

2 The elements of quantum mechanics

In Chapter 1, we discussed some of the evidence for the atomic nature of matter. We also learned that the classical Newtonian form of mechanics could not describe phenomena on the atomic scale. In particular, we saw that experiments involving the diffraction of electrons or atoms by crystals demonstrate that particles exhibit wave properties. Equally, experimental evidence for the quantisation of the radiation field, from phenomena as different as black body radiation, the photoelectric effect and the Compton effect, points to a fundamental wave-particle duality which must be taken into account in satisfactory theories of both matter and radiation. For material particles, such a theory was developed in the years 1925 and 1926 by W. Heisenberg, M. Born and P. Jordan in a form known as *matrix mechanics*. An equivalent form of the theory, called *wave mechanics*, was proposed at the same time by E. Schrödinger, following the ideas put forward in 1924 by L. de Broglie. A more abstract form of quantum mechanics, which includes both matrix mechanics and wave mechanics, was published by P.A.M. Dirac in 1930.

In this chapter, we shall outline the main results and approximation methods of quantum mechanics, which will be used in the detailed discussions of atomic and molecular phenomena following in later chapters. In the space available, only those aspects of the subject which find immediate application will be discussed. For detailed accounts of quantum theory, we refer the reader to the texts of Dirac (1958), Landau and Lifshitz (1965), Messiah (1968), Schiff (1968), Merzbacher (1998) and Bransden and Joachain (2000).

2.1

Waves and particles, wave packets and the uncertainty principle

The experiments on the corpuscular nature of the electromagnetic radiation, which we discussed in the previous chapter, require that with the electromagnetic field we must associate a particle, the photon, whose energy E and magnitude of momentum p are related to the frequency ν and wavelength λ of the electromagnetic radiation by

$$E = h\nu, \quad p = h/\lambda \quad (2.1)$$

On the other hand, the experiments on the wave-like properties of particles, also discussed in Chapter 1, imply that we associate with each particle a wave or

matter field, the de Broglie relations which link the frequency ν and the wavelength λ of the wave with the particle energy E and magnitude of momentum p being also given by (2.1). As a consequence, we can assume that the relations (2.1) hold for all types of particles and field quanta. Introducing the angular frequency $\omega = 2\pi\nu$, the wave number $k = 2\pi/\lambda$ and the reduced Planck constant $\hbar = h/(2\pi)$, we may write the relations (2.1) in the more symmetric form

$$E = \hbar\omega, \quad p = \hbar k \quad (2.2)$$

As a first step in formulating non-relativistic wave mechanics for material particles, let us consider a free particle of mass m , having a well-defined momentum $\mathbf{p} = p_x \hat{x}$ directed along the positive X direction and a non-relativistic energy $E = p_x^2/(2m)$. Guided by (2.2), we associate with this particle a plane wave

$$\Psi(x, t) = A \exp[i(kx - \omega t)] \quad (2.3)$$

where A is a constant. This plane wave travels in the positive X direction, has a wave number $k = 2\pi/\lambda = p_x/\hbar$ and an angular frequency $\omega = E/\hbar$ which we may also write as

$$\omega = \frac{\hbar k^2}{2m} \quad (2.4)$$

The function $\Psi(x, t)$ is known as a *wave function*, and we shall discuss its significance shortly. For the moment, we note that for a free particle, represented by the plane wave (2.3), we have

$$-i\hbar \frac{\partial}{\partial x} \Psi = p_x \Psi \quad (2.5)$$

and

$$i\hbar \frac{\partial}{\partial t} \Psi = E \Psi \quad (2.6)$$

This one-dimensional treatment is easily extended to three dimensions. To a free particle of mass m , having a well-defined momentum $\mathbf{p}$ and an energy $E = p^2/(2m)$, we now associate a plane wave

$$\Psi(\mathbf{r}, t) = A \exp[i(\mathbf{k} \cdot \mathbf{r} - \omega t)] \quad (2.7)$$

where the *propagation vector* $\mathbf{k}$ is related to the momentum $\mathbf{p}$ by

$$\mathbf{p} = \hbar \mathbf{k} \quad (2.8)$$

with

$$k = |\mathbf{k}| = \frac{|\mathbf{p}|}{\hbar} = \frac{2\pi}{\lambda} \quad (2.9)$$

