Chapter 8

Vector Spaces in Quantum Mechanics

We have seen in the previous Chapter that there is a sense in which the state of a quantum system can be thought of as being made up of other possible states. The aim here is to use the example of the Stern-Gerlach experiment to develop this idea further, and to show that the states of a quantum system can be represented by vectors in a complex vector space. To begin, we review some of the basic ideas of vectors, using the example of vectors in real two dimensional space.

8.1 Vectors in Two Dimensional Space

Below is a summary of the important properties of vectors in physical space based on their interpretation as mathematical objects that have both magnitude and direction. Position, velocity, force and so on are examples of such vectors. The intention is not to give a complete discussion, but to highlight a number of important properties of such vectors that have analogues in the case of quantum states, including the property that two vectors can be combined to produce another vector, and that 'how much' of one vector is contained in another can be measured via the inner product of two vectors.

8.1.1 Linear Combinations of Vectors – Vector Addition

Consider two non-collinear vectors $\hat{\mathbf{v}}_1$ and $\hat{\mathbf{v}}_2$, as illustrated in Fig. (8.1). Multiples of these pair of vectors can be added together in the fashion illustrated in Fig. (8.1) to form another vector. Conversely, any other vector $\mathbf{v}$ can be expressed in terms of $\hat{\mathbf{v}}_1$ and $\hat{\mathbf{v}}_2$ using appropriate values for the components a and b of $\mathbf{v}$, i.e.

$$\mathbf{v} = a\hat{\mathbf{v}}_1 + b\hat{\mathbf{v}}_2. \tag{8.1}$$

The right hand side of this equation is known as a *linear combination* of the vectors $\hat{\mathbf{v}}_1$ and $\hat{\mathbf{v}}_2$. The particular points to take away from this is that combining vectors produces other vectors, analogous to our observation above that combining states produces other states, and that the components a and b are a measure of how much of $\hat{\mathbf{v}}_1$ and $\hat{\mathbf{v}}_2$ respectively go towards making up the vector $\mathbf{v}$.

These vectors are *real* vectors in that the coefficients a and b are real numbers. We can readily generalize the ideas above by allowing a and b to be complex numbers. This makes it difficult to draw figures such as Fig. (8.1), but the idea can be still retained that the absolute values $|a|$ and $|b|$ will then represent the extent to which $\mathbf{v}$ is made up of the two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$.

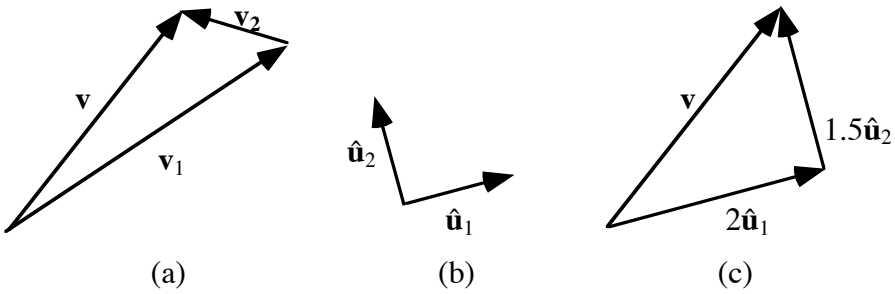

Figure 8.1: Examples of vector addition. (a) Two arbitrary basis vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ combining to give $\mathbf{v}$; (b) A pair of orthonormal basis vectors; (c) The vector $\mathbf{v}$ expressed as a linear combination of the basis vectors $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$: $\mathbf{v} = 2\hat{\mathbf{u}}_1 + 1.5\hat{\mathbf{u}}_2$.

The two vectors $\hat{\mathbf{v}}_1$ and $\hat{\mathbf{v}}_2$ introduced above are arbitrary except insofar as they are not collinear. They are known as *basis vectors* in the sense that any vector can be written as a linear combination of them. There is effectively an infinite number of choices for the basis vectors, and in fact it is possible to choose three or more vectors to be basis vectors. But the minimum number is two, if we wish to be able to describe any vector in the plane as a linear combination of basis vectors. The collection of all the vectors that can be constructed by taking linear combinations of these basis vectors using any complex numbers a and b as components is known as a complex vector space, and since two basis vectors are needed, the vector space is said to be of dimension two.

This vector space possess more structure than that implied by simply forming various linear combinations. Vectors that can be drawn in a plane, as in Fig. (8.1), i.e. for which the coefficients a and b are real, can be clearly seen to have different lengths and relative orientations. These properties of vectors are encompassed in the definition of the inner, scalar or dot product of pairs of vectors.

8.1.2 Inner or Scalar Products

The inner or scalar product of two *real* vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ is defined by

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = v_1 v_2 \cos\theta \tag{8.2}$$

where v_1 and v_2 are the lengths of $\mathbf{v}_1$ and $\mathbf{v}_2$ respectively, and θ is the angle between them. Using the idea of an inner or scalar product $\mathbf{v}_1 \cdot \mathbf{v}_2$ it is possible to introduce a particularly useful pair of basis vectors. To this end, consider two vectors $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ that satisfy

$$\hat{\mathbf{u}}_1 \cdot \hat{\mathbf{u}}_1 = \hat{\mathbf{u}}_2 \cdot \hat{\mathbf{u}}_2 = 1 \tag{8.3}$$

i.e. they have unit length, hence they are known as unit vectors, and

$$\hat{\mathbf{u}}_1 \cdot \hat{\mathbf{u}}_2 = 0 \tag{8.4}$$

i.e. they are orthogonal. This pair of vectors $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ is said to be orthonormal. As before, they form a basis in that multiples of each vector can be added together in the fashion illustrated in Fig. (8.1) to form another vector. Conversely, any vector $\mathbf{v}$ can be expressed using appropriate values for the components of $\mathbf{v}$, i.e.

$$\mathbf{v} = a\hat{\mathbf{u}}_1 + b\hat{\mathbf{u}}_2. \tag{8.5}$$

The components a and b, which represent 'how much' of the vector $\mathbf{v}$ is made up of the vectors $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ are given by

$$a = \hat{\mathbf{u}}_1 \cdot \mathbf{v} \quad \text{and} \quad b = \hat{\mathbf{u}}_2 \cdot \mathbf{v} \tag{8.6}$$

A well known example of a vector expressed as a linear combination of a pair of unit vectors is given by the position vector $\mathbf{r}$ written with respect to the unit vectors $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$:

$$\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}}. \tag{8.7}$$

Once again, if we allow for the possibility of complex vectors, then the definition of the inner product is changed to allow for this. Thus the inner product of two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ is now $\mathbf{v}_1^* \cdot \mathbf{v}_2$, and orthonormal vectors satisfy

$$\begin{aligned} \mathbf{u}_1^* \cdot \mathbf{u}_1 &= \mathbf{u}_2^* \cdot \mathbf{u}_2 = 1 \\ \mathbf{u}_1^* \cdot \mathbf{u}_2 &= \mathbf{u}_2^* \cdot \mathbf{u}_1 = 0. \end{aligned} \tag{8.8}$$

Complex basis vectors are not exotic mathematical entities. A simple example is

$$\begin{aligned} \mathbf{u}_1 &= \frac{3\mathbf{i} + 4i\mathbf{j}}{5} \\ \mathbf{u}_2 &= \frac{4\mathbf{i} - 3i\mathbf{j}}{5} \end{aligned} \tag{8.9}$$

though it is true to say that they are not often encountered in mechanics, except at an advanced level, but are, as we shall see, routinely found in quantum mechanics. Because of that, it is convenient to introduce a new notation for inner product (one that is commonly used in pure mathematics) which is to write the inner product of two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ as $(\mathbf{v}_1, \mathbf{v}_2)$, so that we would have here

$$(\mathbf{v}_1, \mathbf{v}_2) = \mathbf{v}_1^* \cdot \mathbf{v}_2. \tag{8.10}$$

The value of this notation is simply that it is more general and does not tie us to the geometrical notion of inner product that we are used to using with position vectors, velocity vectors and the like.

8.2 Spin Half Quantum States as Vectors

We now need to examine the way in which the quantum states for a spin half system can be seen to fit in with the idea of being considered as vectors. To see how this comes about, we

will show that there is a perfect analogy between Eqs. (7.43) and (7.44) and corresponding relationships for ordinary (complex) vectors. Returning to the Stern-Gerlach example discussed in the preceding Chapter we obtained there an expression Eq. (7.44)

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

for the state of a spin half atom expressed in terms of the states $|\pm\rangle$, which are states for which the atom has a z component of spin equal to $\pm\frac{1}{2}\hbar$, and $\langle\pm|S\rangle$ are probability amplitudes (complex numbers) whose magnitude in some sense tells us 'how much' of the states $|\pm\rangle$ are to be found in the state $|S\rangle$. This result was obtained by 'cancelling' the common factor '$\langle S'|$' from the result

$$\langle S'|S\rangle = \langle S'|+\rangle\langle +|S\rangle + \langle S'|-\rangle\langle -|S\rangle.$$

What we should bear in mind here is that we can recover this relationship between probability amplitudes by reintroducing '$\langle S'|$' into Eq. (7.44) for any chosen final state, yielding an expression for the probability amplitudes as needed. Thus, as has been said before, Eq. (7.44) effectively represents a 'template' into which we insert the appropriate information in order to recover the required probability amplitudes. We can also note that there is nothing sacred about choosing to cancel the the common factor $|S\rangle$ – we could equally as well cancel the factor $|S\rangle$, yielding

$$\langle S'| = \langle S'|+\rangle\langle +| + \langle S'|-\rangle\langle -|. \tag{8.11}$$

Having carried out this cancellation procedure, what has reappeared is the state of a quantum system i.e. $|S\rangle$ which was introduced earlier in a different context, specifically as being nothing more than a way of writing down all that we knew about the state of a quantum system. There, the notation had no mathematical significance, but in the manner in which it appears here, it seems to have acquired a mathematical meaning of some kind. The aim is to see what this meaning might be, and in doing so, we will show that the expression for $|S\rangle$ has many of the properties that we associate with expressing a vector as a sum of its components.

We begin by considering the probability amplitudes $\langle S'|S\rangle$ themselves. These are complex numbers in general for arbitrary spin directions, (but they were real in the particular Stern-Gerlach example used above), such that their modulus squared $|\langle S'|S\rangle|^2$ is the probability $P(S'|S)$ of observing the spin to be in the state $|S'\rangle$ given that it was in the state $|S\rangle$. In particular, $\langle S|S\rangle$ is the probability amplitude of observing the spin to be in the state $|S\rangle$ given that it was in the state $|S\rangle$. This will have to be unity, i.e. $P(S|S) = |\langle S|S\rangle|^2 = 1$. Thus we can conclude that

$$\langle S|S\rangle = e^{i\eta} \tag{8.12}$$

where η is an arbitrary phase. It turns out that this phase always cancels out in any calculation of observable quantities, so it is conventionally set to zero, and hence

$$\langle S|S\rangle = 1. \tag{8.13}$$

The state $|S\rangle$ is said to be normalized to unity. As a particular case, this last result implies that

$$\langle +|+\rangle = 1. \tag{8.14}$$

We can now consider the probability amplitude $\langle +|S\rangle$ obtained by replacing S' by $+$ in the above expression for $\langle S'|S\rangle$:

$$\langle +|S\rangle = \langle +|+\rangle\langle +|S\rangle + \langle +|-\rangle\langle -|S\rangle. \tag{8.15}$$

We have seen that we can put $\langle +|+\rangle = 1$, so we have

$$\langle +|-\rangle\langle -|S\rangle = 0 \tag{8.16}$$

which has to be true no matter what the state $|S\rangle$ happens to be, i.e. no matter what value the probability amplitude $\langle -|S\rangle$ is. Thus we conclude that

$$\langle +|-\rangle = 0. \tag{8.17}$$

Similarly we can show that

$$\langle -|-\rangle = 1 \quad \text{and} \quad \langle -|+\rangle = 0. \tag{8.18}$$

