1\hat{\mathbf{j}}$ and $\mathbf{r}_2 = x_2\hat{\mathbf{i}} + y_2\hat{\mathbf{j}}$ then $\mathbf{r}_1 + \mathbf{r}_2 = (x_1 + x_2)\hat{\mathbf{i}} + (y_1 + y_2)\hat{\mathbf{j}}$.

It is important to note that while a vector $\hat{\mathbf{r}}$ is a unique geometrical object, there is no unique choice of basis vectors, and correspondingly the components of the vector will change depending on the choice of basis vectors. Thus we could equally well have chosen the basis vectors $\hat{\mathbf{i}}'$ and $\hat{\mathbf{j}}'$, as illustrated in Fig. (10.1) so that the *same* vector \mathbf{r} can be written

$$\mathbf{r} = x'\hat{\mathbf{i}}' + y'\hat{\mathbf{j}}' = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} \quad (10.1)$$

with $x' \neq x$ and $y' \neq y$.

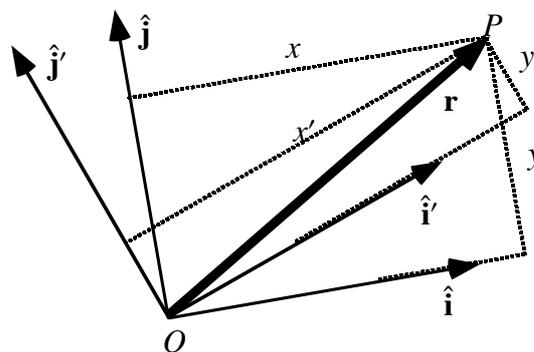


Figure 10.1: The position vector \mathbf{r} written as a linear combination of two different pairs of orthonormal basis vectors. Although the coordinates of \mathbf{r} are different with respect to the two sets of basis vectors, the vector \mathbf{r} remains the same.

Once a choice of basis vectors has been made, it proves to be very convenient to work solely with the coordinates. There is a useful notation by which this can be done. In this notation, the vector \mathbf{r} is written as

$$\mathbf{r} \doteq \begin{pmatrix} x \\ y \end{pmatrix} \quad (10.2)$$

what is known as a *column vector*. We then say that this column vector is a *representation* of the vector \mathbf{r} with respect to the basis vectors $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$. It is important to note that we do not say that \mathbf{r} equals the column vector, in fact it is not an equal sign that is used in Eq. (10.2), rather the symbol ‘ \doteq ’ is used, which is to be read as ‘is represented by’. The reason for this is that, as mentioned above, while the vector \mathbf{r} is a unique geometrical object, its components are not – they depend on the choice of basis vectors. We could have equally chosen basis vectors $\hat{\mathbf{i}}'$ and $\hat{\mathbf{j}}'$, and since the components x' and y' will be, in general, different from x and y , we end up with a different column vector representing the same vector:

$$\mathbf{r} \doteq \begin{pmatrix} x' \\ y' \end{pmatrix}. \quad (10.3)$$

i.e. two apparently different column vectors representing the *same* vector \mathbf{r} . Equivalently, if we had two column vectors with exactly the same numbers in the two positions, we could not conclude that they represent the same vector unless we were told that the basis vectors were the same in each case. Thus if there is any chance of ambiguity, we need to make it clear when we use the column vector notation, exactly what the basis vectors are. The terminology then is to say that the vector \mathbf{r} is given by the column vector in Eq. (10.2) in the $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}\}$ representation.

Once a choice of basis vectors has been settled on, and consistently used, we can proceed with vector calculations using the new notation. Thus, for instance, we can add two vectors:

$$\mathbf{r} = \mathbf{r}_1 + \mathbf{r}_2 \quad (10.4)$$

which becomes, using the $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}\}$ representation

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix}. \quad (10.5)$$

10.1.2 Row Vectors

The scalar product $(\mathbf{r}_1, \mathbf{r}_2) = \mathbf{r}_1 \cdot \mathbf{r}_2$ can be calculated by using the usual rule $\mathbf{r}_1 \cdot \mathbf{r}_2 = r_1 r_2 \cos \theta$, but it can also be expressed in terms of the components of \mathbf{r}_1 and \mathbf{r}_2 in, say, the $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}\}$ representation, though note that the same numerical result is obtained whatever representation is used. The result is simply

$$(\mathbf{r}_1, \mathbf{r}_2) = \mathbf{r}_1 \cdot \mathbf{r}_2 = x_1 x_2 + y_1 y_2. \quad (10.6)$$