The vector $\mathbf{r}$ is the position vector of the particle and the angular frequency ω is still given by $\omega = E/\hbar = \hbar k^2/(2m)$. The equation (2.6) remains unchanged for the plane wave (2.7), while (2.5) is now replaced by its obvious generalisation

$$-i\hbar \nabla \Psi = \mathbf{p} \Psi \quad (2.10)$$

where ∇ is the gradient operator, having Cartesian components $(\partial/\partial x, \partial/\partial y, \partial/\partial z)$. The relations (2.6) and (2.10) show that for a free particle the energy and momentum can be represented by the differential operators

$$E_{\text{op}} = i\hbar \frac{\partial}{\partial t}, \quad \mathbf{p}_{\text{op}} = -i\hbar \nabla \quad (2.11)$$

acting on the wave function Ψ . It is a *postulate* of wave mechanics that when the particle is not free the dynamical variables E and $\mathbf{p}$ are still represented by these differential operators.

The plane waves (2.3) or (2.7) represent particles having a definite momentum, but since their amplitude is constant, these plane waves correspond to a complete absence of localisation of the particle in space. To describe a particle which is confined in a certain spatial region, a *wave packet* can be formed by superposing plane waves of different wave numbers. For example, in order to describe a free particle which is confined in a region of the X axis, we shall superpose plane waves of the form (2.3) to obtain the wave packet

$$\Psi(x, t) = (2\pi\hbar)^{-1/2} \int_{-\infty}^{+\infty} \exp[i(p_x x - Et)/\hbar] \phi(p_x) dp_x \quad (2.12)$$

where the factor in front of the integral has been chosen for future convenience. Writing $\psi(x) \equiv \Psi(x, t=0)$ we see that the functions

$$\psi(x) = (2\pi\hbar)^{-1/2} \int_{-\infty}^{+\infty} \exp(ip_x x/\hbar) \phi(p_x) dp_x \quad (2.13a)$$

and

$$\phi(p_x) = (2\pi\hbar)^{-1/2} \int_{-\infty}^{+\infty} \exp(-ip_x x/\hbar) \psi(x) dx \quad (2.13b)$$

are just *Fourier transforms* of each other. More generally, we write at time t

$$\Psi(x, t) = (2\pi\hbar)^{-1/2} \int_{-\infty}^{+\infty} \exp(ip_x x/\hbar) \Phi(p_x, t) dp_x \quad (2.14a)$$

and

$$\Phi(p_x, t) = (2\pi\hbar)^{-1/2} \int_{-\infty}^{+\infty} \exp(-ip_x x/\hbar) \Psi(x, t) dx \quad (2.14b)$$

so that the functions $\Psi(x, t)$ and $\Phi(p_x, t)$ are also mutual Fourier transforms. The functions $\Phi(p_x, t)$ and $\phi(p_x) \equiv \Phi(p_x, t=0)$ are called *wave functions in momentum space*.