Thus we can set up a comparison:

$$\begin{aligned} \langle +|+\rangle = 1 \quad &\longleftrightarrow \quad \hat{\mathbf{u}}_1^* \cdot \hat{\mathbf{u}}_1 = 1 \\ \langle +|-\rangle = 0 \quad &\longleftrightarrow \quad \hat{\mathbf{u}}_2^* \cdot \hat{\mathbf{u}}_1 = 0 \\ \langle -|-\rangle = 1 \quad &\longleftrightarrow \quad \hat{\mathbf{u}}_2^* \cdot \hat{\mathbf{u}}_2 = 1 \\ \langle -|+\rangle = 0 \quad &\longleftrightarrow \quad \hat{\mathbf{u}}_2^* \cdot \hat{\mathbf{u}}_1 = 0 \end{aligned} \tag{8.19}$$

where we have chosen to make the comparison between the probability amplitudes and the inner product of complex unit vectors as we are dealing with probability amplitudes that are, in eneral, complex numbers. This comparison implies the following correspondences:

$$\begin{aligned} |+\rangle &\longleftrightarrow \hat{\mathbf{u}}_1 \quad |-\rangle \quad \longleftrightarrow \hat{\mathbf{u}}_2 \\ \langle +| &\longleftrightarrow \hat{\mathbf{u}}_1^* \quad \langle -| \quad \longleftrightarrow \hat{\mathbf{u}}_2^*. \end{aligned} \tag{8.20}$$

We know that $\langle +|S\rangle$ and $\langle -|S\rangle$ are both just complex numbers, so call them a and b respectively. If we now write

$$|S\rangle = a|+\rangle + b|-\rangle \tag{8.21}$$

we establish a perfect correspondence with the expression

$$\mathbf{v} = a\,\hat{\mathbf{u}}_1 + b\,\hat{\mathbf{u}}_2. \tag{8.22}$$

On the basis of this result, we are then tempted to interpret the ket $|S\rangle$ as a vector expressed as a linear combination of two orthonormal basis vectors $|\pm\rangle$. We can push the analogy further if we once again use the fact that $\langle S|S\rangle = 1$, so that

$$\langle S|S\rangle = 1 = \langle S|-\rangle\langle -|S\rangle + \langle S|+\rangle\langle +|S\rangle \tag{8.23}$$

On the other hand, the total probability of observing the system in either of the states $|\pm\rangle$ must add up to unity, which means that

$$P(+|S) + P(-|S) = |\langle +|S\rangle|^2 + |\langle -|S\rangle|^2 = 1. \tag{8.24}$$

By comparing the last two equations, and noting that

$$|\langle \pm|S\rangle|^2 = \langle \pm|S\rangle\langle \pm|S\rangle^* \tag{8.25}$$

we conclude that

$$\langle \pm | S \rangle = \langle S | \pm \rangle^*. \tag{8.26}$$

If we now consider

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

and use the result, Eq. (8.26), $\langle \pm | S' \rangle = \langle S' | \pm \rangle^*$, we can write this as

$$\langle S' | = \langle + | S' \rangle^* \langle + | + \langle - | S' \rangle^* \langle - | \tag{8.27}$$

or, expressed in terms of $a' = \langle + | S' \rangle$ and $b' = \langle - | S' \rangle$, we have

$$\langle S' | = a'^* \langle + | + b'^* \langle - | \tag{8.28}$$

which has a perfect correspondence with an ordinary vector $\hat{\mathbf{v}}'$ in the form

$$\mathbf{v}'^* = a'^* \hat{\mathbf{u}}_1^* + b'^* \hat{\mathbf{u}}_2^*. \tag{8.29}$$

So the bra $\langle S' |$ is itself a vector, a bra vector, which can be thought of as being just the complex conjugate of the corresponding ket vector $| S' \rangle$. But while it is occasionally useful to think this way, it is not strictly true mathematically, and this way of viewing a bra vector will not be employed here. Instead, as will be shown shortly, an interpretation of ket vectors as column vectors leads to the interpretation of bra vectors as row vectors. A more mathematical standpoint also leads to interpretation of bra vectors as 'linear functionals', that is, a bra is a mathematical operator that acts on a ket vector to produce a complex number.

Finally, to complete the correspondence, we note that the probability amplitude $\langle S' | S \rangle$ can be written

$$\langle S' | S \rangle = a'^* a + b'^* b \tag{8.30}$$

which can be compared with the inner product $\mathbf{v}'^* \mathbf{v}$, or written in the more formal notation:

$$(\mathbf{v}', \mathbf{v}) = a'^* a + b'^* b \tag{8.31}$$

which tells us that the probability amplitude can be considered as being simply the inner product of the two vectors $| S' \rangle$ and $| S \rangle$, i.e.

$$\langle S' | S \rangle = (| S' \rangle, | S \rangle). \tag{8.32}$$

In other words, we have a perfect analogy between the two dimensional complex vector space formed by linear combinations of the unit vectors $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ discussed in Section 8.1 and a complex vector space consisting of all the linear combinations of the states $| \pm \rangle$. The ket vectors $| \pm \rangle$ are referred to as *basis states*, analogous to $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ being referred to as basis vectors.

Different spin states can be constructed by forming linear combinations $| S \rangle = a | + \rangle + b | - \rangle$ of these basis states, with a and b being allowed to range over all the complex numbers, though we have only been looking at linear combinations with real coefficients. By limiting a and b to be real numbers, we are constructing states whose measured component of spin all lie in the same plane, which, with the system of axes we have been using here, is the XZ plane. If the coefficients a and b are complex, the state $| S \rangle$ represents a state in which the measured component $S = \mathbf{S} \cdot \mathbf{n}$ is along a direction $\hat{\mathbf{n}}$ that points out of this

plane, though we will not attempt to prove this here (see Eq. (8.46)). So, any linear combination in which a and b are any complex numbers also defines a possible spin state of the spin half system. Thus all the possible linear combinations of $|\pm\rangle$, i.e. combinations of the form $a|+\rangle + b|-\rangle$ where a and b are complex numbers form a *complex vector space* known as the *state space* of the system.

The quantum state vectors can also be 'multiplied' together – the inner product of the two vectors $|S'\rangle$ and $|S\rangle$ is just the probability amplitude $\langle S'|S\rangle$. In particular, the basis states are normalized to unity, i.e. they are unit vectors, and they are orthogonal to each other, i.e. they form a pair of *orthonormal basis states*.

The terminology often adopted is to say that the state vector $|S\rangle = a|+\rangle + b|-\rangle$ is a 'linear superposition' of the two states $|\pm\rangle$. The probability amplitudes $a = \langle +|S\rangle$ and $b = \langle -|S\rangle$ represent 'how much' of the states $|\pm\rangle$ are contained within the state $|S\rangle$ to the extent that $|\langle \pm|S\rangle|^2$ is the probability of the z component of spin being found to have the values $\pm\frac{1}{2}\hbar$. One difference between ordinary vectors and quantum state vectors is the importance of the 'normalization condition', i.e. the requirement that $\langle S|S\rangle = 1$, which must hold true given the interpretation of the inner product as a probability amplitude. But how can this be reconciled with the statement above that *any* linear combination of the basis states is a possible state of the system? How can a state vector such as $|\widetilde{S}\rangle = |+\rangle + |-\rangle$ which has the property

$$\langle \widetilde{S}|\widetilde{S}\rangle = 2 \tag{8.33}$$

be a physically acceptable state vector as it seems to be saying that the probability of finding the system in the state $|\widetilde{S}\rangle$ given that it is in the state $|\widetilde{S}\rangle$ is 4, which does not make sense. But, if we define a new vector $|S\rangle$ by

$$|S\rangle = \frac{|\widetilde{S}\rangle}{\sqrt{\langle \widetilde{S}|\widetilde{S}\rangle}} = \frac{1}{\sqrt{2}}|\widetilde{S}\rangle \tag{8.34}$$

then automatically $|S\rangle$ will have the required normalization property – it is said to be normalized to unity. So, rather than abandoning giving a physical interpretation of state vectors which are not normalized to unity, the approach adopted is that we can multiply any state vector by any factor and say that it still represents the same *physical* state, i.e. $|S\rangle$ and $|\widetilde{S}\rangle = a|S\rangle$, where a is any complex number, represent the same physical state. However, it is only the normalized state $|S\rangle$ that should be used in any calculations in order to be certain that probability is properly accounted for.

This can be summarized in a table:

	Classical vector	Quantum state vector for spin half system
Basis Vectors	$\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2$	$\|+\rangle, \|-\rangle$
Inner Product	$(\mathbf{v}_1, \mathbf{v}_2) = \mathbf{v}_1^* \cdot \mathbf{v}_2$	$(\|S'\rangle, \|S\rangle) = \langle S'\|S\rangle$
Orthonormality	$\hat{\mathbf{u}}_1^* \cdot \hat{\mathbf{u}}_1 = \hat{\mathbf{u}}_2^* \cdot \hat{\mathbf{u}}_2 = 1$ $\hat{\mathbf{u}}_1^* \cdot \hat{\mathbf{u}}_2 = \hat{\mathbf{u}}_2^* \cdot \hat{\mathbf{u}}_1 = 0$	$\langle +\|+\rangle = \langle -\|-\rangle = 1$ $\langle +\|-\rangle = \langle -\|+\rangle = 0$
Linear combination	$\hat{\mathbf{v}} = a\hat{\mathbf{u}}_1 + b\hat{\mathbf{u}}_2$ a and b complex numbers	$\|S\rangle = a\|+\rangle + b\|-\rangle$ a and b complex numbers
Normalisation	—	$\langle S\|S\rangle = 1$

A More Detailed Analysis One of the important properties of vectors is that two or more of them can be combined as a 'linear combination' to produce a third. If we are to consider quantum states as vectors, then this very basic property must also be possessed by quantum states. In the above, we have not really shown that any linear combination of the basis states $|\pm\rangle$, does indeed represent a possible state of the system, even if we restrict ourselves to the case of the measuring spin components in the XZ plane. But a more detailed argument than that just presented can be used to strengthen this conclusion. To see how this comes about, we return to the above expression Eq. (7.18) for the probability amplitude $\langle S_f = \frac{1}{2}\hbar|S_i = \frac{1}{2}\hbar\rangle$:

$$\begin{aligned}\langle S_f = \tfrac{1}{2}\hbar|S_i = \tfrac{1}{2}\hbar\rangle =& \langle S_f = \tfrac{1}{2}\hbar|S_I = \tfrac{1}{2}\hbar\rangle\langle S_I = \tfrac{1}{2}\hbar|S_i = \tfrac{1}{2}\hbar\rangle \\ &+ \langle S_f = \tfrac{1}{2}\hbar|S_I = -\tfrac{1}{2}\hbar\rangle\langle S_I = -\tfrac{1}{2}\hbar|S_i = \tfrac{1}{2}\hbar\rangle.\end{aligned} \tag{8.35}$$

We can 'cancel' $\langle S_f = \frac{1}{2}\hbar|$ from this expression and write

$$|S_i = \tfrac{1}{2}\hbar\rangle = |S_I = \tfrac{1}{2}\hbar\rangle\langle S_I = \tfrac{1}{2}\hbar|S_i = \tfrac{1}{2}\hbar\rangle + |S_I = -\tfrac{1}{2}\hbar\rangle\langle S_I = -\tfrac{1}{2}\hbar|S_i = \tfrac{1}{2}\hbar\rangle \tag{8.36}$$