At this point we note that, if we use the rules of matrix multiplication, this last result can be written

$$(\mathbf{r}_1, \mathbf{r}_2) = \mathbf{r}_1 \cdot \mathbf{r}_2 = \begin{pmatrix} x_1 & y_1 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \quad (10.7)$$

where we note the appearance of the column vector representing the vector \mathbf{r}_2 , but \mathbf{r}_1 , the first factor in the scalar product, has been represented by a *row vector*. If the components of \mathbf{r}_1 and \mathbf{r}_2 were complex, then we would write the inner product as

$$(\mathbf{r}_1, \mathbf{r}_2) = \mathbf{r}_1^* \cdot \mathbf{r}_2 = x_1^* x_2 + y_1^* y_2 = \begin{pmatrix} x_1^* & y_1^* \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \quad (10.8)$$

The use of a row vector to represent \mathbf{r}_1 can be looked on here as a convenience so that the rules of matrix multiplication can be applied, but there is a deeper significance to its use¹ that will become apparent when we look at the column and row vector representations of ket and bra vectors.

10.2 Representations of State Vectors and Operators

The procedure here is identical to that which was followed in the case of the position vector, i.e. we introduce a complete set of orthonormal basis states $\{|\varphi_n\rangle; n = 1, 2, \dots\}$ that span the state space of the quantum system, and then work with the components of the ket and bra vectors, and the operators. Of course, we now do not have the luxury of interpreting these basis vectors as representing physical directions in real space – rather they are abstract vectors in a multi-dimensional complex vector space, but much of what has been said above in connection with vectors in ordinary Euclidean space can be carried over to this more abstract situation.

10.2.1 Row and Column Vector Representations for Spin Half State Vectors

To set the scene, we will look at the particular case of spin half state vectors for which, as we have seen earlier, Sec. 8.2, an arbitrary state $|S\rangle$ can be written

$$|S\rangle = |-\rangle\langle -|S\rangle + |+\rangle\langle +|S\rangle,$$

i.e. the state $|S\rangle$ is expressed as a linear combination of the two basis states $|\pm\rangle$. We further saw that the ket vectors as $|+\rangle, |-\rangle$ could be put into direct correspondence with the (complex) unit vectors $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ respectively, and that the probability amplitudes

¹Effectively, what is going on is that corresponding to any vector \mathbf{r} represented by a column vector, there corresponds another vector \mathbf{r}_* known as its dual which is represented by a row vector. The original vector is the ‘physical’ vector while its dual is an abstract mathematical companion. The original vector and its dual belong to two different vector spaces.

$\langle \pm | S \rangle$ are the components of $|S\rangle$ in the ‘direction’ of the basis states $|\pm\rangle$. We can complete the analogy with the case of ordinary vectors by noting that we could then write this ket vector as a column vector, i.e.

$$|S\rangle \doteq \begin{pmatrix} \langle - | S \rangle \\ \langle + | S \rangle \end{pmatrix}. \quad (10.9)$$

If we pursue this line further, we can get an idea of how to interpret bra vectors. To do this, consider the more general probability amplitudes $\langle S' | S \rangle$. This we can write as

$$\langle S' | S \rangle = \langle S' | - \rangle \langle - | S \rangle + \langle S' | + \rangle \langle + | S \rangle. \quad (10.10)$$

If we now use

$$\langle \pm | S \rangle = \langle S | \pm \rangle^* \quad (10.11)$$

this becomes

$$\langle S' | S \rangle = \langle - | S' \rangle^* \langle - | S \rangle + \langle + | S' \rangle^* \langle + | S \rangle \quad (10.12)$$

which we can write as

$$\langle S' | S \rangle = \left(\langle - | S' \rangle^* \quad \langle + | S' \rangle^* \right) \begin{pmatrix} \langle - | S \rangle \\ \langle + | S \rangle \end{pmatrix}. \quad (10.13)$$

In other words, the bra vector $\langle S' |$ is represented by the row vector

$$\langle S' | \doteq \left(\langle - | S' \rangle^* \quad \langle + | S' \rangle^* \right). \quad (10.14)$$

This shows that a bra vector is more than just the ‘complex conjugate’ of a ket vector, since a row vector is not the same as a column vector.