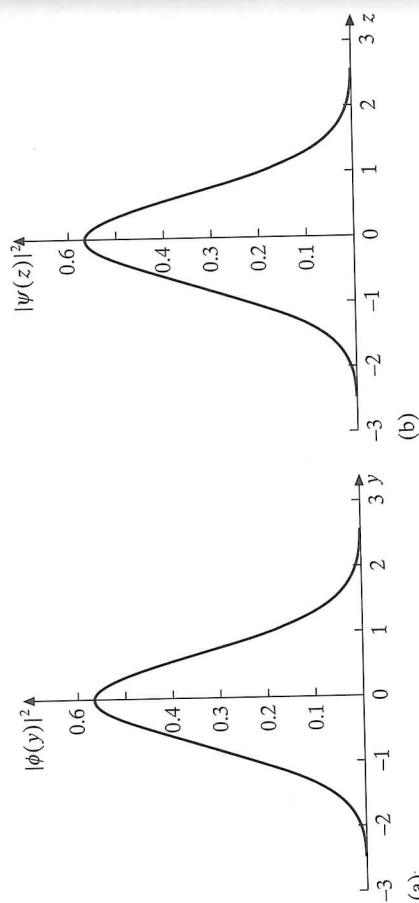


Figure 2.1 (a) The function $|\phi(y)|^2 = \pi^{-1/2} \exp(-y^2)$, where $y = (p_x - p_0)/\Delta p_x$. (b) The function $|\psi(z)|^2 = \pi^{-1/2} \exp(-z^2)$, where $z = (\Delta p_x/\hbar)x$. Note that the functions $\phi(y)$ and $\psi(z)$ have been normalised so that $\int_{-\infty}^{+\infty} |\phi(y)|^2 dy = \int_{-\infty}^{+\infty} |\psi(z)|^2 dz = 1$.

The Heisenberg uncertainty principle

Consider the case for which the wave function $\phi(p_x)$ is localised in a certain region of the p_x variable. As a simple example, we shall assume that $\phi(p_x)$ is the Gaussian function

$$\phi(p_x) = \exp\left[-\frac{(p_x - p_0)^2}{2(\Delta p_x)^2}\right] \quad (2.15)$$

where Δp_x , which we shall call the 'width' of the distribution in p_x , is a constant such that $|\phi(p_x)|^2$ drops to $1/e$ of its maximum value at $p_x = p_0 \pm \Delta p_x$ (see Fig. 2.1(a)). Using the known result

$$\int_{-\infty}^{+\infty} \exp(-\alpha u^2) \exp(-\beta u) du = \left(\frac{\pi}{\alpha}\right)^{1/2} \exp[\beta^2/(4\alpha)] \quad (2.16)$$

we find from (2.13) that

$$\psi(x) = \hbar^{-1/2} \Delta p_x \exp(ip_0 x/\hbar) \exp[-(\Delta p_x)^2 x^2/(2\hbar^2)]. \quad (2.17)$$

Apart from the phase factor $\exp(ip_0 x/\hbar)$, this function is again a Gaussian. We remark that $|\psi(x)|^2$ has a maximum value at $x = 0$ and falls to $1/e$ of its maximum value at $x = \pm \Delta x$, where $\Delta x = \hbar/\Delta p_x$ is the 'width' of the distribution in the x variable (see Fig. 2.1(b)). Given the above definitions of the 'widths' Δx and Δp_x , we see that for a Gaussian wave packet $\Delta x \Delta p_x = \hbar$. Thus, if we decrease Δp_x , so that the wave function in momentum space, $\phi(p_x)$, is more sharply peaked about $p_x = p_0$, then Δx will increase and $\psi(x)$ becomes increasingly 'delocalised'. Conversely, if Δp_x is increased, so that $\phi(p_x)$ is 'delocalised' in momentum space, then $\psi(x)$ will become more strongly localised about $x = 0$. Similar conclusions can be

derived for functions $\phi(p_x)$ having different shapes (Problem 2.1). In fact, it is a general result of Fourier analysis that if $\psi(x)$ and $\phi(p_x)$ are mutual Fourier transforms as in (2.13)–(2.14), then

$$\Delta x \Delta p_x \geq \hbar \quad (2.18)$$

This is the *Heisenberg uncertainty relation for position and momentum*, according to which the uncertainty Δx in measuring the x coordinate of a particle is related to the uncertainty Δp_x in measuring the x component of the momentum, the product of the uncertainties being larger than a quantity of order $\hbar$. A precise definition of the uncertainties Δx and Δp_x will be given below.

