5

Symmetry in Quantum Mechanics

We begin this chapter with a review, in section 5.1, of the basic concepts of quantum mechanics. It is assumed that the reader has met this subject before (see bibliography) and the main purpose of the review is to present those features which we shall need in illustrating the consequences of symmetry. Following a careful definition of symmetry in section 5.2 we proceed to study, one by one, these different consequences. By way of illustration, in section 5.6 we take a single particle moving in a fixed field with various simple symmetries, $C_3, D_3, S_3$, and $S_2$. The use of symmetry in approximate methods is described towards the end of the chapter. The later chapters of the book will deal with more complicated groups and more realistic physical systems but all the important consequences of symmetry in quantum systems are introduced in this chapter.

5.1 Brief review of the framework of quantum mechanics

In quantum mechanics the behaviour of a system with $n$ degrees of freedom
Symmetry in Quantum Mechanics

Described by some coordinates \((r_1, r_2, \ldots, r_n) \equiv r\) is completely specified by its wave function (or state vector) \(\psi(r, t)\). This wave function takes on a value, which is generally a complex number, for each value of \(r\) and \(t\). The real number \(\int \psi(r, t) \overline{\psi(r, t)} \, dV\) is interpreted as the probability that, at time \(t\), the coordinates have values lying in volume \(dV\) at \(r\). The normalisation of the wave functions must therefore be chosen such that the total probability of being in some position is equal to unity, i.e.,

\[
\int \vert \psi(r, t) \vert^2 \, dV = 1 \tag{5.1}
\]

where the integral runs over all values of the coordinates.

The wave functions are found by solving the Schrödinger equation

\[
H(r, t) \psi(r, t) = i\hbar \frac{\partial}{\partial t} \psi(r, t) \tag{5.2}
\]

where \(H(r, t)\) is the Hamiltonian (operator) corresponding to the classical Hamiltonian \(H = T + V\), where \(T\) and \(V\) are the kinetic and potential energies, respectively. There are general rules for obtaining a quantum mechanical operator from its classical analogue—for example the Cartesian coordinate \(x\) of a particle is replaced by the operator \(x\), while the conjugate momentum \(p_x\) is replaced by the operator \(-i\hbar \frac{\partial}{\partial x}\). The boundary conditions imposed on the differential equation \(5.2\), for example that \(\psi\) should vanish at the spatial boundaries, are sufficient to ensure that \(\psi(r, t)\) is uniquely determined for all time by equation \(5.2\) if the function \(\psi(r, t_0)\) is known at some time \(t_0\).

Many important problems in quantum mechanics are described by a time-independent Hamiltonian. In such cases it follows immediately from equation \(5.2\) that there are solutions in which the time dependence appears as a factor

\[
\psi_E(r, t) = \psi_E(r) \exp(-iEt/\hbar) \tag{5.3}
\]

where \(\psi_E(r)\) is a solution (eigenfunction) of the time-independent eigenvalue equation

\[
H\psi_E(r) = E\psi_E(r) \tag{5.4}
\]

and \(E\) is the corresponding eigenvalue. The infinite set of eigenfunctions \(\psi_E(r)\) is a complete set in the sense that any continuous function with the same boundary conditions may be expanded as an infinite series in the \(\psi_E(r)\). Hence any solution of the Schrödinger equation may be expanded

\[
\psi(r, t) = \sum_E \psi_E(r) \exp(-iEt/\hbar) \tag{5.5}
\]

The simple solutions \(5.3\) are called ‘stationary’ states, since the probability density \(\vert \psi_E(r, t) \vert^2\) is time independent. The eigenvalue \(E\) is interpreted physically as the energy of the system.

The set of all continuous functions satisfying the boundary conditions imposed on equation \(5.2\) form a function space, usually called the Hilbert
5.1 Symmetry in Quantum Mechanics

space, and it is convenient to define a scalar product of two functions $\psi$ and $\phi$ by the integral

$$
(\psi, \phi) = \int \phi^*(r) \psi(r) dV
$$

(5.6)

taken over all possible values of the coordinates $r$. The Hamiltonian must be Hermitian with respect of this definition of scalar product in order that the energies $E$ shall be real. As a consequence, the set of eigenfunctions $\psi_E(r)$ are orthogonal with respect to this scalar product and therefore provide a convenient basis. In accordance with equation (5.1) it is usual to define the $\psi_E(r)$ to be normalised to unity. Notice that for a stationary state the condition (5.1) becomes $(\psi_E, \psi_E) = 1$.