We can now extend the idea to the more general situation of a state space of dimension $n > 2$.

10.2.2 Representation of Ket and Bra Vectors

In terms of the basis states $\{|\varphi_n\rangle; n = 1, 2, \dots\}$, an arbitrary state vector $|\psi\rangle$ can be written as

$$|\psi\rangle = \sum_n |\varphi_n\rangle \langle \varphi_n | \psi \rangle. \quad (10.15)$$

Let us now write

$$\langle \varphi_n | \psi \rangle = \psi_n. \quad (10.16)$$

We then have, by analogy with the position vector:

$$|\psi\rangle \doteq \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \vdots \end{pmatrix}. \quad (10.17)$$

This is a representation of $|\psi\rangle$ as a column vector with respect to the set of basis states $\{|\varphi_n\rangle; n = 1, 2, \dots\}$. In particular, the basis state $|\varphi_m\rangle$ will have components

$$(\varphi_m)_n = \langle\varphi_n|\varphi_m\rangle = \delta_{nm} \quad (10.18)$$

and so they will be represented by column vectors of the form

$$|\varphi_1\rangle \doteq \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \quad |\varphi_2\rangle \doteq \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} \quad \dots \quad (10.19)$$

i.e. the m^{th} component φ_{nm} of $|\varphi_n\rangle$ is zero except in the m^{th} position where $\varphi_{mm} = 1$.

Now form the inner product $\langle\chi|\psi\rangle$:

$$\langle\chi|\psi\rangle = \sum_n \langle\chi|\varphi_n\rangle \langle\varphi_n|\psi\rangle. \quad (10.20)$$

We know that $\langle\chi|\varphi_n\rangle = (\langle\varphi_n|\chi\rangle)^*$, and following on from the notation introduced above, we write $\chi_n = \langle\varphi_n|\chi\rangle$ so that

$$\langle\chi|\varphi_n\rangle = \chi_n^* \quad (10.21)$$

and hence

$$\langle\chi|\psi\rangle = \sum_n \chi_n^* \psi_n \quad (10.22)$$

which we can write as

$$\langle\chi|\psi\rangle = \begin{pmatrix} \chi_1^* & \chi_2^* & \chi_3^* & \dots \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \vdots \end{pmatrix} \quad (10.23)$$

which we evaluate by the usual rules of matrix multiplication. Note that here we have made the identification of the bra vector $\langle\chi|$ as a row vector:

$$\langle\chi| \doteq \begin{pmatrix} \chi_1^* & \chi_2^* & \chi_3^* & \dots \end{pmatrix} \quad (10.24)$$

with respect to the set of basis states $\{|\varphi_n\rangle; n = 1, 2, \dots\}$. This can be compared with the

representation of the ket vector $|\chi\rangle$ as a column vector:

$$|\chi\rangle \doteq \begin{pmatrix} \chi_1^* \\ \chi_2^* \\ \chi_3^* \\ \vdots \end{pmatrix}. \quad (10.25)$$

This difference in appearance of the representation of a bra and a ket vector, the first as a row vector, the second as a column vector, perhaps emphasizes the point made in Section 8.5.2 that the bra vectors form a vector space, the dual Hilbert space \mathcal{H}^* related to, but distinct from, the Hilbert space \mathcal{H} of the ket vectors. In a sense, a bra vector can be thought of as something akin to being the ‘complex conjugate’ of its corresponding ket vector.

10.2.3 Representation of Operators

Now turn to the operator equation

$$\hat{A}|\psi\rangle = |\phi\rangle \quad (10.26)$$

which we can write as

$$|\phi\rangle = \hat{A}|\psi\rangle = \hat{A} \sum_n |\varphi_n\rangle \langle \varphi_n | \psi \rangle = \sum_n \hat{A} |\varphi_n\rangle \langle \varphi_n | \psi \rangle. \quad (10.27)$$