The foregoing discussion is readily generalised to more than one dimension. By superposing plane waves of the form (2.7) we obtain the wave packet

$$\begin{aligned} \Psi(\mathbf{r}, t) &= (2\pi\hbar)^{-3/2} \int \exp[i(\mathbf{p} \cdot \mathbf{r} - Et)/\hbar] \phi(\mathbf{p}) d\mathbf{p} \\ &= (2\pi\hbar)^{-3/2} \int \exp(i\mathbf{p} \cdot \mathbf{r}/\hbar) \Phi(\mathbf{p}, t) d\mathbf{p} \end{aligned} \quad (2.19)$$

where the wave functions in momentum space $\Phi(\mathbf{p}, t)$ and $\phi(\mathbf{p}) \equiv \Phi(\mathbf{p}, t=0)$ are the Fourier transforms of $\Psi(\mathbf{r}, t)$ and $\psi(\mathbf{r}) \equiv \Psi(\mathbf{r}, t=0)$, respectively. That is,

$$\Phi(\mathbf{p}, t) = (2\pi\hbar)^{-3/2} \int \exp(-i\mathbf{p} \cdot \mathbf{r}/\hbar) \Psi(\mathbf{r}, t) d\mathbf{r} \quad (2.20)$$

and

$$\phi(\mathbf{p}) = (2\pi\hbar)^{-3/2} \int \exp(-i\mathbf{p} \cdot \mathbf{r}/\hbar) \psi(\mathbf{r}) d\mathbf{r} \quad (2.21)$$

The three-dimensional form of the Heisenberg uncertainty relations for position and momentum is now

$$\Delta x \Delta p_x \geq \hbar, \quad \Delta y \Delta p_y \geq \hbar, \quad \Delta z \Delta p_z \geq \hbar \quad (2.22)$$

The position-momentum uncertainty relations (2.22) have been derived by using the theory of Fourier analysis. This theory may also be invoked to obtain a *time-energy uncertainty relation*. Indeed, according to Fourier analysis, a wave packet of duration Δt must be composed of plane-wave components whose angular frequencies extend over a range $\Delta\omega$ such that $\Delta t \Delta\omega \geq 1$. Since $E = \hbar\omega$, we therefore have

$$\Delta t \Delta E \geq \hbar \quad (2.23)$$

which is the *Heisenberg uncertainty relation for time and energy*. It connects the uncertainty ΔE in the determination of the energy of a system with the time interval Δt available for this energy determination. Thus, if a system does not stay longer than a time Δt in a given state of motion, its energy in that state will be uncertain by an amount $\Delta E \geq \hbar/\Delta t$.

The Heisenberg uncertainty relations (2.22) and (2.23), which are fundamental to any wave theory of matter, are particular cases of the *uncertainty principle* formulated by W. Heisenberg in 1927. This principle states that it is impossible to specify precisely and simultaneously the values of 'complementary' dynamical variables such as x and p_x , y and p_y , z and p_z , or t and E .

The uncertainty relations (2.22) and (2.23) may be used to estimate various basic quantities which occur in quantum physics. For example, it is possible by using the position-momentum uncertainty relations (2.22) to obtain an estimate of the size and of the energy of the hydrogen atom in its ground state (Problem 2.2). On the other hand, the time-energy uncertainty relation (2.23) is very useful in studying the lifetime of unstable systems.

Interpretation of the wave function

To make progress in the theory, we need to interpret the meaning of the wave function $\Psi(\mathbf{r}, t)$. In searching for this interpretation, we must remember that $\Psi(\mathbf{r}, t)$ is in general a complex function, and that $|\Psi(\mathbf{r}, t)|$ is large where the particle is likely to be found, and small everywhere else. We also recall that diffraction patterns, made by light, depend on the intensity of the radiation determined by the Poynting vector, which can be interpreted as measuring the flux of photons, and which depends on the square of the vector potential. In a similar way, M. Born in 1926 made the fundamental postulate that if a particle is described by a wave function $\Psi(\mathbf{r}, t)$, the probability of finding the particle within the volume element $d\mathbf{r} = dx dy dz$ about the point $\mathbf{r}$ at the time t is

$$P(\mathbf{r}, t) d\mathbf{r} = |\Psi(\mathbf{r}, t)|^2 d\mathbf{r} \quad (2.24)$$