To each classical observable $O$, like the energy, momentum, angular momentum, etc., there is a corresponding Hermitian operator $O$ (usually time independent) which operates in the Hilbert space of the wave functions. Each operator will have its own set of eigenfunctions and eigenvalues $O\psi_\lambda^O(r) = \lambda_\lambda^O \psi_\lambda^O(r)$ and each set $\psi_\lambda^O(r)$ will form an orthonormal basis. The only definite value resulting from a measurement of the observable $O$ are the eigenvalues $\lambda$ and a particular value $\lambda$ will be observed only if the wave function of the system at that particular time is precisely the corresponding eigenfunction $\psi_\lambda^O(r)$ of $O$. In general, the wave function $\psi(r, t)$ will not coincide with any particular eigenfunction of $O$ but, because of the completeness of the $\psi_\lambda^O(r)$, we can always expand

$$
\psi(r, t) = \sum_\lambda c_\lambda(t) \psi_\lambda^O(r)
$$

(5.7)

The result of a measurement of the observable $O$ in such a state is not completely determined but $|c_\lambda(t)|^2$ is the probability of obtaining the value $\lambda$. The mean value or, as it is usually called, 'the expectation value' of $O$ in the state $\psi$ is given by the sum $\sum_\lambda c_\lambda(t)^2 \lambda$. This expectation value may be expressed very conveniently as a scalar product since

$$
(\psi, O\psi) = \sum_\lambda \sum_\lambda c_\lambda^*(t) c_\lambda(t) \lambda (\psi_\lambda^O, \psi_\lambda^O) = \sum_\lambda |c_\lambda(t)|^2 \lambda
$$

(5.8)

using the orthonormality of the eigenfunctions $\psi_\lambda^O(r)$. It is customary in quantum mechanics to use the notation $\langle \psi | O | \psi \rangle$ for the scalar product $\langle \psi, O\psi \rangle$. Notice that the stationary states give precise values $\lambda$ for a measurement of the energy, the corresponding operator being the Hamiltonian. To be consistent in our notation we should have used $\psi_E^* | \psi_E \rangle$ for the stationary states but for brevity we used simply $\psi_E$, in view of the special significance of the energy operator $H$.

Assuming that the spectrum of an operator (the set of eigenvalues) is discrete we have the phenomenon of quantisation in which the measured values of the
operator are restricted to that discrete set of eigenvalues. For some operators it may happen that the spectrum is continuous, in which case the expansion (5.7) becomes

$$\psi(r, t) = \int c(\lambda, t)\psi^0(\lambda, r)d\lambda$$

where the integration runs over the possible range of eigenvalues \(\lambda\).

From the experimental point of view one may investigate a quantum system by measuring its energies and the expectation values of various operators. These can then be compared with the values calculated from some assumed Hamiltonian. Another very important kind of experiment is the measurement of the transition rate, or transition probability, for a process in which a system changes from an initial state \(\psi_i\) to a final state \(\psi_f\). It may be argued that the transition probability \(W_{if}\) is proportional to the square of an off-diagonal matrix element

$$W_{if} \propto |\langle \psi_f, O\psi_i \rangle|^2 \quad (5.9)$$

where \(O\) is the operator relevant to the particular process. It is, of course, by the measurement of energies, expectation values and transition probabilities that one learns about the symmetry of a quantum system and in the subsequent sections of this chapter we shall show how these quantities are affected by symmetry. The interaction of a system, for example an atom or atomic nucleus, with an electromagnetic field is a particularly convenient and fruitful source of information. It provides both expectation values, through the perturbation of energies, and transition probabilities through the absorption and emission of radiation. Without going into the theory in any detail, we now describe the main features of the electromagnetic interaction and list the various operators which occur.