Then

$$\langle \varphi_m | \phi \rangle = \sum_n \langle \varphi_m | \hat{A} | \varphi_n \rangle \langle \varphi_n | \psi \rangle \quad (10.28)$$

which we can write as

$$\phi_m = \sum_n A_{mn} \psi_n \quad (10.29)$$

where

$$A_{mn} = \langle \varphi_m | \hat{A} | \varphi_n \rangle. \quad (10.30)$$

We can write this as a matrix equation:

$$\begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & A_{13} & \dots \\ A_{21} & A_{22} & A_{23} & \dots \\ A_{31} & A_{32} & A_{33} & \dots \\ \vdots & \vdots & \vdots & \dots \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \vdots \end{pmatrix} \quad (10.31)$$

where the operator \hat{A} is represented by a *matrix*:

$$\hat{A} \doteq \begin{pmatrix} A_{11} & A_{12} & A_{13} & \dots \\ A_{21} & A_{22} & A_{23} & \dots \\ A_{31} & A_{32} & A_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \quad (10.32)$$

The quantities A_{mn} are known as the matrix elements of the operator \hat{A} with respect to the basis states $\{|\varphi_n\rangle; n = 1, 2, \dots\}$.

It is important to keep in mind that the column vectors, row vectors, and matrices above are constructed with respect to a particular set of basis states. If a different set of basis states are used, then the state vectors and operators remain the same, but the column or row vector, or matrix representing the state vector or operator respectively will change. Thus, to give any meaning to a row vector, or a column vector, or a matrix, it is essential that the basis states be known. An important part of quantum mechanics is the mathematical formalism that deals with transforming between different sets of basis states. However, we will not be looking at transformation theory here.

Ex 10.1 Consider two state vectors

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

where $|\pm\rangle$ are the usual base states for a spin half system. We want to represent these ket vectors as column vectors with respect to the set of basis states $\{|+\rangle, |-\rangle\}$. Firstly, we note that in the general development described above, we assumed that the basis states were named $|\varphi_1\rangle, |\varphi_2\rangle$ and so on. But here we are using a different way of labelling the basis states, which means we have a choice as to which of $|\pm\rangle$ we identify with $|\varphi_1\rangle$ and $|\varphi_2\rangle$. It makes no difference what we choose: we make the choice to suit ourselves, and provided we use it consistently then no problems should arise. Thus, here, we will choose $|\varphi_1\rangle = |+\rangle$ and $|\varphi_2\rangle = |-\rangle$. Thus we can write

$$|+\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |-\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

We can then express the states $|1\rangle$ and $|2\rangle$ in column vector notation as

$$|1\rangle \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} -i \\ 1 \end{pmatrix}$$

which can also be written as

$$|1\rangle \doteq -\frac{i}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The corresponding bra vectors are

$$\langle 1| = \frac{1}{\sqrt{2}}[\langle -| + i\langle +|] \quad \langle 2| = \frac{1}{\sqrt{2}}[\langle -| - i\langle +|]$$

or, as row vectors

$$\langle 1| \doteq (i \quad 1) \quad \text{and} \quad \langle 2| \doteq (-i \quad 1).$$

We can calculate inner products as follows:

$$\begin{aligned} \langle 1|2\rangle &= \frac{1}{2}(i \quad 1) \begin{pmatrix} i \\ 1 \end{pmatrix} \\ &= 0, \\ \langle 1|1\rangle &= \frac{1}{2}(i \quad 1) \begin{pmatrix} -i \\ 1 \end{pmatrix} \\ &= 1. \end{aligned}$$

and so on.

Ex 10.2 We can also look at the operator \hat{A} defined by

$$\hat{A}|\pm\rangle = \pm\frac{1}{2}i\hbar|\mp\rangle$$

which can be written out in matrix form as

$$\hat{A} \doteq \begin{pmatrix} \langle +|\hat{A}|+\rangle & \langle +|\hat{A}|-\rangle \\ \langle -|\hat{A}|+\rangle & \langle -|\hat{A}|-\rangle \end{pmatrix} = \begin{pmatrix} 0 & -\frac{1}{2}i\hbar \\ \frac{1}{2}i\hbar & 0 \end{pmatrix}$$

so that, for instance

$$\begin{aligned} \hat{A}|1\rangle &\doteq \begin{pmatrix} 0 & -\frac{1}{2}i\hbar \\ \frac{1}{2}i\hbar & 0 \end{pmatrix} \begin{pmatrix} -\frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \\ &= -\frac{1}{2}i\hbar \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{pmatrix} \\ &= \frac{1}{2}\hbar \begin{pmatrix} -\frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}. \end{aligned}$$

Thus we have $\hat{A}|1\rangle = \frac{1}{2}\hbar|1\rangle$, which incidentally shows that $|1\rangle$ is an eigenstate of \hat{A} .