so that

$$P(\mathbf{r}, t) = |\Psi(\mathbf{r}, t)|^2 = \Psi^*(\mathbf{r}, t)\Psi(\mathbf{r}, t) \quad (2.25)$$

is the (position) *probability density*. Born's statistical interpretation of the wave function has to be justified by the success of the theory built on it. For $P(\mathbf{r}, t)$ to be unique everywhere, one must require that the wave function $\Psi(\mathbf{r}, t)$ should be continuous and single-valued. It is also worth noting that since $|\Psi(\mathbf{r}, t)|^2$ is the physically significant quantity, two wave functions which differ from each other by a constant multiplicative factor of modulus one (that is, a constant phase factor of the form $\exp(i\alpha)$, where α is a real constant) are equivalent.

Since the probability of finding the particle somewhere must be unity, we deduce from (2.24) that the wave function $\Psi(\mathbf{r}, t)$ should be *normalised* so that

$$\int |\Psi(\mathbf{r}, t)|^2 d\mathbf{r} = 1 \quad (2.26)$$

where the integral extends over all space. Wave functions satisfying this condition are said to be *square integrable*. However, not every wave function can be normalised like this, for example the plane wave (2.7). In this case $|\Psi(\mathbf{r}, t)|^2 d\mathbf{r}$ can be interpreted as the *relative* probability of finding the particle at time t in a volume

element $d\mathbf{r}$ centred about $\mathbf{r}$, so that the ratio $|\Psi(\mathbf{r}_1, t)|^2/|\Psi(\mathbf{r}_2, t)|^2$ gives the probability of finding the particle within a volume element centred about $\mathbf{r} = \mathbf{r}_1$, compared with that of finding it within the same volume element at $\mathbf{r} = \mathbf{r}_2$.

There are alternative methods of normalising wave functions which are not square integrable. For example, we may assume that the system is confined to a large box of volume V , in which case the normalisation condition (2.26) reads

$$\int_V |\Psi(\mathbf{r}, t)|^2 d\mathbf{r} = 1 \quad (2.27)$$

where the integration now extends over the volume V . It can be shown that calculated physical quantities are independent of V , provided the box is sufficiently large. Plane waves such as (2.7) can also be 'normalised' with the help of the *Dirac delta function* $\delta(x)$, which is defined by the relation

$$f(0) = \int_{-\infty}^{+\infty} f(x) \delta(x) dx \quad (2.28)$$

where $f(x)$ is an arbitrary well-behaved function. There are several ways of representing the delta function, one of which is

$$\delta(x) = \lim_{\alpha \rightarrow \infty} \frac{\sin \alpha x}{\pi x} \quad (2.29)$$

When used in (2.28) the infinitely rapid oscillations of this function when $x \neq 0$ imply that the only contribution to the integral arises from the neighbourhood of $x = 0$, and

$$\int_{-\infty}^{+\infty} f(x) \delta(x) dx = f(0) \lim_{\alpha \rightarrow \infty} \int_{-\infty}^{+\infty} \frac{\sin \alpha x}{\pi x} dx = f(0) \quad (2.30)$$

Other useful representations of the Dirac delta function are

$$\delta(x) = (2\pi)^{-1} \lim_{K \rightarrow \infty} \int_{-K}^{+K} \exp(ikx) dk \quad (2.31a)$$

$$= (2\pi)^{-1} \int_{-\infty}^{+\infty} \exp(ikx) dk$$

$$\delta(x) = \lim_{\varepsilon \rightarrow 0^+} (2\pi)^{-1} \int_{-\infty}^{+\infty} \exp(ikx) \exp(-\varepsilon|k|) dk$$

$$= \lim_{\varepsilon \rightarrow 0^+} \frac{\varepsilon}{\pi(x^2 + \varepsilon^2)} \quad (2.31b)$$

$$\delta(x) = \lim_{\beta \rightarrow \infty} \frac{1 - \cos \beta x}{\pi \beta x^2} \quad (2.31c)$$

$$\delta(x) = \lim_{\varepsilon \rightarrow 0} \frac{\theta(x + \varepsilon) - \theta(x)}{\varepsilon} \quad (2.31d)$$

where in the last equation $\theta(x)$ denotes the step function such that $\theta(x) = 1$ for $x > 0$, and $\theta(x) = 0$ for $x < 0$.