In Section 7.3, Eq. (7.33), the quantity $\langle S_f = \frac{1}{2}\hbar|S_i = \frac{1}{2}\hbar\rangle$ was shown to be given by

$$\langle S_f = \tfrac{1}{2}\hbar|S_i = \tfrac{1}{2}\hbar\rangle = \cos[(\theta_f - \theta_i)/2]$$

with the phase factor $\exp(i\Phi)$ put equal to unity, and with $\hat{\mathbf{n}}$ and $\hat{\mathbf{m}}$, the directions of the magnetic fields, confined to the XZ plane. The probability amplitudes for the system to pass through the intermediate states $|S_I = \pm\frac{1}{2}\hbar\rangle$, that is, $\langle S_I = \pm\frac{1}{2}\hbar|S_i = \frac{1}{2}\hbar\rangle$ are likewise given by

$$\begin{aligned}\langle S_I = \tfrac{1}{2}\hbar|S_i = \tfrac{1}{2}\hbar\rangle &= \cos[\tfrac{1}{2}(\theta_I - \theta_i)] \\ \langle S_I = -\tfrac{1}{2}\hbar|S_i = \tfrac{1}{2}\hbar\rangle &= \cos[\tfrac{1}{2}(\theta_I + \pi - \theta_i)] = -\sin[\tfrac{1}{2}(\theta_I - \theta_i)].\end{aligned} \tag{8.37}$$

so that Eq. (8.36) becomes

$$|S_i = \tfrac{1}{2}\hbar\rangle = \cos[\tfrac{1}{2}(\theta_I - \theta_i)]|S_I = +\tfrac{1}{2}\hbar\rangle - \sin[\tfrac{1}{2}(\theta_I - \theta_i)]|S_I = -\tfrac{1}{2}\hbar\rangle. \tag{8.38}$$

To further simplify things, we will assume that $\theta_I = 0$, i.e. the $\hat{\mathbf{I}}$ vector is in the z direction. In addition, we will make the notational simplification already used above defined by

$$\begin{aligned} |S_i = \tfrac{1}{2}\hbar\rangle &\rightarrow |S\rangle \\ |S_I = \pm\tfrac{1}{2}\hbar\rangle &\rightarrow |\pm\rangle \\ \langle S_I = \pm\tfrac{1}{2}\hbar | S_i = \tfrac{1}{2}\hbar\rangle &\rightarrow \langle \pm | S\rangle \end{aligned}$$

so that $|\pm\rangle = |S_z = \pm\frac{1}{2}\hbar\rangle$, and Eq. (8.36) becomes

$$|S\rangle = |+\rangle\langle +|S\rangle + |-\rangle\langle -|S\rangle. \tag{8.39}$$

or, using Eq. (8.37)

$$|S\rangle = \cos(\tfrac{1}{2}\theta_i)|+\rangle + \sin(\tfrac{1}{2}\theta_i)|-\rangle. \tag{8.40}$$

We are now at the point at which we can begin to supply an interpretation to this equation. What this equation is saying is that the combination $\cos(\frac{1}{2}\theta_i)|+\rangle + \sin(\frac{1}{2}\theta_i)|-\rangle$, and $|S\rangle$, both represent the same thing – the atomic spin is in a state for which $S_i = \frac{1}{2}\hbar$. In other words, if we were presented with the combination:

$$\frac{1}{\sqrt{2}}|+\rangle + \frac{1}{\sqrt{2}}|-\rangle \tag{8.41}$$

we immediately see that $\cos(\frac{1}{2}\theta_i) = 1/\sqrt{2}$ and $\sin(\frac{1}{2}\theta_i) = 1/\sqrt{2}$, and hence $\theta_i = 90°$. Thus the magnetic field is pointing in the direction 90° to the z direction, i.e. in the x direction, and hence the spin state of the atom is the state $|S\rangle = |S_x = \frac{1}{2}\hbar\rangle$.

But what if we were presented with the combination $2|+\rangle + 2|-\rangle$? Here, we cannot find any angle θ_i, so it appears that this combination is not a possible state of the atomic spin. But we can write this as

$$2\sqrt{2}\left[\frac{1}{\sqrt{2}}|+\rangle + \frac{1}{\sqrt{2}}|-\rangle\right] \tag{8.42}$$

which we can now understand as representing $2\sqrt{2}|S_x = \frac{1}{2}\hbar\rangle$. Is $2\sqrt{2}|S_x = \frac{1}{2}\hbar\rangle$ a different physical state of the system to $|S_x = \frac{1}{2}\hbar\rangle$? Well, it is our notation, so we can say what we like, and what turns out to be preferable is to say that $\alpha|S\rangle$ describes the same physical state as $|S\rangle$, for any value of the constant α. Thus, we can say that $2|+\rangle + 2|-\rangle$ is also a state of the system, namely $2\sqrt{2}|S_x = \frac{1}{2}\hbar\rangle$, which represents the same physical information about the state of the system as $|S_x = \frac{1}{2}\hbar\rangle$.

Thus any combination $C_+|+\rangle + C_-|-\rangle$ where $C_\pm$ are real numbers will always represent some state of the system, in general given by

$$\sqrt{C_+^2 + C_-^2}\,|S_i = \tfrac{1}{2}\hbar\rangle \tag{8.43}$$

where

$$S_i = \mathbf{S}\cdot\hat{\mathbf{n}} \tag{8.44}$$

and where $\hat{\mathbf{n}}$ is a unit vector in the direction defined by the angle

$$\theta_i = 2\tan^{-1}\left(\frac{C_-}{C_+}\right). \tag{8.45}$$

Conversely, given any state of the system, we can work out how to write it in the form $C_+|+\rangle + C_-|-\rangle$. Further, we can repeat the whole of the above discussion for any other choice of the intermediate states $|S_I = \pm\frac{1}{2}\hbar\rangle$.

It is this last fact that a state $|S\rangle$ can be written as the linear combination or linear superposition $C_+|+\rangle + C_-|-\rangle$ of two other states, analogous to Eq. (8.1) for the arbitrary vector $\mathbf{v}$, and conversely that any linear superposition of states is itself another state is the essential property that these states need to possess in order for them to be interpreted as vectors belonging to some vector vector space, specifically here a real vector space.

In the more general case of the magnetic fields vectors not all being in the XZ plane, the probability amplitudes $\langle\pm|S\rangle$ will be, in general, complex numbers. For instance, it can be shown that the state $|S\rangle = |S_n = \frac{1}{2}\hbar\rangle$ where $S_n = \mathbf{S}\cdot\hat{\mathbf{n}}$, and where $\hat{\mathbf{n}} = \sin\theta\cos\phi\,\hat{\mathbf{i}} + \sin\theta\sin\phi\,\hat{\mathbf{j}} + \cos\theta\,\hat{\mathbf{k}}$ is a unit vector oriented in a direction defined by the spherical polar angles θ, ϕ, is given by

$$|S\rangle = \cos(\tfrac{1}{2}\theta)|+\rangle + e^{i\phi}\sin(\tfrac{1}{2}\theta)|-\rangle \qquad (8.46)$$

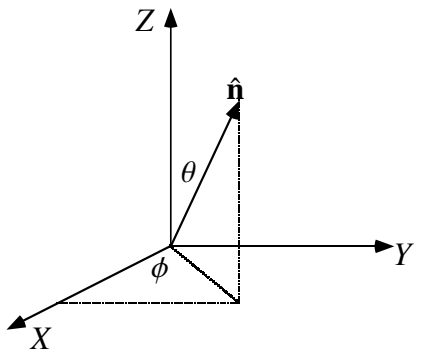

Figure 8.2: Polar angles for defining direction of unit vector $\hat{\mathbf{n}}$

Nevertheless, the arguments presented above continue to apply. In particular *any* linear combination $C_-|-\rangle + C_+|+\rangle$, where $C_\pm$ are complex numbers, will be a possible spin state of an atom. But since the coefficients are complex numbers, the vector space, or state space, is a *complex* vector space. This more general case, is what is usually encountered in quantum mechanics and below, we will assume that the probability amplitudes are, in general, complex.

8.3 The General Case of Many Intermediate States

The above results were based on the particular case of a spin half system which could pass through two possible intermediate spin states, and was developed, at least in part, by analogy with the two slit experiment where a particle could pass through two possible intermediate states defined by the positions of the slits. Both these examples can be readily generalized. For instance, we could consider an interference experiment in which there are multiple slits in the first barrier. Clearly the arguments used in the two slit case would apply again here, with the final result for the probability amplitude of observing an electron striking the screen at position x after having set out from a source S being given by

$$\langle x|S\rangle = \sum_{n=1}^{N}\langle x|n\rangle\langle n|S\rangle \qquad (8.47)$$

where we have supposed that there are N slits in the first barrier, so that the electron can pass through N intermediate states. Likewise, if we repeat the Stern-Gerlach experiment with spin 1 atoms, we would find that the atoms would emerge in one or the other of three beams corresponding to $S_z = -\hbar, 0, \hbar$. More generally, if the atoms have spin s (where

$s = 0$ or $\frac{1}{2}$ or 1 or $\frac{3}{2}, \ldots$ then they will emerge in one of a total of $2s + 1$ different possible beams corresponding to $S_z = -s\hbar$, or $(-s + 1)\hbar$, or $\ldots$, or $(s - 1)\hbar$, or $s\hbar$. We can then write

$$\langle S'|S\rangle = \sum_{n=-s}^{s} \langle S'|n\rangle\langle n|S\rangle \tag{8.48}$$

i.e. the atom can pass through $2s + 1$ intermediate states, where we have written $|n\rangle$ for the state in which the atom has a z component of spin $S_z = n\hbar$ [1]. What we want to do is extract from these two examples a general prescription that we can apply to any system in order to build up a quantum mechanical description of that system. In order to set the scene for this development, we need to reconsider again the general features that apply to the intermediate states appearing in the above two expressions. The important properties of these intermediate states are as follows:

1. Each intermediate state represents a mutually exclusive possibility, that is, if the system is observed to be in one of the intermediate states, it is definitely the case that it will not also be observed in any of the others: the electron can be observed to pass through one slit only, or a spin s atom would be seen to emerge from a Stern-Gerlach apparatus in one beam only. It is this mutual exclusiveness that allows us to write such things as $\langle -|+\rangle = 0$ or $\langle +|+\rangle = 1$ and so on.

2. The list of intermediate states covers all possibilities: if there are N slits in the barrier, then there will be N intermediate states $|n\rangle$ – there are no other slits for the electron to pass through. Likewise, if an atom with spin s enters a Stern-Gerlach apparatus, it emerges in one of $2s + 1$ possible beams. Thus the states $|n\rangle$, $n = -s, -s + 1, \ldots, s - 1, s$ account for all the possible states that the atom can emerge in.

3. Finally, the claim is made that the probability amplitude for finding the system in the state $\langle\phi|$ given that it was in the state $|\psi\rangle$ can be written as

$$\langle\phi|\psi\rangle = \sum_{n} \langle\phi|n\rangle\langle n|\psi\rangle \tag{8.49}$$

which allows us to write the arbitrary state $|\psi\rangle$ as

$$|\psi\rangle = \sum_{n} |n\rangle\langle n|\psi\rangle \tag{8.50}$$

where $\langle n|\psi\rangle$, the probability amplitude of finding the system in the state $|n\rangle$, given that it was in the state $|\psi\rangle$, is now viewed as a 'weighting' for the state $|\psi\rangle$ to be observed in the state $|n\rangle$.

The above three properties of intermediate states shows us how to generalize the results that were obtained for the two slit and spin half examples used to develop the ideas. However, the ideas developed ought to be applicable to systems other than these two possibilities. Thus, we need to extract out of what has been said so far ideas that can be

[1] Note, n is not necessarily an integer here. If $s = \frac{1}{2}$ for instance, then n will have the two possible values $n = \pm\frac{1}{2}$, and the states $|\pm\frac{1}{2}\rangle$ are just the states we have earlier called $|\pm\rangle$.

applied to any physical system so as to generate a quantum description of such a system. We do this by noting the singular role played in the above discussion of the 'intermediate states'. If we can identify such intermediate states for any physical system, then we are in a position to formulate a quantum description of such a system.