As one might expect from simple classical ideas, the interaction of a system with an electromagnetic field depends on the positions and momenta of the constituent charges. Next to the total charge of the system, the most important quantities are what are known as the electric dipole and magnetic dipole moments, which for a system of particles \(i\) with charges \(e_i\) are given by \(\sum e_i r_i\) and \(\sum (e_i/2Mc)r_i \wedge p_i\), respectively. (For particles with spin there is an additional contribution to the magnetic moment, see section 8.4.) In an atom, where all the \(Z\) electrons have the same charge, these expressions simplify to

$$-ZeR \quad \text{and} \quad -(e/2Mc)L$$

where \(R\) is the centre-of-mass vector and \(L\) the total angular momentum. These operators are involved both in the energy perturbations and in the transition probabilities for the quantum system. In particular, the energy shift due to a uniform magnetic field \(B\) is given by \(eL \cdot B/2Mc\). For transitions, the electric dipole process is dominant in atoms, and for unpolarised radiation all three components of \(R\) will contribute. However, for the absorption of polarised light, only certain components of \(R\) are relevant. For example, in what is known as plane-polarised light, the
electric field vector points in a fixed direction $\mathbf{p}$ which is perpendicular to the light beam. The relevant operator is then the component of $\mathbf{R}$ in the direction $\mathbf{p}$ and we refer to this case as polarisation in the $\mathbf{p}$-direction.

### 5.2 Definition of symmetry in a quantum system

Consider a system governed by a time-independent Hamiltonian $H(r)$ with $\psi(r)$ denoting an arbitrary wave function—not necessarily an eigenfunction. As in the previous section, the coordinate vector $r$ describes the $n$ degrees of freedom of the system. Any group $G$ of coordinate transformations $r \to r'$ will define a corresponding set of induced transformations $T(G_a)$ in the space of wave functions through the definition (3.37):

$$T(G_a) \psi(r) = \psi'(r) = \psi(G_a^{-1} r')$$

It also defines, through equation (3.24), the transformed Hamiltonian operator $T(G_a)HT^{-1}(G_a) = H'$. If the Hamiltonian is invariant under these transformations, i.e. if

$$T(G_a)HT^{-1}(G_a) = H$$

for all $G_a$ in the group $G$, then $G$ is said to be a symmetry group of the Hamiltonian. We shall soon see how the existence of a symmetry group leads to a variety of important consequences. By multiplying on the right by $T$, the condition (5.10) takes the equivalent form $T(G_a)H - HT(G_a) = 0$, or $[T(G_a), H] = 0$, i.e. the Hamiltonian commutes with all the induced transformations of the group.

The group elements $G_a$ of a symmetry group are called symmetry elements. In practice, if one can find one or two symmetry elements of a Hamiltonian, a symmetry group may be generated by multiplying them together and taking powers until no new elements are produced. (It follows immediately from equation (5.10) that the product of two symmetry elements is another symmetry element.) It is, however, rather difficult to be sure that one has found all the symmetry elements of a given Hamiltonian and not simply the elements belonging to a subgroup of the full symmetry group.

For many transformations, such as rotations, reflections and translations, the kinetic energy operator is an invariant so that the invariance of the Hamiltonian rests on the condition

$$V(r) = V(G_a r)$$

on the potential energy $V$.

In the presentation above, we have been fairly explicit in the definition of the induced transformation but it is possible to imagine a more abstract set of transformations $T(G_a)$ in the function space which are not related to coordinate transformations but which nevertheless satisfy the multiplication rules for a group and therefore define a symmetry group.
5.3 Degeneracy and the labelling of energies and eigenfunctions

The first two consequences of the existence of a symmetry group $\mathcal{G}$ of the Hamiltonian are intimately related and we discuss them together. First we state the results and then show how they are derived.

1. The eigenfunctions $\psi$ and eigenvalues $E$ of $H$ may be labelled by the irreducible representations $T(\alpha)$ of the symmetry group $\mathcal{G}$, so that we write $\psi^{\alpha}$ and $E^{\alpha}$.

2. The energy $E^{\alpha}$ will be (at least) $s_\alpha$-fold degenerate, where $s_\alpha$ is the dimension of $T(\alpha)$.