We see that the Dirac delta function is an unusual function (it is in fact a 'distribution') which is such that, in effect,

$$\delta(x) = \begin{cases} 0 & \text{if } x \neq 0 \\ +\infty & \text{if } x = 0 \end{cases} \quad (2.32a)$$

and

$$\int_{-\infty}^{+\infty} \delta(x) dx = 1. \quad (2.32b)$$

The main properties of the Dirac delta functions are the following:

$$\int_a^b f(x) \delta(x - x_0) dx = f(x_0) \quad \begin{matrix} \text{if } a < x_0 < b \\ = 0 & \text{if } x_0 < a \text{ or } x_0 > b \end{matrix} \quad (2.33a)$$

$$(2.33b)$$

$$\delta(x) = \delta(-x) \quad (2.33c)$$

$$(2.33d)$$

$$\delta(ax) = \frac{1}{|a|} \delta(x), \quad a \neq 0$$

$$(2.33e)$$

$$f(x) \delta(x - a) = f(a) \delta(x - a)$$

$$(2.33f)$$

$$\int \delta(a - x) \delta(x - b) dx = \delta(a - b)$$

$$(2.33g)$$

$$\delta[g(x)] = \sum_i \frac{1}{|g'(x_i)|} \delta(x - x_i)$$

where in the last relation x_i is a zero of $g(x)$ and $g'(x_i) \neq 0$. Particular cases of (2.33g) are

$$\delta[(x - a)(x - b)] = \frac{1}{|a - b|} [\delta(x - a) + \delta(x - b)], \quad a \neq b \quad (2.33h)$$

and

$$\delta(x^2 - a^2) = \frac{1}{2|a|} [\delta(x - a) + \delta(x + a)] \quad (2.33i)$$

Using the representation (2.31a) of the Dirac delta function, we find that the plane waves

$$\psi_k(\mathbf{r}) = (2\pi)^{-3} \exp(i\mathbf{k} \cdot \mathbf{r}) \quad (2.34)$$

satisfy the relation

$$\begin{aligned} \int \psi_k^*(\mathbf{r}) \psi_{k'}(\mathbf{r}) d\mathbf{r} &= (2\pi)^{-3} \int_{-\infty}^{+\infty} \exp[i(k_x - k'_x)x] dx \\ &\times \int_{-\infty}^{+\infty} \exp[i(k_y - k'_y)y] dy \int_{-\infty}^{+\infty} \exp[i(k_z - k'_z)z] dz \\ &= \delta(k_x - k'_x) \delta(k_y - k'_y) \delta(k_z - k'_z) \\ &= \delta(\mathbf{k} - \mathbf{k}') \end{aligned} \quad (2.35)$$

which is known as an *orthonormality relation*. In addition, the plane waves (2.34) satisfy the *closure relation*

$$\int \psi_k^*(\mathbf{r}') \psi_k(\mathbf{r}) d\mathbf{k} = \delta(\mathbf{r} - \mathbf{r}') \quad (2.36)$$

The normalisation condition expressed by (2.35) is often referred to as *k-normalisation*. Other normalisations for wave functions which are not square integrable may be used, for example the *p-normalisation* and the *energy normalisation*, which are the subject of Problem 2.3.

Expectation values

Consider a one-dimensional system in which the normalised wave function is $\Psi(x, t)$. Since $|\Psi(x, t)|^2 dx$ is the probability of finding the particle at the position x in the interval dx , the *average* or *expectation value* of x may be defined as

$$\langle x \rangle = \int_{-\infty}^{+\infty} x |\Psi(x, t)|^2 dx \quad (2.37)$$

Likewise, the expectation value of any function $f(\mathbf{r})$ of the coordinates x, y, z , in a state represented by the normalised wave function $\Psi(\mathbf{r}, t)$ is

$$\langle f \rangle = \int f(\mathbf{r}) |\Psi(\mathbf{r}, t)|^2 d\mathbf{r} \quad (2.38)$$

where the integral is taken over all space.