The characteristics of the intermediate states that are of greatest importance are the following:

1. They represent mutually exclusive alternatives, that is, if a measurement is made, a system is observed in only one of these possible intermediate states;
2. The intermediate states cover all possibilities, that is, if a measurement is made, the system is always observed to be in one of these intermediate states;
3. If the system is initially prepared in some initial state $|\psi\rangle$, then the probability amplitude of observing it in some arbitrary final state $|\phi\rangle$ is given by a sum of probability amplitudes of passing through each of these intermediate states, in a manner analogous to Eq. (8.48).

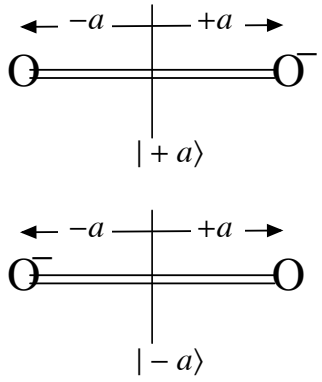

Figure 8.3: O_2^- ion with the two possible positions of the electron, corresponding to the two states $|\pm a\rangle$.

We will now see how these ideas can be applied to more general kinds of physical systems. To begin, we have to set up these intermediate states for a given system. So suppose we perform an *exhaustive* series of measurements of some observable property of the system – call it Q. For example, we could determine through which slits it is possible for an electron to pass in the two slit experiment above, or the possible values of the Z component of the spin of a particle, or the possible values of the position of an electron on an O_2^- ion, as in the adjacent figure. We will give many more examples later. Whatever the system, what we mean by exhaustive is that we determine *all* the possible values that the observed quantity Q might have. For instance, we determine that, in the two slit interference experiment, the electron can pass through, (that is, be observed to be at) the position of one or the other of the two slits, or that a spin half particle would be observed to have either of the values $S_z = \pm\frac{1}{2}\hbar$, and no other value. Or for the O_2^- ion, the electron can be found *either* on the atom at position $x = -a$ *or* on the atom at position $x = +a$. In general, for an arbitrary observable Q, let us represent these observed values by $q_1, q_2, q_3, \ldots$ all of which will be real numbers. Of course, there might be other observable properties of the system that we might be able to measure, but for the present we will suppose that we only need concern ourselves with just one.

In keeping with the way that we developed the idea of state earlier, we then say that as a result of the measurement producing, for instance, the outcome q_1, we then say that the system ends up in the state $|q_1\rangle$, and similarly for all the others possibilities. Moreover, there is now the possibility, if the system is initially prepared in the initial state $|\psi\rangle$, of observing it in any one of these states $|q_n\rangle$, but once again, we are unable to predict with certainty which of these states we will observe the system to be in – we can only assign

a probability $P(q_n|\psi)$ of finding the system in the state $|q_n\rangle$ given it was in the state $|\psi\rangle$. Likewise, there will be a probability $P(\phi|q_n)$ of finding the system in the final state $|\phi\rangle$ given that it was in the state $|q_n\rangle$. It is at this point that we reach the result that, in a sense, defines quantum mechanics. If the possibility exists of finding the system in any of the states $|q_1\rangle$, $|q_2\rangle$, ..., then according to classical ideas of probability, we would be safe in assuming the intuitive formula

$$P(\phi|\psi) = \sum_n P(\phi|q_n)P(q_n|\psi) \quad \text{— a result of classical probability theory}$$

i.e. the system starts in the state $|\psi\rangle$, and has the probability $P(q_n|\psi)$ that it will end up in the state $|q_n\rangle$, and once it 'arrives there' we then have the probability $P(\phi|q_n)$ that it will then end up in the state $|\phi\rangle$ given that it was in the state $|q_n\rangle$. Multiplying these two probabilities will then give the probability of the system starting from $|\psi\rangle$ and passing through $|q_n\rangle$ on its way to the final state $|\phi\rangle$. We then sum over all the possible intermediate states $|q_n\rangle$ to give the total probability of arriving in the state $|\phi\rangle$. However, what is found experimentally is that *if the system is never observed in any of the intermediate states* $|q_n\rangle$, this probability is not given by this classical result – the measurements show evidence of interference effects, which can only be understood if this probability is given as the square of a 'probability amplitude'.

Thus, we find we must write $P(\phi|\psi) = |\langle\phi|\psi\rangle|^2$, where $\langle\phi|\psi\rangle$ is a complex number, generally referred to as the probability amplitude of finding the system in state $|\phi\rangle$ given that it was in state $|\psi\rangle$, and this probability amplitude is then given by

$$\langle\phi|\psi\rangle = \sum_n \langle\phi|q_n\rangle\langle q_n|\psi\rangle \tag{8.51}$$

where the sum is over the probability amplitudes of the system passing through all the possible states associated with all the possible observed values of the quantity Q.

If we square this result we find that

$$\begin{aligned} P(\phi|\psi) = |\langle\phi|\psi\rangle|^2 &= \sum_n |\langle\phi|q_n\rangle|^2|\langle q_n|\psi\rangle|^2 + \text{cross terms} \\ &= \sum_n P(\phi|q_n)P(q_n|\psi) + \text{cross terms.} \end{aligned} \tag{8.52}$$

We note that this expression consists, in part, of the classical result Eqn. (8.3), but there is, in addition, cross terms. It is these terms that are responsible for the interference effects observed in the two slit experiment, or in the Stern-Gerlach experiment. If we observe which state $|q_n\rangle$ the system 'passes through' it is always found that these interference terms are washed out, reducing the result to the classical result Eqn. (8.3). This we have seen in the case of two slit interference wherein if the slit through which the particle passes is observed, then the interference pattern on the observation screen is washed out.

What remains to be found is ways to calculate these 'probability amplitudes'. We can go a long way towards doing this by recognizing, from a purely physical perspective, that the probability amplitudes $\langle q_m|q_n\rangle$ must have the following properties:

- $\langle q_m|q_n\rangle = 0$ if $m \neq n$. This amounts to stating that if the system is in the state $|q_n\rangle$, i.e. wherein the observable Q is known to have the value q_n, then there is zero possibility of finding it in the state $|q_m\rangle$.

- $\langle q_n|q_n\rangle = 1$. This asserts that if the system is in the state for which the quantity Q has the value q_n, then it is certain to be found in the state in which it has the value q_n. Thus the states $\{|q_n\rangle; n = 1, 2, \ldots\}$ are *mutually exclusive*.
- The states $\{|q_n\rangle; n = 1, 2, \ldots\}$ are also exhaustive in that they cover *all* the possible values that could be observed of the observable Q. These states are said to be *complete* – simply because they cover all possibilities – and the result Eqn. (8.51) as the 'closure relation' for these states.

Then we propose the following set of ideas and concepts:

1. $\langle\phi|\psi\rangle$ = the probability amplitude of the system being found in the state $|\phi\rangle$ given that it was in state $|\psi\rangle$.
2. $|\langle\phi|\psi\rangle|^2$ = the probability of the system being found in the state $|\phi\rangle$ given that it was in state $|\psi\rangle$.
3. The states $\{|q_n\rangle; n = 1, 2, \ldots\}$ are mutually exclusive i.e.

$$\begin{aligned} \langle q_m|q_n\rangle = \delta_{mn} &= 1 \quad n = m \\ &= 0 \quad n \neq m \end{aligned} \tag{8.53}$$

   where δ_{mn} is known as the Kronecker delta. The states $\{|q_n\rangle; n = 1, 2, \ldots\}$ are said to be *orthonormal*.

   This condition satisfied by these states is a mathematical expression of the physical fact that if the system is in one of the states $|q_n\rangle$, then there is no possibility of finding it in some other state $|q_m\rangle$, $m \neq n$, hence $\langle q_m|q_n\rangle = 0$ while the probability of finding it in the state $|q_n\rangle$ must be unity.

4. The fundamental law of quantum mechanics, otherwise known as the *closure relation*:

$$\langle\phi|\psi\rangle = \sum_n \langle\phi|q_n\rangle\langle q_n|\psi\rangle \tag{8.54}$$

   tells us that the total probability amplitude of finding the system in the final state $|\phi\rangle$ is just the sum of the probability amplitudes of the system 'passing through' any of the states $\{|q_n\rangle; n = 1, 2, \ldots\}$. This is the 'sum over paths' idea alluded to above.

It is at this point that we perform the 'cancellation' trick which results in the states being expressed in terms of other states as in the following, which ultimately leads us to interpreting these states as vectors.

5. The completeness relations:

$$|\psi\rangle = \sum_n |q_n\rangle\langle q_n|\psi\rangle \tag{8.55}$$

$$\langle\phi| = \sum_n \langle\phi|q_n\rangle\langle q_n| \tag{8.56}$$

   which tells us that either the initial or final states can be expressed as a linear combination of the intermediate states $\{|q_n\rangle; n = 1, 2, \ldots\}$. This is a statement that these states are *complete* in that there is enough of them such that *any* state of the system can be written in the above form in terms of these states.

6. $\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*$.

7. The normalization condition:

$$\langle\psi|\psi\rangle = \sum_n \langle\psi|q_n\rangle\langle n|\psi\rangle = \sum_n |\langle q_n|\psi\rangle|^2 = 1. \tag{8.57}$$

This last result tells us that the probability of finding the system in any of the states $|n\rangle$ adds up to unity. Moreover, we know that $\langle q_m|q_n\rangle = 0$ if $n \neq m$ so that a system in state $|q_n\rangle$ can never be found in some other state $|q_m\rangle$. Consequently, the set of states $\{|q_n\rangle; n = 1, 2, \dots\}$ represent a complete set of possible alternative final states, complete in the sense that the total probability of ending up in any of the mutually exclusive possible final states $|q_n\rangle$ adds up to unity – there is nowhere else for the system to be found.

By a simple extension of the arguments presented in Section 8.2 in the case of spin half quantum states it is now possible to show how the above properties 1 – 7 are completely analogous to the properties of vectors in a complex vector space. This mathematical formalism will be developed more fully in Section 8.5, but for the present we can summarize the essential ideas based on what we have already put forward earlier. The importnat points then are as follows:

1. The collection of all the possible state vectors of a quantum system forms a complex vector space known as the *state space* of the system.

2. The probability amplitudes are identified as the inner product of these state vectors.

3. The intermediate states $\{|q_n\rangle; n = 1, 2, \dots\}$ form a complete orthonormal set of basis states of this state space, i.e. any state vector $|\psi\rangle$ can be written as a linear combination of these basis states.

4. The number of basis states is known as the dimension of the state space.

8.4 Constructing a State Space

The ideas developed above can now be applied to constructing a state space for a physical system. The basic idea is as discussed in Section 8.3 which enables us to define a set of basis states for the state space of the system. By establishing a set of basis states, in a sense, we 'bring the state space into existence', and once this is done, we are free to use all the mathematical machinery available for analysing the properties of the state space so constructed. The question can be asked as to whether or not the ideas presented in Section 8.3, admittedly extracted from only a handful of examples, can be applied with success to any other system. This is a question that can only be answered by applying the rules formulated there and considering the consequences. In Section 8.7 we will discuss where these ideas, if naively applied, fail to work. Otherwise, these ideas, when fully formed, constitute the basis of quantum physics.

In accordance with the ideas developed in Section 8.3, constructing a state space for a physical system can be carried out by recognizing the intermediate states through which a system can pass as it makes its way from some initial state to some observed final state,

as was done in the case of the two slit, or spin half systems. Thus, in the two slit example, the two possible intermediate states are those for which the particle is to be found at the position of either of the two slits. In the spin half example, the two intermediate states are those in which the spin is observed to have either of the two values $S_z = \pm\frac{1}{2}\hbar$; these are the states we have been calling $|\pm\rangle$. These intermediate states are states of the system that can be identified through an argument based on the idea that some physical property of the system can be exhaustively measured to yield a set of values that we then use to label a complete set of basis states for the state space of the system.