Using the representations of bra vectors and operators, it is straightforward to see what the action of an operator on a bra vector is given by. Thus, we have:

$$\langle \psi | \hat{A} \doteq (\psi_1 \quad \psi_2 \quad \psi_3 \quad \dots) \begin{pmatrix} A_{11} & A_{12} & A_{13} & \dots \\ A_{21} & A_{22} & A_{23} & \dots \\ A_{31} & A_{32} & A_{33} & \dots \\ \vdots & \vdots & \vdots & \end{pmatrix}$$

$$= (\psi_1 A_{11} + \psi_2 A_{21} + \dots \quad \psi_2 A_{12} + \psi_2 A_{22} + \dots \quad \psi_1 A_{13} + \psi_2 A_{23} + \dots \quad \dots). \quad (10.33)$$

The final result can then be written in the corresponding bra vector notation if desired. This can be illustrated by example.

Ex 10.3 Evaluate $\langle 2 | \hat{A}$ using the representations of the bra vector $\langle 2 |$ and operator \hat{A} :

$$\begin{aligned} \langle 2 | \hat{A} &\doteq \frac{1}{\sqrt{2}} (-i \quad 1) \begin{pmatrix} 0 - \frac{1}{2} i \hbar \\ \frac{1}{2} i \hbar \end{pmatrix} \\ &= \frac{1}{\sqrt{2}} (\frac{1}{2} i \hbar \quad -\frac{1}{2} \hbar) \\ &= -\frac{1}{2} \hbar \cdot \frac{1}{\sqrt{2}} (-i \quad 1) \end{aligned}$$

which can be written as $\langle 2 | \hat{A} = -\frac{1}{2} \hbar \langle 2 |$.

10.2.4 Properties of Matrix Representations of Operators

Many of the properties of operators can be expressed in terms of the properties of their representative matrices. Most of these properties are straightforward, and will be presented below without comment.

Equality

Two operators are equal if their corresponding operator matrix elements are equal, i.e. $\hat{A} = \hat{B}$ if $A_{mn} = B_{mn}$.

Unit and Zero Operator

The unit operator $\hat{1}$ is the operator such that $\hat{1}|\psi\rangle = |\psi\rangle$ for all states $|\psi\rangle$. It has the matrix elements $\hat{1}_{mn} = \delta_{mn}$, i.e. the diagonal elements are all unity, and the off-diagonal elements are all zero. The unit operator has the same form in all representations, i.e. irrespective of the choice of basis states. The zero operator $\hat{0}$ is the operator such that $\hat{0}|\psi\rangle = 0$ for all states $|\psi\rangle$. Its matrix elements are all zero.

Addition of Operators

Given two operators \hat{A} and \hat{B} with matrix elements A_{mn} and B_{mn} , then the matrix elements of their sum $\hat{S} = \hat{A} + \hat{B}$ are given by

$$S_{mn} = A_{mn} + B_{mn}. \quad (10.34)$$

Multiplication by a Complex Number

If λ is a complex number, then the matrix elements of the operator $\hat{C} = \lambda\hat{A}$ are given by

$$C_{mn} = \lambda A_{mn}. \quad (10.35)$$

Product of Operators

Given two operators \hat{A} and \hat{B} with matrix elements A_{mn} and B_{mn} , then the matrix elements of their product $\hat{P} = \hat{A}\hat{B}$ are given by

$$P_{mn} = \sum_k A_{mk} B_{kn} \quad (10.36)$$

i.e. the usual rule for the multiplication of two matrices. Matrix multiplication, and hence operator multiplication, is not commutative, i.e. in general $\hat{A}\hat{B} \neq \hat{B}\hat{A}$. The difference, $\hat{A}\hat{B} - \hat{B}\hat{A}$, known as the commutator of \hat{A} and \hat{B} and written $[\hat{A}, \hat{B}]$, can be readily evaluated using the matrix representations of \hat{A} and \hat{B} .