How should we calculate the average value of the momentum of the particle? The wavelength, which determines the magnitude of the momentum of a particle through (2.1), has only a precise value for a plane wave such as (2.7). For such a wave the momentum is obtained by operating on the wave function with the differential operator $-i\hbar \nabla$, as in (2.10). When the wave function is not a plane wave, we shall still assume that the momentum is represented by this operator (see (2.11)) and the average value of the momentum in a state represented by a wave function $\Psi(\mathbf{r}, t)$ normalised to unity is

$$\langle \mathbf{p} \rangle = \int \Psi^*(\mathbf{r}, t) \mathbf{p}_{\text{op}} \Psi(\mathbf{r}, t) d\mathbf{r} = \int \Psi^*(\mathbf{r}, t) (-i\hbar \nabla) \Psi(\mathbf{r}, t) d\mathbf{r} \quad (2.39)$$

In general, if an operator A depends on $\mathbf{r}$, $\mathbf{p}$ and t , its expectation value is defined as

$$\langle A \rangle = \int \Psi^*(\mathbf{r}, t) A(\mathbf{r}, -i\hbar\nabla, t) \Psi(\mathbf{r}, t) d\mathbf{r} \quad (2.40)$$

The uncertainties Δx , Δp_x and so on which appeared in the uncertainty relations (2.22) can now be defined as standard deviations by the formulae

$$\Delta x = \langle (x - \langle x \rangle)^2 \rangle^{1/2}, \quad \Delta p_x = \langle (p_x - \langle p_x \rangle)^2 \rangle^{1/2} \quad (2.41)$$

2.2 The Schrödinger equation

In order to compute the average values of functions of the position and momentum of a particle, which are the quantities determined by a series of measurements, some way is needed of calculating wave functions. We first notice that the wave packet (2.19) for a free particle satisfies the differential equation

$$i\hbar \frac{\partial}{\partial t} \Psi = -\frac{\hbar^2}{2m} \nabla^2 \Psi \quad (2.42)$$

because $E(p) = p^2/(2m)$, and this is called the Schrödinger equation for a free particle. What is the Schrödinger equation for a particle moving in a potential $V(\mathbf{r}, t)$? Using (2.11), the Schrödinger equation for a free particle can be written

$$E_{\text{op}} \Psi = \frac{1}{2m} (\mathbf{p}_{\text{op}})^2 \Psi \quad (2.43)$$

The operator $(\mathbf{p}_{\text{op}})^2/(2m)$ represents the kinetic energy T of the particle. The total energy or Hamiltonian operator of a particle is just

$$H = T + V = -\frac{\hbar^2}{2m} \nabla^2 + V \quad (2.44)$$

and the generalisation of the Schrödinger equation for a particle in a potential is

$$E_{\text{op}} \Psi = (T + V) \Psi \quad (2.45)$$

or

$$i\hbar \frac{\partial}{\partial t} \Psi = H \Psi \quad (2.46)$$

which explicitly reads

$$i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}, t) = \left[-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}, t) \right] \Psi(\mathbf{r}, t) \quad (2.47)$$

This is the *time-dependent Schrödinger equation* for a particle moving in a potential. We remark that this equation is *linear* in $\Psi(\mathbf{r}, t)$. As a result, it satisfies

the *superposition principle*, according to which a linear superposition of possible wave functions is also a possible wave function. The fact that the superposition principle applies is directly related to the wave nature of matter, and in particular to the existence of interference effects for de Broglie waves. We also note that the Schrödinger equation (2.47) is of first order in the time derivative $\partial/\partial t$, so that once the initial value of the wave function Ψ is given at some time t_0 , namely $\Psi(\mathbf{r}, t_0)$, its values at all other times can be found.

The time-dependent Schrödinger equation (2.47) will now be used to obtain two important results concerning probability conservation and the time variation of expectation values.