To derive these results we first note that the set of all degenerate eigenfunctions at some chosen eigenvalue $E$ forms a vector space $V$. (If $\phi$ and $\psi$ both have energy $E$ then so does any linear combination of them.) Next, we argue that $V$ must be an invariant space with respect to the transformations $T(G_\alpha)$ induced by $\mathcal{G}$. To show the invariance we simply define the transformed function $\psi' = T(G_\alpha)\psi$ in the usual way and then deduce that, if $\psi$ is an eigenfunction of $H$ with energy $E$, then so is $\psi'$, because

$$H\psi' = HT(G_\alpha)\psi = T(G_\alpha)H\psi = ET(G_\alpha)\psi = E\psi'$$

where we have made use of equation (5.10). The vector space $V$ therefore provides a representation of the symmetry group through the transformations $T(G_\alpha)$. This representation is either irreducible or may be reduced into its irreducible constituents so that in either case a basis for $V$ may be chosen in which the basis vectors are written as $\psi^{\alpha,i}$, being labelled by an irreducible representation label $\alpha$ of the group $\mathcal{G}$ and by a row label $i$. The degeneracy is given by the dimension of $V$ which is at least $s_\alpha$ and will be greater if $V$ is reducible.

The argument above justifies the results given at the beginning of this section but leaves one important point to be clarified. If the representation provided by $V$ is not irreducible then that energy will carry not one but a number of labels $\alpha$ corresponding to the irreducible representations which occur in its reduction. In practice, if one is using the full symmetry group, it is very rare to find such a situation—as rare as it would be to find equal eigenvalues for an arbitrary matrix. One therefore uses the expression 'normal degeneracy' for one described by an irreducible representation and 'accidental degeneracy' when two or more irreducible representations occur at the same energy. An accidental degeneracy can be produced by the variation of some parameter in $H$ (such as a magnetic or electric field strength) so that, for some particular values of this parameter, two normally non-degenerate levels cross. If, for a particular Hamiltonian, one finds systematic accidental degeneracies when using a symmetry group $\mathcal{G}$ there is usually an explanation in the existence of a larger symmetry group. As we have seen in section 4.18, an irreducible

\footnote{The symbol $V$ used on this page for the vector space should not be confused with the same symbol $V$ which was used for the potential energy on the previous page.}
representation of a group \( G \) will generally reduce into a sum of several irreducible representations of any subgroup. Thus the accidentally degenerate levels are brought together within a single irreducible representation of the larger group. Some famous examples of this are given in chapter 19 of volume 2. Examples of the much more important case of normal degeneracy are given in section 5.6.

### 5.4 Selection rules and matrix elements of operators

In a system with a symmetry group \( G \) we have seen that the eigenfunctions \( \psi \) of the Hamiltonian may be labelled by the irreducible representations of \( G \). In section 4.20 we saw that if an operator \( O^{(1)} \) was defined to transform according to an irreducible representation \( T^{(1)} \) then its matrix elements, taken between functions which also transform irreducibly under the same group, have some simple features. It is therefore clear that we shall be able to draw some simple conclusions about the physically interesting matrix elements if we analyse the operators into their irreducible parts. The most striking result is that if a particular transition process is governed by an operator \( O^{(i)} \) then a transition from a state \( \psi^{(i)} \) may take place only to those final states \( \psi^{(j)} \) for which the representation \( T^{(j)} \) occurs in the reduction (4.44):

\[
T^{(i)} \times T^{(j)} = \sum_{\ell} m_{\ell} T^{(\ell)}
\]

Thus transitions from the state \( \psi^{(i)} \) via the operator \( O^{(i)} \) select those final states \( \psi^{(j)} \) which transform according to any one of the representations on the right-hand side of equation (4.44). For other final states, the matrix element vanishes and no transition can take place. One refers to this phenomenon as a 'selection rule' and transitions which cannot take place are said to be 'forbidden'. Notice that a selection rule does not tell you that a transition will occur but only that, so far as the symmetry is concerned, it may occur. In particular cases there can always be other reasons which cause particular matrix elements to vanish. To find the full set of selection rules one must take the irreducible representations \( T^{(\ell)} \) one by one and deduce the coefficients \( m_\ell \) from equation (4.45) using the known character table for the symmetry group. The transition may then take place for every \( \ell \) for which \( m_\ell \neq 0 \).