Negatively Charged Ions Here the system is a molecule which has acquired an extra electron, which can be assumed to found only on any one of the atoms making up the molecule. This is, of course, an approximation. The electron could be found anywhere around the atoms, or in the space between the atoms, in a way that depends on the nature of the chemical bond between the atoms. Here we are making use of a coarse notion of position, i.e. we are assuming that the electron can be observed to reside on one atom or the other, and we do not really care about exactly where on each atom the electron might be found. The idea is best illustrated by the simple example of the O_2^- ion in which the electron can be found on one or the other of the oxygen atoms (see Fig. (8.3)). If we let $x = \pm a$ be the positions of the two atomic nuclei with respect to a point midway between them, then we can use these two possible positions of the electron to label the possible states of the system. The two states are then $|-a\rangle$ and $|+a\rangle$. If the electron is observed to be on the oxygen atom at $x = +a$, then it has unit probability of being observed on the atom at $x = +a$, but zero probability of being observed on the atom at $x = -a$. These two states represent mutually exclusive possibilities, and so they are distinct orthonormal states:

$$\begin{aligned} \langle +a|+a\rangle &= \langle -a|-a\rangle = 1 \\ \langle +a|-a\rangle &= \langle -a|+a\rangle = 0 \end{aligned} \tag{8.58}$$

These two states are also complete in that, within the limits of our model, there is nowhere else for the electron to be found – it is either on the atom at $x = +a$ or on the atom at $x = -a$. These two states form a complete set of orthonormal basis states for the state space of the ion, so that *any* state of the ion can be expressed in the form

$$|\psi\rangle = c_1|+a\rangle + c_2|-a\rangle \tag{8.59}$$

where c_1, c_2 are complex numbers – this is just the completeness relation for this system. Likewise, any linear combination of $\{|+a\rangle, |-a\rangle\}$ will represent a possible state of the ion. As there are two basis states, the state space for the system has dimension 2.

This kind of model can be generalized to situations involving different geometries, such as atoms arranged in a ring e.g. an ozone ion O_3^-. In this case, the state space will be spanned by three basis states corresponding to the three possible positions at which the electron can be observed. This model (and its generalizations to an arbitrary number of atoms arranged in a ring) is valuable as it gives rise to results that serve as an approximate treatment of angular momentum in quantum mechanics.

Ammonia molecule Here the system is the ammonia molecule NH_3 in which the nitrogen atom is at the apex of a triangular pyramid with the three hydrogen atoms forming an equilateral triangle as the base. The nitrogen atom can be positioned either above or below the plane of the hydrogen atoms, these two possibilities we take as two possible states of the ammonia molecule. (The N atom can move between these two positions by 'quantum 'tunnelling' through the potential barrier lying in the plane of the hydrogen atoms.) Once again, this is a state space of dimension 2.

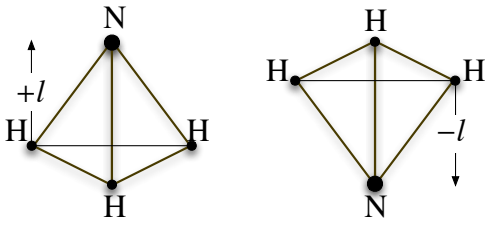

Figure 8.4: Ammonia molecule in two states distingusished by the position of the nitrogen atom, either above or below the plane of the hydrogen atoms, corresponding to the states $|+l\rangle$ and $|-l\rangle$ respectively.

Spin Flipping In this case, we have a spin half particle (for instance) in a constant magnetic field, so the two possible states are the familiar spin up or spin down states. If, in addition, we add a *rotating* magnetic field at right angles to the constant field, there arises a time dependent probability of the spin flipping from one orientation to the other. As the spin up and spin down states are of different energies, this represents a change in energy of the particle, a change that can be detected, and is the basis of the electron spin and nuclear magnetic resonance imaging much used in medical work. Obviously this is a state space of dimension two.

The Qubit Any two state system can be used to represent the quantum version of a binary numbers: spin up and spin down, an atom excited or not, and so on. Overall the two states can be represented by $|0\rangle$ and $|1\rangle$, corresponding to the two binary numbers 0 and 1. A linear combination

$$|\psi\rangle = c_0|0\rangle + c_1|1\rangle.$$

can be formed which represents the possibility of a memory registering a bit of information, not as either a 0 or a 1, which is all that can happen classically, but simultaneously, registering both possibilities of 0 or 1. Such a state is known as a *qubit*. Obviously, the state space is of dimension two, and much that we have said above about spin half systems applies. Quantum computation then involves manipulating the whole state $|\psi\rangle$, which, in effect, amounts to performing two calculations at once, differing by the initial setting of the memory bit. The idea introduced above can be readily extended. Thus, if we have two two level atoms, we have such possibilities as $|00\rangle$, $|01\rangle$, $|10\rangle$, and $|11\rangle$ where, for instance, $|10\rangle$ is the state in which the first atom is in its excited state and the second is in its ground state. Obviously, the state $|00\rangle$ represents the number zero, $|01\rangle$ the number one, $|10\rangle$ the number two, and $|11\rangle$ the number three. We now have two qubits, and a state space of dimension four, and we can set up linear combinations such as

$$|\psi\rangle = c_{00}|00\rangle + c_{01}|01\rangle + c_{10}|10\rangle + c_{11}|11\rangle \tag{8.60}$$

and we can then perform calculations making use, simultaneously, of four different possible values for whatever quantity the states are intended to represent. With three atoms, or

four and so on, the state space becomes much larger: of dimension 2^N in fact where N is the number of qubits, and the basis states represent the numbers ranging from 0 to $2^N - 1$ in binary notation.

Benzene Molecule An example of quite a different character is that of the benzene molecule, illustrated in Fig. 8.5. The two states of the molecule are distinguished by the positioning of the double bonds between pairs of carbon atoms. The molecule, at least with regard to the arrangements of double bonds can be found in two different states which, for want of a better name, we will call $|\alpha\rangle$ and $|\beta\rangle$. The state space is therefore of dimension 2, and an arbitrary state of the molecule would be given by

$$|\psi\rangle = a|\alpha\rangle + b|\beta\rangle. \qquad (8.61)$$

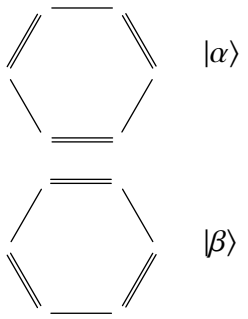

Figure 8.5: Two arrangements of the double bonds in a benzene moelcule corresponding to two states $|\alpha\rangle$ and $|\beta\rangle$.

All the examples given above yield state spaces of finite dimension. Much the same argument can be applied to construct state spaces of infinite dimension. A couple of examples follow.

The Tight-Binding Model of a Crystalline Metal The examples given above of an electron being positioned on one of a (finite) number of atoms can be readily generalized to a situation in which there are an infinite number of such atoms. This is not a contrived model in any sense, as it is a good first approximation to modelling the properties of the conduction electrons in a crystalline solid. In the free electron model of a conducting solid, the conduction electrons are assumed to be able to move freely (and without mutual interaction) through the crystal, i.e. the effects of the background positive potentials of the positive ions left is ignored. A further development of this model is to take into account the fact that the electrons will experience some attraction to the periodically positioned positive ions, and so there will be a tendency for the electrons to be found in the neighbourhood of these ions. The resultant model – with the basis states consisting of a conduction electron being found on any one of the ion sites – is obviously similar to the one above for the molecular ion. Here however, the number of basis states is infinite (for an infinite crystal), so the state space is of infinite dimension. Representing the set of basis states by $\{|n\rangle, n = 0, \pm 1, \pm 2, \dots\}$ where na is the position of the n^{th} atom, and a is the separation between neighbouring atoms, then any state of the system can then be written as

$$|\psi\rangle = \sum_{n=-\infty}^{+\infty} c_n |n\rangle. \qquad (8.62)$$

By taking into account the fact that the electrons can make their way from an ion to one of its neighbours, much of the band structure of semiconducting solids can be obtained.

Free Particle We can generalize the preceding model by supposing that the spacing between the neighbouring atoms is allowed to go to zero, so that the positions at which the electron can be found become continuous. This then acts as a model for the description of a particle free to move anywhere in one dimension. In setting up this model, we find that as well as there being an infinite number of basis states – something we have already encountered – we see that these basis states are not discrete, i.e. a particle at position x will be in the basis state $|x\rangle$, and as x can vary continuously over the range $-\infty < x < \infty$, there will be a non-denumerably infinite, that is, a continuous range of such basis states. As a consequence, the completeness relation ought to be written as an integral:

$$|\psi\rangle = \int_{-\infty}^{+\infty} |x\rangle\langle x|\psi\rangle \, dx. \tag{8.63}$$

The states $|x\rangle$ and $|x'\rangle$ will be orthonormal if $x \neq x'$, but in order to be able to retain the completeness relation in the form of an integral, it turns out that these basis states have to have an infinite norm. However, there is a sense in which we can continue to work with such states, as will be discussed in Section 8.6.

Particle in an Infinitely Deep Potential Well We saw in Section 5.3 that a particle of mass m in an infinitely deep potential well of width L can have the energies $E_n = n^2\pi^2\hbar^2/2mL^2$ where n is a positive integer. This suggests that the basis states of the particle in the well be the states $|n\rangle$ such that if the particle is in state $|n\rangle$, then it has energy E_n. The probability amplitude of finding the particle at position x when in state $|n\rangle$ is then $\langle x|n\rangle$ which, from Section 5.3 we can identify with the wave function ψ_n, i.e.

$$\begin{aligned} \psi_n(x) = \langle x|n\rangle &= \sqrt{\frac{2}{L}} \sin(n\pi x/L) \quad 0 < x < L \\ &= 0 \qquad\qquad\qquad\qquad x < 0, \quad x > L. \end{aligned} \tag{8.64}$$

The state space is obviously of infinite dimension.

It has been pointed out before that a state space can have any number of sets of basis states, i.e. the states $|n\rangle$ introduced here do not form the sole possible set of basis states for the state space of this system. In this particular case, it is worthwhile noting that we could have used as the base states the states labelled by the position of the particle in the well, i.e. the states $|x\rangle$.

As we have seen, there are an infinite number of such states which in a way is to be expected as we have already seen that the state space is of infinite dimension. But the difference between this set of states and the set of states $|n\rangle$ is that in the latter case, these states are discrete, i.e. they can be labelled by the integers, while the states $|x\rangle$ are continuous, they are labelled by the continuous variable x. Thus, something new emerges from this example: for state spaces of infinite dimension, it is possible to have a denumerably infinite number of basis states (i.e. the discrete states $|n\rangle$) or non-denumerably infinite number of basis states (i.e. the states $|x\rangle$.) This feature of state spaces of infinite dimension, plus others, are discussed separately below in Section 8.6.

A System of Identical Photons Many other features of a quantum system not related to the position or energy of the system can be used as a means by which a set of basis states

can be set up. An important example is one in which the system consists of a possibly variable number of identical particles. One example is a 'gas' of photons, all of the same frequency and polarization. Such a situation is routinely achieved in the laboratory using suitably constructed hollow superconducting metallic cavities designed to support just one mode (i.e. a single frequency and polarization) of the electromagnetic field. The state of the electromagnetic field can then be characterized by the number n of photons in the field which can range from zero to positive infinity, so that the states of the field (known as number states) can be written $|n\rangle$ with $n = 0, 1, 2, \ldots$. The state $|0\rangle$ is often referred to as the vacuum state. These states will then constitute a complete, orthonormal set of basis states (called Fock or number states), i.e.

$$\langle n|m\rangle = \delta_{nm} \tag{8.65}$$

and as n can range up to infinity, the state space for the system will be infinite dimensional. An arbitrary state of the cavity field can be then be written

$$|\psi\rangle = \sum_{n=0}^{\infty} c_n|n\rangle \tag{8.66}$$

so that $|c_n|^2$ will be the probability of finding n photons in the field. In terms of these basis states, it is possible to describe the processes in which particles are created or destroyed. For instance if there is a single atom in an excited energy state in the cavity, and the cavity is in the vacuum state $|0\rangle$, then the state of the combined atom field system can be written $|e, 0\rangle$, where the e indicates that the atom is in an excited state. The atom can later lose this energy by emitting it as a photon, so that at some later time the state of the system will be $a|e, 0\rangle + b|g, 1\rangle$, where now there is the possibility, with probability $|b|^2$, of the atom being found in its ground state, and a photon having been created.