Ex 10.4 Three operators $\hat{\sigma}_1$, $\hat{\sigma}_2$ and $\hat{\sigma}_3$, known as the Pauli spin matrices, that occur in the theory of spin half systems (and elsewhere) have the matrix representations with respect to the $\{|+\rangle, |-\rangle\}$ basis given by

$$\hat{\sigma}_1 \doteq \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \hat{\sigma}_2 \doteq \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \hat{\sigma}_3 \doteq \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The commutator $[\hat{\sigma}_1, \hat{\sigma}_2]$ can be readily evaluated using these matrices:

$$\begin{aligned} [\hat{\sigma}_1, \hat{\sigma}_2] &= \hat{\sigma}_1\hat{\sigma}_2 - \hat{\sigma}_2\hat{\sigma}_1 \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} - \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} - \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \\ &= 2i \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned}$$

The final matrix can be recognized as the representation of $\hat{\sigma}_3$, so that overall we have shown that

$$[\hat{\sigma}_1, \hat{\sigma}_2] = 2i\hat{\sigma}_3.$$

Cyclic permutation of the subscripts then gives the other two commutators.

Functions of Operators

If we have a function $f(x)$ which we can expand as a power series in x :

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots = \sum_{n=0}^{\infty} a_nx^n \quad (10.37)$$

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

$$f(\hat{A}) = a_0 + a_1\hat{A} + a_2\hat{A}^2 + \cdots = \sum_{n=0}^{\infty} a_n\hat{A}^n. \quad (10.38)$$

Once again, using the matrix representation of \hat{A} , it is possible, in certain cases, to work out what the matrix representation is of $f(\hat{A})$.

Ex 10.5 One of the most important functions of an operator that is encountered is the exponential function. To illustrate what this means, we will evaluate here the exponential function $\exp(i\phi\hat{\sigma}_3)$ where $\hat{\sigma}_1$ is one of the Pauli spin matrices introduced above, for which

$$\hat{\sigma}_1 \doteq \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

and ϕ is a real number. Using the power series expansion of the exponential function, we have

$$e^{i\phi\hat{\sigma}_1} = \sum_{n=0}^{\infty} \frac{i\phi^n}{n!} \hat{\sigma}_1^n.$$

It is useful to note that

$$\hat{\sigma}_3^2 \doteq \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

i.e. $\hat{\sigma}_1^2 = \hat{1}$, the identity operator. Thus we can always write

$$\hat{\sigma}_1^{2n} = \hat{1} \quad \hat{\sigma}_1^{2n+1} = \hat{\sigma}_1.$$

Thus, if we separate the infinite sum into two parts:

$$e^{i\phi\hat{\sigma}_1} = \sum_{n=0}^{\infty} \frac{(i\phi)^{2n}}{(2n)!} \hat{\sigma}_1^{2n} + \sum_{n=0}^{\infty} \frac{(i\phi)^{2n+1}}{(2n+1)!} \hat{\sigma}_1^{2n+1}$$

where the first sum is over all the even integers, and the second over all the odd integers, we get

$$\begin{aligned}
 e^{i\phi\hat{\sigma}_1} &= \hat{1} \sum_{n=0}^{\infty} (-1)^n \frac{\phi^{2n}}{(2n)!} + i\hat{\sigma}_1 \sum_{n=0}^{\infty} (-1)^n \frac{\phi^{2n+1}}{(2n+1)!} \\
 &= \cos \phi + i\hat{\sigma}_1 \sin \phi \\
 &\doteq \begin{pmatrix} \cos \phi & 0 \\ 0 & \cos \phi \end{pmatrix} + \begin{pmatrix} 0 & i \sin \phi \\ i \sin \phi & 0 \end{pmatrix} \\
 &= \begin{pmatrix} \cos \phi & i \sin \phi \\ i \sin \phi & \cos \phi \end{pmatrix}.
 \end{aligned}$$

Inverse of an Operator

Finding the inverse of an operator, given its matrix representation, amounts to finding the inverse of the matrix, provided, of course, that the matrix has an inverse.

Ex 10.6 The inverse of $\exp(i\phi\hat{\sigma}_1)$ can be found by taking the inverse of its representative matrix:

$$(e^{i\phi\hat{\sigma}_1})^{-1} \doteq \begin{pmatrix} \cos \phi & i \sin \phi \\ i \sin \phi & \cos \phi \end{pmatrix}^{-1} = \begin{pmatrix} \cos \phi & -i \sin \phi \\ -i \sin \phi & \cos \phi \end{pmatrix}.$$

This inverse can be recognized as being just

$$\begin{pmatrix} \cos \phi & -i \sin \phi \\ -i \sin \phi & \cos \phi \end{pmatrix} = \begin{pmatrix} \cos(-\phi) & i \sin(-\phi) \\ i \sin(-\phi) & \cos(-\phi) \end{pmatrix}$$

which means that

$$(e^{i\phi\hat{\sigma}_1})^{-1} = e^{-i\phi\hat{\sigma}_1}$$

a perhaps unsurprising result in that it is a particular case of the fact that the inverse of $\exp(\hat{A})$ is just $\exp(-\hat{A})$, exactly as is the case for the exponential function of a complex variable.