So far we have assumed that the transition is governed by an operator \( O^{(i)} \) transforming according to an irreducible representation \( x \). In practice it may happen that the operator \( O \) transforms according to a reducible representation. In this case we may either analyse it into its irreducible parts

\[
O = \sum_x O^{(x)}
\]

and deduce the transitions allowed separately for each irreducible component, or, if we know the character \( \chi(G_x) \) of the 'reducible' representation directly.
then the coefficients \( m \) are given by

\[
m_i = \frac{1}{d} \sum \chi^i(G_w) \chi(G_w) \chi^i(G_w)
\]

in place of equation (4.45).

The operators \( R = (X, Y, Z) \) for electric dipole transitions transform like a vector under rotations and are odd under inversion. They thus form a three-dimensional representation of any group of rotations and inversions. For a proper rotation about the \( z \)-axis, through an angle \( \theta \), the character is simply \((2 \cos \theta + 1)\) from equation (4.6) and, since the trace is independent of basis, this is also the character for rotation about any axis. In the same way for magnetic dipole transitions, the relevant operator \( L = \sum r_i \wedge p_i \) is also a vector and therefore has the same character for rotations. However, whereas \( R \) changes sign under inversion, \( L \) does not, since both \( r \) and \( p \) change sign, so that when improper rotations are considered, the characters for \( R \) and \( L \) differ in sign. One often refers to these two three-dimensional representations of any group of rotations and inversions as the vector and pseudo- (or axial) vector representations, respectively.

The Wigner–Eckart theorem, equation (4.62), goes beyond the selection rule to relate matrix elements which have the same \( \alpha \), \( \beta \) and \( \gamma \) but differ in their row labels \( i \), \( j \) and \( k \)—see equation (4.62). In other words it tells how a matrix element changes when one of the wave functions is replaced by another member of the degenerate multiplet. Since the simple transition processes involve a sum over any degenerate final states, this feature of symmetry is of less importance. We leave any further discussion of applications of the Wigner–Eckart theorem until chapter 8.

### 5.5 Conservation laws

An observable \( O \) is said to be conserved in a given system if its mean value in any state \( \psi(r, t) \) of the system does not vary with time. An equivalent statement is that if, at some time, the wave function of the system coincides with an eigenfunction of \( O \), then it will remain an eigenfunction of \( O \) with the same eigenvalue for all time. We now show that an operator \( O \) has this conservation property if it commutes with the Hamiltonian and is also time independent. We have

\[
\frac{d}{dt} \langle \psi | O \psi \rangle = \left( \frac{\partial \psi}{\partial t} , O \psi \right) + \left( \psi , O \frac{\partial \psi}{\partial t} \right) = \{ - [H\psi] + [\psi, OH] \} / \hbar = 0
\]

(5.13)
using the Schrödinger equation (5.2) and the assumed commutation property 
\[ [H, O] = 0. \]

Now, if the system possesses symmetry, the operators \( T(G_a) \) for any \( G_a \) in \( \mathcal{G} \) will commute with the Hamiltonian and are therefore conserved. At first 
sight this suggests a large number of conserved quantities but they are not 
indeed independent. The conservation of \( T(G_a) \) is clearly not independent of the 
conservation of \( T(G_b) \) and \( T(G_c) \) if \( G_a G_b = G_c \). There is therefore a 
minimum number of independent conserved quantities for each group corresponding to 
the minimum number of group elements needed to generate all group elements 
by multiplication amongst themselves. Thus for a cyclic group, like \( C_p \), there is 
only one conserved quantity since, by definition of a cyclic group, all elements 
are obtained as powers of a single element. For the group \( D_3 \), see table 2.5, one 
needs two elements \( R_1 \) and \( R_2 \) to generate all others and so there are two 
independent conserved quantities. For the continuous group \( \mathcal{R}_2 \), it will be seen 
in subsection 7.3.5 that all group elements may be generated from a single 
operator.

It should also be remembered from the postulates of quantum mechanics 
that only Hermitian operators may represent physical observables. Thus for a 
physically observable conserved quantity one must construct a Hermitian 
operator from the \( T(G_a) \). We shall see that only for the continuous groups do 
these conserved quantities correspond to familiar classical concepts like 
momentum and angular momentum.