8.5 General Mathematical Description of a Quantum System

It was shown in preceding Sections that the mathematical description of this sum of probability amplitudes admits an interpretation of the state of the system as being a vector in a complex vector space, the state space of the system. It is this mathematical picture that is summarized here in the general case introduced in the immediately preceding Section. This idea that the state of a quantum system is to be considered a vector belonging to a complex vector space, which we have developed here in the case of a spin half system, and which has its roots in the sum over paths point of view, is the basis of all of modern quantum mechanics and is used to describe any quantum mechanical system. Below is a summary of the main points as they are used for a general quantum system whose state spaces are of arbitrary dimension (including state spaces of infinite dimension). The emphasis here is on the mathematical features of the theory.

8.5.1 State Space

We have indicated a number of times that in quantum mechanics, the state of a physical system is represented by a vector belonging to a complex vector space known as the state

space of the system. Here we will give a list of the defining conditions of a state space, though we will not be concerning ourselves too much with the formalities. The following definitions and concepts set up the state space of a quantum system.

1. Every physical state of a quantum system is specified by a vector called a state vector, ket vector, or sometimes just state and written $|\ldots\rangle$ where $\ldots$ is a label specifying the physical information known about the state. An arbitrary state is written $|\psi\rangle$, or $|\phi\rangle$ and so on. The set of all state vectors describing a given physical system forms a complex vector space (actually a Hilbert space, see Sec. 8.5.2) $\mathcal{H}$ also known as the state space or ket space for the system.

2. Every linear superposition of two or more state vectors $|\phi_1\rangle$, $|\phi_2\rangle$, $|\phi_3\rangle$, $\ldots$, is also a state of the quantum system i.e. the state $|\psi\rangle$ given by

$$|\psi\rangle = c_1|\phi_1\rangle + c_2|\phi_2\rangle + c_3|\phi_3\rangle + \ldots$$

   is a state of the system for all complex numbers $c_1, c_2, c_3, \ldots$.

These last two points amount to saying that every physical state of a system is represented by a vector in the state space of the system, and every vector in the state space represents a possible physical state of the system. To guarantee this, the following condition is also imposed:

3. If a state of the system is represented by a vector $|\psi\rangle$, then the same state is represented by the vector $c|\psi\rangle$ where c is any non-zero complex number.

Finally, we need the concept of a set of basis states, and of the dimension of the state space.

4. A set of vectors $\{|\varphi_1\rangle, |\varphi_2\rangle, |\varphi_3\rangle, \ldots\}$ is said to form a basis for the state space if every state of the quantum system can be represented as a linear superposition of the $|\varphi_i\rangle$'s i.e. for any state $|\psi\rangle$ we can write

$$|\psi\rangle = \sum_i c_i|\varphi_i\rangle.$$

   The set of vectors $\{|\varphi_i\rangle, i = 1, 2, \ldots\}$ are said to *span* the vector space. The vectors are also termed the *base* states for the vector space. They are also said to be *complete*. What this means, mathematically, is that for every state $|\phi\rangle$ say, at least one of the inner products $\langle\varphi_n|\phi\rangle$ will be non-zero, or conversely, there does not exist a state $|\xi\rangle$ for which $\langle\varphi_n|\xi\rangle = 0$ for all the basis states $|\varphi_n\rangle$. Completeness clearly means that no more basis states are needed to describe any possible physical state of a system.

For example, returning to the spin half system, the two states $|\pm\rangle$ are all that is needed to describe any state of the system, i.e. there are no spin states that cannot be described in terms of these basis states. Thus, these states are said to be complete.

5. The minimum number of vectors needed to form a complete set of basis states is known as the dimension of the state space. [In many, if not most cases of interest in quantum mechanics, the dimension of the state space is infinite.]

It should be noted that there is an infinite number of possible sets of basis states for any state space. The arguments presented above by which we arrive at a set of basis states (or intermediate states) serves as a physically motivated starting point to construct the state space for the system. But once we have defined the state space in this way, there is no reason why we cannot, at least mathematically, construct other sets of basis states. These basis states that we start with are particularly useful as they have an immediate physical meaning; this might not be the case for an arbitrary basis set. But there are other means by which other physically meaningful basis states can be determined: often the choice of basis states is suggested by the physics (such as the set of eigenstates of an observable, see Chapter 9).

8.5.2 Probability Amplitudes and the Inner Product of State Vectors

We obtained a number of properties of probability amplitudes when looking at the case of a spin half system. Some of the results obtained there, and a few more that were not, are summarized in the following.

If $|\phi\rangle$ and $|\psi\rangle$ are any two state vectors belonging to the state space $\mathcal{H}$, then

1. $\langle\phi|\psi\rangle$, a complex number, is the probability amplitude of observing the system to be in the state $|\phi\rangle$ given that it is in the state $|\psi\rangle$.
2. The probability of observing the system to be in the state $|\phi\rangle$ given that it is in the state $|\psi\rangle$ is $|\langle\phi|\psi\rangle|^2$.

The probability amplitude $\langle\phi|\psi\rangle$, can then be shown to have the properties

3. $\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*$.
4. $(\langle\phi|\{c_1|\psi_1\rangle + c_2|\psi_2\rangle\} = c_1\langle\phi|\psi_1\rangle + c_2\langle\phi|\psi_2\rangle$ where c_1 and c_2 are complex numbers.
5. $\langle\psi|\psi\rangle \geq 0$. If $\langle\psi|\psi\rangle$=0 then $|\psi\rangle = 0$, the zero vector.

This last statement is related to the physically reasonable requirement that the probability of a system being found in a state $|\psi\rangle$ given that it is in the state $|\psi\rangle$ has to be unity, i.e. $|\langle\psi|\psi\rangle|^2 = 1$ which means that $\langle\psi|\psi\rangle = \exp(i\eta)$. We now *choose* $\eta = 0$ so that $\langle\psi|\psi\rangle = 1$, which is bigger than zero. But recall that any multiple of a state vector still represents the same physical state of the system, i.e. $|\widetilde{\psi}\rangle = a|\psi\rangle$ still represents the same physical state as $|\psi\rangle$. However, in this case, $\langle\widetilde{\psi}|\widetilde{\psi}\rangle = |a|^2$ which is not necessarily unity, but is is certainly bigger than zero.

6. The quantity $\sqrt{\langle\psi|\psi\rangle}$ is known as the *length* or *norm* of $|\psi\rangle$.
7. A state $|\psi\rangle$ is normalized, or normalized to unity, if $\langle\psi|\psi\rangle = 1$.

Normalized states are states for which probability is properly taken into account. It is mathematically convenient to permit the use of states whose norms are not equal to unity, but it is necessary in order to make use of the probability interpretation to deal only with that state which has norm of unity. Any state that cannot be normalized to unity (i.e. it is of infinite length) cannot represent a physically acceptable state.

8. Two states $|\phi\rangle$ and $|\psi\rangle$ are orthogonal if $\langle\phi|\psi\rangle = 0$.

The physical significance of two states being orthogonal should be understood: for a system in a certain state, there is zero probability of it being observed in a state with which it is orthogonal. In this sense, two orthogonal states are as distinct as it is possible for two states to be.

Finally, a set of orthonormal basis vectors $\{|\varphi_n\rangle; n = 1, 2, \dots\}$ will have the property

9. $\langle\varphi_m|\varphi_n\rangle = \delta_{mn}$ where δ_{mn} is known as the Kronecker delta, and equals unity if $m = n$ and zero if $m \neq n$.

All the above conditions satisfied by probability amplitudes were to a greater or lesser extent physically motivated, but it nevertheless turns out that these conditions are identical to the conditions that are used to define the inner product of two vectors in a complex vector space, in this case, the state space of the system, i.e. we could write, using the usual mathematical notation for an inner product, $\langle\phi|\psi\rangle = (|\phi\rangle, |\psi\rangle)$. The state space of a physical system is thus more than just a complex vector space, it is a vector space on which there is defined an inner product, and so is more correctly termed a complex 'inner product' space. However, it is usually required in quantum mechanics that certain convergency criteria, defined in terms of the norms of sequences of vectors belonging to the state space, must be satisfied. This is not of any concern for spaces of finite dimension, but are important for spaces of infinite dimension. If these criteria are satisfied then the state space is said to be a Hilbert space. Thus rather than referring to the state space of a system, reference is made to the Hilbert space of the system.

It is important to recognize that all the vectors belonging to a Hilbert space have finite norm, or, putting it another way, all the state vectors can be normalized to unity – this state of affairs is physically necessary if we want to be able to apply the probability interpretation in a consistent way. However, as we shall see, we will encounter states which do not have a finite norm and hence neither represent physically realizable states, nor do they belong to the state or Hilbert space of the system. Nevertheless, with proper care regarding their use and interpretation, such states turn out to be essential, and play a crucial role throughout quantum mechanics. Recognizing that a probability amplitude is nothing but an inner product on the state space of the system, leads to a more general way of defining what is meant by a bra vector. The following discussion emphasizes the fact that a bra vector, while it shares many characteristices of a ket vector, is actually a different mathematical entity.

Bra Vectors

We have consistently used the notation $\langle\phi|\psi\rangle$ to represent a probability amplitude, but we have just seen that this quantity is in fact nothing more than the inner product of two state

vectors, which can be written in a different notation, $(|\phi\rangle, |\psi\rangle)$, that is more commonly encountered in pure mathematics. But the inner product can be viewed in another way, which leads to a new interpretation of the expression $\langle\phi|\psi\rangle$, and the introduction of a new class of state vectors. If we consider the equation

$$\langle\phi|\psi\rangle = (|\phi\rangle, |\psi\rangle) \tag{8.67}$$

and 'cancel' the $|\psi\rangle$, we get the result

$$\langle\phi| \bullet = (|\phi\rangle,\ \bullet) \tag{8.68}$$

where the '$\bullet$' is inserted, temporarily, to remind us that in order to complete the equation, a ket vector has to be inserted. By carrying out this procedure, we have introduced a new quantity $\langle\phi|$ which is known as a bra or bra vector, essentially because $\langle\phi|\psi\rangle$ looks like quantities enclosed between a pair of 'bra(c)kets'. It is a vector because, as can be readily shown, the collection of all possible bras form a vector space. For instance, by the properties of the inner product, if

$$|\psi\rangle = a_1|\varphi_1\rangle + a_2|\varphi_2\rangle \tag{8.69}$$

then

$$(|\psi\rangle,\ \bullet) = \langle\psi| \bullet = (a_1|\varphi_1\rangle + a_2|\varphi_2\rangle,\ \bullet) \tag{8.70}$$

$$= a_1^*(|\varphi_1\rangle,\ \bullet) + a_2^*(|\varphi_2\rangle,\ \bullet) = a_1^*\langle\varphi_1| \bullet + a_2^*\langle\varphi_2| \bullet \tag{8.71}$$

i.e., dropping the '$\bullet$' symbols, we have

$$\langle\psi| = a_1^*\langle\varphi_1| + a_2^*\langle\varphi_2| \tag{8.72}$$

so that a linear combination of two bras is also a bra, from which follows (after a bit more work checking that the other requirements of a vector space are also satisfied) the result that the set of all bras is a vector space. Incidentally, this last calculation above shows, once again, that if $|\psi\rangle = a_1|\varphi_1\rangle + a_2|\varphi_2\rangle$ then the corresponding bra is $\langle\psi| = a_1^*\langle\varphi_1| + a_2^*\langle\varphi_2|$. So, in a sense, the bra vectors are the 'complex conjugates' of the ket vectors.