10.2.5 Eigenvectors and Eigenvalues

Operators act on states to map them into other states. Amongst the possible outcomes of the action of an operator on a state is to map the state into a multiple of itself:

$$\hat{A}|\phi\rangle = a_\phi|\phi\rangle \tag{10.39}$$

where $|\phi\rangle$ is, in general, a complex number. The state $|\phi\rangle$ is then said to be an eigenstate or eigenket of the operator \hat{A} with a_ϕ the associated eigenvalue. The fact that operators can possess eigenstates might be thought of as a mathematical fact incidental to the physical content of quantum mechanics, but it turns out that the opposite is the case: the eigenstates and eigenvalues of various kinds of operators are essential parts of the physical interpretation of the quantum theory, and hence warrant close study. Notationally, it is often useful to use the eigenvalue associated with an eigenstate to label the eigenvector, i.e. the notation

$$\hat{A}|a\rangle = a|a\rangle. \quad (10.40)$$

This notation, or minor variations of it, 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 $\hat{A}|\phi\rangle = a_\phi|\phi\rangle$. Written out in terms of the matrix representations of the operator with respect to some set of orthonormal basis vectors $\{|\varphi\rangle; n = 1, 2, \dots\}$, this eigenvalue equation is

$$\begin{pmatrix} A_{11} & A_{12} & \dots \\ A_{21} & A_{22} & \dots \\ \vdots & \vdots & \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \end{pmatrix} = a \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \end{pmatrix}. \quad (10.41)$$

This expression is equivalent to a set of simultaneous, homogeneous, linear equations:

$$\begin{pmatrix} A_{11} - a & A_{12} & \dots \\ A_{21} & A_{22} - a & \dots \\ \vdots & \vdots & \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \end{pmatrix} = 0 \quad (10.42)$$

which have to be solved for the possible values for a , and the associated values for the components ϕ_1, ϕ_2, \dots of the eigenvectors. The procedure is standard. The determinant of coefficients must vanish in order to get non-trivial solutions for the components ϕ_1, ϕ_2, \dots :

$$\begin{vmatrix} A_{11} - a & A_{12} & \dots \\ A_{21} & A_{22} - a & \dots \\ \vdots & \vdots & \end{vmatrix} = 0 \quad (10.43)$$

which yields an equation known as the secular equation, or characteristic equation, that has to be solved to give the possible values of the eigenvalues a . Once these are known, they have to be resubstituted into Eq. (10.41) and the components ϕ_1, ϕ_2, \dots of the associated eigenvectors determined. The details of how this is done properly belongs to a text on linear algebra and will not be considered any further here, except to say that the eigenvectors are typically determined up to an unknown multiplicative constant. This constant is usually fixed by the requirement that these eigenvectors be normalized to unity. In the

case of repeated eigenvalues, i.e. when the characteristic polynomial has multiple roots (otherwise known as degenerate eigenvalues), the determination of the eigenvectors is made more complicated still. Once again, issues connected with these kinds of situations will not be considered here.

In general, for a state space of finite dimension, it is found that the operator \hat{A} will have one or more discrete eigenvalues a_1, a_2, \dots and associated eigenvectors $|a_1\rangle, |a_2\rangle, \dots$. 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*.

For the present we will be confining our attention to operators that have discrete eigenvalue spectra. Modifications needed to handle continuous eigenvalues will be introduced later.

10.2.6 Hermitean Operators

Apart from certain calculational advantages, the representation of operators as matrices makes it possible to introduce in a direct fashion Hermitean operators, already considered in a more abstract way in Section 9.3.1, which have a central role to play in the physical interpretation of quantum mechanics.