5.6 Examples

To illustrate these consequences of symmetry we now consider four examples, 
corresponding to symmetry groups \( C_3, D_3, S_3 \) and \( \mathcal{R}_2 \). Although it is not 
necessary to specify the wave function in detail (this is the strength of the 
symmetry argument) one may imagine these examples to relate to the motion 
of a single particle in a potential \( V(r) \). The kinetic energy operator \( V^2 \) being a 
scalar product, is invariant with respect to any rotation and is therefore an 
invariant in all four examples. On the other hand, the potential may be chosen 
to be invariant only with respect to one of the more restricted rotation groups 
listed above. This simple problem is not entirely artificial since it can represent 
the motion of an electron around the nucleus of an atom. Although the 
Coulomb attraction to the nucleus is spherically symmetric one may imagine 
an externally applied field which reduces the symmetry. Alternatively, if the 
atom is in a crystal then the presence of other atoms of the crystal will reduce 
the symmetry from \( \mathcal{R}_3 \) to some finite group of rotations.

5.6.1 Symmetry group \( C_3 \)

For this group the potential \( V(r) \) is invariant under three-fold rotations \( R_1 \) 
and \( R_2 = R_1^2 \) about the \( z \)-axis, using the notation of section 2.2(c). The
group $C_3$ is a cyclic group and the three one-dimensional irreducible representations $\tau^{(1)}$, $\tau^{(2)}$ and $\tau^{(3)}$, given in table 4.3 of section 4.18, may be written

\[ \tau^{(a)}(R_1) = \exp[2\pi(i\alpha - 1)/3] \]

(5.14)

with $\alpha = 1, 2$ and 3. From the general results of section 5.3 we can therefore deduce that the eigenstates will be non-degenerate, can be labelled by the irreducible representation label $\alpha$ and will have the property $T(R_1)\psi^{(a)} = \exp[2\pi(i\alpha - 1)/3]\psi^{(a)}$. In our particular example of a single particle this implies, from the definition (3.37), that

\[ \psi^{(a)}(\phi - \frac{2\pi}{3}) = \exp[2\pi(i\alpha - 1)/3] \psi^{(a)}(\phi) \]

(5.15)

where $\phi$ is the usual polar coordinate around the $z$-axis. (Without the direct use of group theory we could have derived this result by noticing that the Hamiltonian $H$ commutes with the three-fold rotation operator $\tau(R_1)$ and hence the eigenstates of $H$ can be chosen to make $\tau(R_1)$ diagonal and can be labelled by the eigenvalues of $\tau(R_1)$ which are given by equation (5.14). The operator $\tau(R_1)$ represents the conserved quantity for $C_3$ but it has no simple physical interpretation.)

The symmetry properties of functions which satisfy equation (5.15) are seen more clearly by writing

\[ \psi^{(a)}(\phi) = u_\alpha(\phi)\exp[-i(\alpha - 1)\phi] \]

(5.16)

where $u_\alpha(\phi)$ is a periodic function of $\phi$ with period $2\pi/3$. This may be justified by substituting (5.16) in equation (5.15) whence

\[ u_\alpha(\phi - \frac{2\pi}{3})\exp[-i(\alpha - 1)[\phi - (2\pi/3)]] = \exp[2\pi(i\alpha - 1)/3]u_\alpha(\phi)\exp[-i(\alpha - 1)\phi] \]

i.e.

\[ u_\alpha(\phi - \frac{2\pi}{3}) = u_\alpha(\phi) \]

(5.17)

Equation (5.16) gives $\psi^{(a)}(\phi)$ for all $\phi$ if it is known for the range $0 \leq \phi < 2\pi/3$. Symmetry arguments thus lead to the conclusion that, to solve Schrodinger's equation for a system with $C_3$ symmetry, it is only necessary to solve an equivalent equation for $u_\alpha(\phi)$ for the range $0 \leq \phi < 2\pi/3$ with periodic boundary conditions.

We will now study the selection rules for electric dipole transitions between the eigenstates $\psi^{(a)}(\phi)$. For light polarised along the $z$-axis the appropriate operator (see section 5.1) is $\epsilon z$. Clearly this is invariant under the three-fold rotations and hence belongs to the identity representation $\tau^{(1)}$. Since the product representation $\tau^{(1)} \times \tau^{(a)} = \tau^{(a)}$ this operator can only cause transitions

\[ \boxed{} \]