The vector space of all bra vectors is obviously closely linked to the vector space of all the kets $\mathcal{H}$, and is in fact usually referred to as the dual space, and represented by $\mathcal{H}^*$. To each ket vector $|\psi\rangle$ belonging to $\mathcal{H}$, there is then an associated bra vector $\langle\psi|$ belonging to the dual space $\mathcal{H}^*$. However, the reverse is not necessarily true: there are bra vectors that do not necessarily have a corresponding ket vector, and therein lies the difference between bras and kets. It turns out that the difference only matters for Hilbert spaces of infinite dimension, in which case there can arise bra vectors whose corresponding ket vector is of infinite length, i.e. has infinite norm, and hence cannot be normalized to unity. Such ket vectors can therefore never represent a possible physical state of a system. But these issues will not arise here, so will not be of any concern. The point to be taken away from all this is that a bra vector is not the same kind of mathematical object as a ket vector. In fact, it has all the attributes of an operator in the sense that it acts on a ket vector to produce a complex number, this complex number being given by the appropriate inner product. This is in contrast to the more usual sort of operators encountered in quantum mechanics that act on ket vectors to produce other ket vectors. In mathematical texts a bra vector is usually referred to as a 'linear functional'. Nevertheless, in spite of the mathematical

distinction that can be made between bra and ket vectors, the correspondence between the two kinds of vectors is in most circumstances so complete that a bra vector equally well represents the state of a quantum system as a ket vector. Thus, we can talk of a system being in the state $\langle\psi|$.

We can summarize all this in the general case as follows: The inner product $(|\psi\rangle, |\phi\rangle)$ defines, for all states $|\psi\rangle$, the set of functions (or linear functionals) $(|\psi\rangle, \ )$. The linear functional $(|\psi\rangle, \ )$ maps any ket vector $|\phi\rangle$ into the complex number given by the inner product $(|\psi\rangle, |\phi\rangle)$.

1. The set of all linear functionals $(|\psi\rangle, \ )$ forms a complex vector space $\mathcal{H}^*$, the dual space of $\mathcal{H}$.
2. The linear functional $(|\psi\rangle, \ )$ is written $\langle\psi|$ and is known as a bra vector.
3. To each ket vector $|\psi\rangle$ there corresponds a bra vector $\langle\psi|$ such that if $|\phi_1\rangle \rightarrow \langle\phi_1|$ and $|\phi_2\rangle \rightarrow \langle\phi_2|$ then

$$c_1|\phi_1\rangle + c_2|\phi_2\rangle \rightarrow c_1^*\langle\phi_1| + c_2^*\langle\phi_2|.$$

8.6 State Spaces of Infinite Dimension

Some examples of physical systems with state spaces of infinite dimension were provided in the previous Section. In these examples, we were able to proceed, at least as far as constructing the state space was concerned, largely as was done in the case of finite dimensional state spaces. However, further investigation shows that there are features of the mathematics, and the corresponding physical interpretation in the infinite dimensional case that do not arise for systems with finite dimensional state spaces. Firstly, it is possible to construct state vectors that cannot represent a state of the system and secondly, the possibility arises of the basis states being continuously infinite. This latter state of affairs is not at all a rare and special case – it is just the situation needed to describe the motion of a particle in space, and hence gives rise to the wave function, and wave mechanics.

8.6.1 States of Infinite Norm

To illustrate the first of the difficulties mentioned above, consider the example of a system of identical photons in the state $|\psi\rangle$ defined by Eq. (8.66). As the basis states are orthonormal we have for $\langle\psi|\psi\rangle$

$$\langle\psi|\psi\rangle = \sum_{n=0}^{\infty} |c_n|^2 \tag{8.73}$$

If the probabilities $|c_n|^2$ form a convergent infinite series, then the state $|\psi\rangle$ has a finite norm, i.e. it can be normalized to unity. However, if this series does not converge, then it is not possible to supply a probability interpretation to the state vector as it is not normalizable to unity. For instance, if $c_0 = 0$ and $c_n = 1/\sqrt{n}, n = 1, 2, \ldots$, then

$$\langle\psi|\psi\rangle = \sum_{n=1}^{\infty} \frac{1}{n} \tag{8.74}$$

which is a divergent series, i.e. this state cannot be normalized to unity. In contrast, if $c_n = 1/n, n = 1, 2, \ldots$, then

$$\langle\psi|\psi\rangle = \sum_{n=0}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \tag{8.75}$$

which means we can normalize this state to unity by defining

$$|\tilde{\psi}\rangle = \frac{\sqrt{6}}{\pi}|\psi\rangle. \tag{8.76}$$

This shows that there are some linear combination of states that do not represent possible physical states of the system. Such states do not belong to the Hilbert space $\mathcal{H}$ of the system, i.e. the Hilbert space consists only of those states for which the coefficients c_n satisfy Eq. (8.73) [2]. This is a new feature: the possibility of constructing vectors that do not represent possible physical states of the system. It turns out that some very useful basis states have this apparently undesirable property, as we will now consider.

8.6.2 Continuous Basis States

If we now turn to the example in the previous section of the infinite crystal, we can consider what happens if we suppose that the separation between the neighbouring atoms in the crystal goes to zero, so that the electron can be found anywhere over a range extending from $-\infty$ to ∞. This, in effect, is the continuous limit of the infinite crystal model just presented, and represents the possible positions that a particle free to move anywhere in one dimension, the X axis say, can have. In this case, we could label the possible states of the particle by its X position, i.e. $|x\rangle$, where now, instead of having the discrete values of the crystal model, the position can now assume any of a continuous range of values, $-\infty < x < \infty$. It would seem that we could then proceed in the same way as we have done with the discrete states above, but it turns out that such states cannot be normalized to unity and hence do not represent (except in an idealised sense) physically allowable states of the system.

The aim here is to try to develop a description of the possible basis states for a particle that is not limited to being found only at discrete positions on the X axis. After all, in principle, we would expect that a particle in free space could be found at any position x in the range $-\infty < x < \infty$. We will get at this description by a limiting procedure which is not at all mathematically rigorous, but nevertheless yields results that turn out to be valid.

[2] Note however, that we can still construct a bra vector

$$\langle\psi| = \sum_{n=0}^{n=\infty} c_n^* \langle n|$$

without placing any restrictions on the convergence of the c_n's such as the one in Eq. (8.73). The corresponding ket cannot then represent a possible state of the system, but such inner products as $\langle\psi|\phi\rangle$ where $|\phi\rangle$ is a normalized ket can still be evaluated. The point being made here is that if $\mathcal{H}$ is of infinite dimension, the dual space $\mathcal{H}^*$ can also include bra vectors that do not correspond to normalized ket vectors in $\mathcal{H}$, which emphasizes the fact that $\mathcal{H}^*$ is defined as a set of linear functionals, and not simply as a 'complex conjugate' version of $\mathcal{H}$. The distinction is important in some circumstances, but we will not have to deal with such cases.

Suppose we return to the completeness relation for the states $|na\rangle$ for the one dimensional crystal

$$|\psi\rangle = \sum_{n=-\infty}^{+\infty} |na\rangle\langle na|\psi\rangle. \tag{8.77}$$

If we now put $a = \Delta x$ and $na = x_n$, BUT we write $|na\rangle = \sqrt{a}|x_n\rangle$, this becomes

$$|\psi\rangle = \sum_{n=-\infty}^{+\infty} |x_n\rangle\langle x_n|\psi\rangle\Delta x \tag{8.78}$$

where now

$$\langle x_n|x_m\rangle = \frac{\delta_{nm}}{a} \tag{8.79}$$

i.e. each of the states $|x_n\rangle$ is not normalized to unity, but we can nevertheless identify such a state as being that state for which the particle is at position x_n – recall if a state vector is multiplied by a constant, it still represents the same physical state of the system.

If we put to one side any concerns about the mathematical legitimacy of what follows, we can now take the limit $\Delta x \to 0$, i.e. $a \to 0$, then Eq. (8.78) can be written as an integral, i.e.

$$|\psi\rangle = \int_{-\infty}^{+\infty} |x\rangle\langle x|\psi\rangle\, dx \tag{8.80}$$

We can identify the state $|x\rangle$ with the physical state of affairs in which the particle is at the position x, and the expression Eq. (8.80) is consistent with the completeness requirement i.e. that the states $\{|x\rangle, -\infty < x < \infty\}$ form a complete set of basis states, so that any state of the one particle system can be written as a superposition of the states $|x\rangle$, though the fact that the label x is continuous has forced us to write the completeness relation as an integral. The difficulty with this procedure is that the states $|x\rangle$ are no longer normalized to unity. This we can see from Eq. (8.79) which tells us that $\langle x|x'\rangle$ will vanish if $x \neq x'$, but for $x = x'$ we see that

$$\langle x|x\rangle = \lim_{a\to 0} \frac{1}{a} = \infty \tag{8.81}$$

i.e. the state $|x\rangle$ has infinite norm! This means that there is a price to pay for trying to set up the mathematics in such a manner as to produce the completeness expression Eq. (8.80), which is that we are forced to introduce basis states which have infinite norm, and hence cannot represent a possible physical state of the particle! Nevertheless, provided care is taken, it is still possible to work with these states as if they represent the state in which the particle is at a definite position. To see this, we need to look at the orthonormality properties of these states, and in doing so we are lead to introduce a new kind of function, the Dirac delta function.

8.6.3 The Dirac Delta Function

We have just seen that the inner product $\langle x|x'\rangle$ vanishes if $x \neq x'$, but appears to be infinite if $x = x'$. In order to give some mathematical sense to this result, we return to Eq. (8.80) and look more closely at the properties that $\langle x|x'\rangle$ must have in order for the completeness relation also to make sense.

The probability amplitude $\langle x|\psi\rangle$ appearing in Eq. (8.80) are functions of the continuous variable x, and is often written $\langle x|\psi\rangle = \psi(x)$, which we identify as the wave function of the particle. If we now consider the inner product

$$\langle x'|\psi\rangle = \int_{-\infty}^{+\infty} \langle x'|x\rangle\langle x|\psi\rangle dx \tag{8.82}$$

or

$$\psi(x') = \int_{-\infty}^{+\infty} \langle x'|x\rangle\psi(x) dx \tag{8.83}$$

we now see that we have an interesting difficulty. We know that $\langle x'|x\rangle = 0$ if $x' \neq x$, so if $\langle x|x\rangle$ is assigned a finite value, the integral on the right hand side will vanish, so that $\psi(x) = 0$ for all x!! But if $\psi(x)$ is to be a non-trivial quantity, i.e. if it is not to be zero for all x, then it cannot be the case that $\langle x|x\rangle$ is finite. In other words, $\langle x'|x\rangle$ must be infinite for $x = x'$ in some sense in order to guarantee a non-zero integral. The way in which this can be done involves introducing a new 'function', the Dirac delta function, which has some rather unusual properties.

What we are after is a 'function' $\delta(x - x_0)$ with the property that

$$f(x_0) = \int_{-\infty}^{+\infty} \delta(x - x_0) f(x) dx \tag{8.84}$$

for all (reasonable) functions $f(x)$.