To begin with, suppose we have an operator \hat{A} with matrix elements A_{mn} with respect to a set of orthonormal basis states $\{|\varphi_n\rangle; n = 1, 2, \dots\}$. From the matrix representing this operator, we can construct a new operator by taking the transpose and complex conjugate of the original matrix:

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} & \dots \\ A_{21} & A_{22} & A_{23} & \dots \\ A_{31} & A_{32} & A_{33} & \dots \\ \vdots & \vdots & \vdots & \end{pmatrix} \rightarrow \begin{pmatrix} A_{11}^* & A_{21}^* & A_{31}^* & \dots \\ A_{12}^* & A_{22}^* & A_{32}^* & \dots \\ A_{13}^* & A_{23}^* & A_{33}^* & \dots \\ \vdots & \vdots & \vdots & \end{pmatrix}. \quad (10.44)$$

The new matrix will represent a new operator which is obviously related to \hat{A} , which we will call \hat{A}^\dagger , i.e.

$$\hat{A}^\dagger \doteq \begin{pmatrix} (A^\dagger)_{11} & (A^\dagger)_{12} & (A^\dagger)_{13} & \dots \\ (A^\dagger)_{21} & (A^\dagger)_{22} & (A^\dagger)_{23} & \dots \\ (A^\dagger)_{31} & (A^\dagger)_{32} & (A^\dagger)_{33} & \dots \\ \vdots & \vdots & \vdots & \end{pmatrix} = \begin{pmatrix} A_{11}^* & A_{21}^* & A_{31}^* & \dots \\ A_{12}^* & A_{22}^* & A_{32}^* & \dots \\ A_{13}^* & A_{23}^* & A_{33}^* & \dots \\ \vdots & \vdots & \vdots & \end{pmatrix}. \quad (10.45)$$

i.e.

$$\langle \varphi_m | \hat{A}^\dagger | \varphi_n \rangle = (\langle \varphi_n | \hat{A} | \varphi_m \rangle)^*. \quad (10.46)$$

The new operator that we have created, \hat{A}^\dagger , can be recognized as the Hermitean adjoint of \hat{A} . The Hermitean adjoint has a useful property which we can most readily see if we use the matrix representation of the operator equation

$$\hat{A}|\psi\rangle = |\phi\rangle \quad (10.47)$$

which we earlier showed could be written as

$$\phi_m = \sum_n \langle \varphi_m | \hat{A} | \varphi_n \rangle \psi_n. \quad (10.48)$$

If we now take the complex conjugate of this expression we find

$$\phi_m^* = \sum_n \psi_n^* \langle \varphi_n | \hat{A}^\dagger | \varphi_m \rangle \quad (10.49)$$

which we can write in row vector form as

$$(\phi_1^* \quad \phi_2^* \quad \dots) = (\psi_1^* \quad \psi_2^* \quad \dots) \begin{pmatrix} (A^\dagger)_{11} & (A^\dagger)_{12} & \dots \\ (A^\dagger)_{21} & (A^\dagger)_{22} & \dots \\ \vdots & \vdots & \end{pmatrix} \quad (10.50)$$

which is the matrix version of

$$\langle \phi | = \langle \psi | \hat{A}^\dagger. \quad (10.51)$$

In other words we have shown that if $\hat{A}|\psi\rangle = |\phi\rangle$, then $\langle \psi | \hat{A}^\dagger = \langle \phi |$, a result that we used earlier to motivate the definition of the Hermitean adjoint in the first place. Thus, there are two ways to approach this concept: either through a general abstract argument, or in terms of matrix representations of an operator.

Ex 10.7 Consider the operator \hat{A} which has the representation in some basis

$$\hat{A} \doteq \begin{pmatrix} 1 & i \\ 0 & -1 \end{pmatrix}.$$

Then

$$\hat{A}^\dagger \doteq \begin{pmatrix} 1 & 0 \\ -i & -1 \end{pmatrix}.$$

To be noticed in this example is that $\hat{A} \neq \hat{A}^\dagger$.

Ex 10.8 Now consider the operator

$$\hat{A} \doteq \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

Then

$$\hat{A}^\dagger \doteq \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

i.e. $\hat{A} = \hat{A}^\dagger$.

This is an example of a situation in which \hat{A} and \hat{A}^\dagger are identical. In this case, the operator is said to be selfadjoint, or Hermitean. Once again, we encountered this idea in a more general context in Section 9.4.2, where it was remarked that Hermitean operators have a number of very important properties which leads to their playing a central role in the physical interpretation of quantum mechanics. These properties of Hermitean operators lead to the identification of Hermitean operators as representing the physically observable properties of a physical system, in a way that will be discussed in the next Chapter.