So what is $\delta(x - x_0)$? Perhaps the simplest way to get at what this function looks like is to examine beforehand a sequence of functions defined by

$$\begin{aligned} D(x, \epsilon) &= \epsilon^{-1} \qquad -\epsilon/2 < x < \epsilon/2 \\ &= 0 \qquad x < -\epsilon/2, x > \epsilon/2. \end{aligned} \tag{8.85}$$

What we first notice about this function is that it defines a rectangle whose area is always unity for any (non-zero) value of ϵ, i.e.

$$\int_{-\infty}^{+\infty} D(x,\epsilon)dx = 1. \tag{8.86}$$

Secondly, we note that as ϵ is made smaller, the rectangle becomes taller and narrower. Thus, if we look at an integral

$$\int_{-\infty}^{+\infty} D(x,\epsilon)f(x)dx = \epsilon^{-1}\int_{-\epsilon/2}^{\epsilon/2} f(x)dx \tag{8.87}$$

where $f(x)$ is a reasonably well behaved function (i.e. it is continuous in the neighbourhood of $x = 0$), we see that as $\epsilon \to 0$, this tends to the limit $f(0)$. We can summarize this by the equation

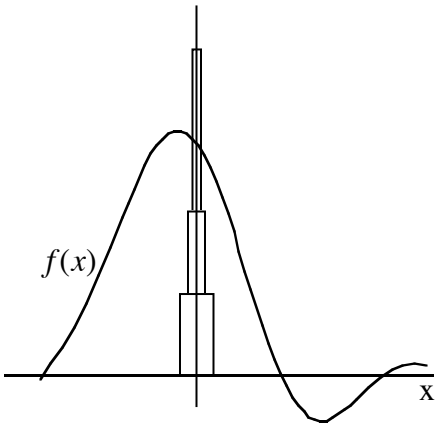

Figure 8.6: A sequence of rectangles of decreasing width but increasing height, maintaining a constant area of unity approaches an infinitely high 'spike' at $x = 0$..

$$\lim_{\epsilon\to 0}\int_{-\infty}^{+\infty} D(x,\epsilon)f(x)dx = f(0). \tag{8.88}$$

Taking the limit inside the integral sign (an illegal mathematical operation, by the way), we can write this as

$$\int_{-\infty}^{+\infty}\lim_{\epsilon\to 0} D(x,\epsilon)f(x)dx = \int_{-\infty}^{+\infty}\delta(x)f(x) = f(0) \tag{8.89}$$

where we have introduced the 'Dirac delta function' $\delta(x)$ defined as the limit

$$\delta(x) = \lim_{\epsilon\to 0} D(x,\epsilon), \tag{8.90}$$

a function with the unusual property that it is zero everywhere except for $x = 0$, where it is infinite.

The above defined function $D(x,\epsilon)$ is but one 'representation' of the Dirac delta function. There are in effect an infinite number of different functions that in an appropriate limit behave as the rectangular function here. Some examples are

$$\begin{aligned}\delta(x-x_0) &= \lim_{L\to\infty}\frac{1}{\pi}\frac{\sin L(x-x_0)}{x-x_0}\\ &= \lim_{\epsilon\to 0}\frac{1}{\pi}\frac{\epsilon}{(x-x_0)^2+\epsilon^2}\\ &= \lim_{\lambda\to\infty}\tfrac{1}{2}\lambda e^{-\lambda|x-x_0|}.\end{aligned} \tag{8.91}$$

In all cases, the function on the right hand side becomes narrower and taller as the limit is taken, while the area under the various curves remains the same, that is, unity.

The first representation above is of particular importance. It arises by via the following integral:

$$\frac{1}{2\pi}\int_{-L}^{+L} e^{ik(x-x_0)}dk = \frac{e^{iL(x-x_0)} - e^{iL(x-x_0)}}{2\pi i(x-x_0)} = \frac{1}{\pi}\frac{\sin L(x-x_0)}{x-x_0} \tag{8.92}$$

In the limit of $L \to \infty$, this then becomes

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{ik(x-x_0)}dk = \delta(x - x_0). \tag{8.93}$$

The delta function is not to be thought of as a function as it is usually defined in pure mathematics, but rather it is to be understood that a limit of the kind outlined above is implied whenever the delta function appears in an integral[3]. However, such mathematical niceties do not normally need to be a source of concern in most instances. It is usually sufficient to be aware of the basic property Eq. (8.84) and a few other rules that can be proven using the limiting process, such as

$$\begin{aligned}
\delta(x) &= \delta(-x) \\
\delta(ax) &= \frac{1}{|a|}\delta(x) \\
x\delta(x) &= 0 \\
\int_{-\infty}^{+\infty} \delta(x - x_0)\delta(x - x_1) &= \delta(x_0 - x_1) \\
\int_{-\infty}^{+\infty} f(x)\delta'(x - x_0)dx &= -f'(x_0).
\end{aligned}$$

The limiting process should be employed if there is some doubt about any result obtained. For instance, it can be shown that the square of a delta function cannot be given a satisfactory meaning.

Delta Function Normalization

Returning to the result

$$\psi(x') = \int_{-\infty}^{+\infty} \langle x'|x\rangle\psi(x)dx \tag{8.94}$$

we see that the inner product $\langle x'|x\rangle$, must be interpreted as a delta function:

$$\langle x'|x\rangle = \delta(x - x'). \tag{8.95}$$

The states $|x\rangle$ are said to be delta function normalized, in contrast to the orthonormal property of discrete basis states. One result of this, as has been pointed out earlier, is that states such as $|x\rangle$ are of infinite norm and so cannot be normalized to unity. Such states states cannot represent possible physical states of a system, though it is often convenient, with caution, to speak of such states as if they were physically realizable. Mathematical (and physical) paradoxes can arise if care is not taken. 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

$$|\psi\rangle = \int_{-\infty}^{+\infty} |x\rangle\langle x|\psi\rangle dx, \tag{8.96}$$

[3]This raises the question as to whether or not it would matter what representation of the delta function is used. Provided the function $f(x)$ is bounded over $(-\infty, +\infty)$ there should be no problem, but if the function $f(x)$ is unbounded over this interval, e.g. $f(x) = \exp(x^2)$, then only the rectangular representation of the delta function will give a sensible answer.

then

$$\langle\psi|\psi\rangle = \int_{-\infty}^{+\infty} \langle\psi|x\rangle\langle x|\psi\rangle dx. \tag{8.97}$$

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

$$\langle\psi|\psi\rangle = \int_{-\infty}^{\infty} |\psi(x)|^2 dx. \tag{8.98}$$

Provided $|\psi(x)|^2$ is a well behaved function, in particular that it vanish as $x \to \pm\infty$, this integral will converge to a finite result, so that the state $|\psi\rangle$ can indeed be normalized to unity, and if so, then we can interpret $|\psi(x)|^2 dx$ as the probability of finding the particle in the region $(x, x + dx)$, which is just the standard Born interpretation of the wave function.

8.6.4 Separable State Spaces

We have seen that state spaces of infinite dimension can be set up with either a denumerably infinite number of basis states, i.e. the basis states are discrete but infinite in number, or else a non-denumerably infinite number of basis states, i.e. the basis states are labelled by a continuous parameter. Since a state space can be spanned by more than one set of basis states, it is worthwhile investigating whether or not a space of infinite dimension can be spanned by a set of denumerable basis states, as well as a set of non-denumerable basis states. An example of where this is the case was given earlier, that of a particle in an infinitely deep potential well, see p95. It transpires that not all vector spaces of infinite dimension have this property, i.e. that they can have both a denumerable and a non-denumerable set of basis states. Vector spaces which can have both kinds of basis states are said to be separable, and in quantum mechanics it is assumed that state spaces are separable.

8.7 States of Macroscopic Systems

In the examples given above, it was assumed that an exhaustive list of results that could be obtained in the measurement of some observable of a quantum system could be used to set up the basis states for the state space of the system. The value of doing this is, of course, to be determine by the success or otherwise of these ideas. That quantum mechanics is such an overwhelmingly successful theory indicates that there is something correct in this procedure, but the question that arises is this: why does it not appear to work for macroscopic systems, i.e. for systems which we know can be fully adequately explained by standard classical physics? The answer appears to lie in the fact that in all the examples discussed above, whether or not the Hilbert space is of finite or infinite dimension, i.e. whether or not we are talking about spin up or spin down of a spin half particle, or the position of a particle in space, the implicit assumption is that the system we are considering is totally isolated from all other systems, in particular from any influence of the surrounding environment. After all, when we talked about a system, such as an O_2^- ion, we are ignoring all the other physical influences that could act on this system, i.e. we do not need to mention, in our specification of the state of the system, anything other than properties that directly pertain to the system of interest. The assumption is

made, as it is in classical physics, that such influences are sufficiently weak that they can be ignored to a good approximation. In effect, we are supposing that the systems under consideration are isolated systems, that is, systems that are isolated from the effect of any external perturbations.

Classically, at the macroscopic level, we can usually continue to ignore weak perturbing influences when specifying the state of a system. In fact, when defining a 'system' we typically include in what we refer to as the 'system', all the participants in the physical process being described that interact strongly with each other. Anything else that weakly affects these constitutents is ignored. For instance, when describing the orbital dynamics of the Earth as it revolves around the Sun, we might need to take into account the gravitational pull of the Moon – the system is the Earth, the Sun and the Moon. But we do not really need to take into account the effect of the background microwave radiation left over after the Big Bang. Or, when describing the collision between two billiard balls, it is probably necessary to include the effect of rolling friction, but it not really necessary to take into account the frictional drag due to air resistance. Of course, sometimes it is necessary to include external influences even when weak: to describe a system coming to thermal equilibrium with its surroundings it is necessary to extend the system by including the environment in the dynamical model. In any of these examples, the same classical physics methods and philosophy applies.

There is a subtle difference when it comes to trying to apply the quantum ideas developed so far to macroscopic systems. The same, weak perturbations that can be put to one side in a classical description of a macroscopic system turn out to have a far-reaching effect if included in a quantum description of the same system. If we were to attempt to describe a macroscopic system according to the laws of quantum mechanics, we would find that any linear superposition of different possible states of the system evolves on a fantastically short time scale to a classical mixture of the different possibilities. For instance, if we were to attempt to describe the state of a set of car keys in terms of two possibilities: in your pocket $|p\rangle$ or in your brief case $|b\rangle$, then a state of the form

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|p\rangle + |b\rangle) \tag{8.99}$$

could be used to represent a possible 'quantum state' of the keys. But this quantum state would be exceedingly short lived (on a time scale $\sim 10^{-40}$ sec), and would evolve into the two alternative possibilities: a 50% chance of the keys being in the state $|p\rangle$, i.e. a 50% chance of finding your keys in your pocket, and a 50% chance of being in the state $|b\rangle$, i.e. a 50% chance of finding them in your brief case. But this is no longer a superposition of these two states. Instead, the keys are *either* in the state $|p\rangle$ *or* the state $|b\rangle$. What this effectively means is that randomness is still there, i.e. repeating an experiment under identical conditions can give randomly varying results. But the state of the keys is no longer represented by a state vector, so there are no longer any quantum interference effects present. The randomness can then be looked upon as being totally classical in nature, i.e. as being due to our ignorance of information that is in principle there, but impossibly difficult to access. In effect, the quantum system behaves like a noisy classical system.

The process that washes out the purely quantum effects is known as decoherence. Since it is effectively impossible to isolate any macroscopic system from the influence of its

surrounding environment[4], all macroscopic systems (with very few exceptions such as superfluid SQUID junctions) are subject to decoherence. This process is believed to play a crucial role in why, at the macroscopic level, physical systems, which are *all* intrinsically quantum mechanical, behave in accordance with the classical laws of physics. It is also one of the main corrupting influences that prevent a quantum computer from functioning as it should. Quantum computers rely for their functioning on the 'qubits' remaining in linear superpositions of states, but the ever-present decohering effects of the environment will tend to destroy these delicate quantum states before a computation is completed, or else at the very least introduce errors as the computation proceeds. Controlling decoherence is therefore one of the major challenges in the development of viable quantum computers.

So, the bottom line is that it is only for protected isolated systems that quantum effects are most readily observed, and it is for microscopic systems that this state of affairs is to be found. But that is not to say that quantum effects are not present at the macroscopic level. Peculiar quantum effects associated with the superposition of states are not to be seen, but the properties of matter in general, and indeed the properties of the forces of nature, are all intrinsically quantum in origin.

[4] 'No man is an Iland, intire of itselfe' – J. Donne, *Devotions upon Emergent Occasions* Meditation XVII (1693)