Probability conservation

The probability of finding a particle within a fixed volume V is seen from (2.24) to be

$$\int_V P(\mathbf{r}, t) d\mathbf{r} = \int_V |\Psi(\mathbf{r}, t)|^2 d\mathbf{r} \quad (2.48)$$

The rate of change of this probability can be obtained with the help of (2.47):

$$\begin{aligned} \frac{\partial}{\partial t} \int_V P(\mathbf{r}, t) d\mathbf{r} &= \int_V \left(\Psi^* \frac{\partial \Psi}{\partial t} + \frac{\partial \Psi^*}{\partial t} \Psi \right) d\mathbf{r} \\ &= \frac{i\hbar}{2m} \int_V [\Psi^* \nabla^2 \Psi - (\nabla^2 \Psi^*) \Psi] d\mathbf{r} \\ &= - \int_V \nabla \cdot \mathbf{j} d\mathbf{r} \end{aligned} \quad (2.49)$$

where we have introduced the vector

$$\mathbf{j}(\mathbf{r}, t) = \frac{\hbar}{2mi} [\Psi^* \nabla \Psi - (\nabla \Psi^*) \Psi] \quad (2.50)$$

Using Green's theorem, we find that

$$\frac{\partial}{\partial t} \int_V P(\mathbf{r}, t) d\mathbf{r} = - \int_S \mathbf{j} \cdot d\mathbf{S} \quad (2.51)$$

where the integral on the right-hand side is over the surface S bounding the volume V . Since the rate of change of the probability of finding the particle in the region V is equal to the probability flux passing through the surface S , the vector $\mathbf{j}$ can be interpreted as a *probability current density*. The equation

$$\nabla \cdot \mathbf{j} + \frac{\partial P}{\partial t} = 0 \quad (2.52)$$

which follows from (2.49) is analogous to the continuity equation of electromagnetism, expressing charge conservation.

Time variation of expectation values

The rate of change of the expectation value (2.40) of an operator A is given by

$$\begin{aligned}\frac{d}{dt}\langle A \rangle &= \frac{d}{dt} \int (\Psi^* A \Psi) d\mathbf{r} \\ &= \int \left(\frac{\partial \Psi^*}{\partial t} A \Psi + \Psi^* \frac{\partial A}{\partial t} \Psi + \Psi^* A \frac{\partial \Psi}{\partial t} \right) d\mathbf{r} \\ &= \left\langle \frac{\partial A}{\partial t} \right\rangle + \frac{1}{i\hbar} \int \Psi^* (AH - HA) \Psi d\mathbf{r}\end{aligned}\quad (2.53)$$

where we have used the Schrödinger equation (2.46) and the fact that H is Hermitian (see below). The *commutator* of two operators A and B is defined as

$$[A, B] = AB - BA \quad (2.54)$$

and two operators are said to commute if their commutator vanishes. Using this definition (2.53) can be written as

$$\frac{d}{dt}\langle A \rangle = \left\langle \frac{\partial A}{\partial t} \right\rangle + \frac{1}{i\hbar} \langle [A, H] \rangle \quad (2.55)$$

In particular, if the operator A does not depend explicitly on time (that is, $\partial A/\partial t = 0$), we have

$$i\hbar \frac{d}{dt} \langle A \rangle = \langle [A, H] \rangle \quad (2.56)$$

and if A commutes with H , A is a *constant of the motion* ($d\langle A \rangle/dt = 0$).

Using the result (2.56) one can prove (Problem 2.6) *Ehrenfest's theorem*, according to which the expectation values of the variables $\mathbf{r}$ and $\mathbf{p}$ obey equations corresponding to Newton's equations of motion. That is,

$$m \frac{d\langle \mathbf{r} \rangle}{dt} = \langle \mathbf{p} \rangle, \quad \frac{d\langle \mathbf{p} \rangle}{dt} = -\langle \nabla V \rangle \quad (2.57)$$

Time-independent Schrödinger equation and energy eigenfunctions

When the potential does not depend on the time, the Schrödinger equation (2.47) admits *stationary state solutions* of the form

$$\Psi(\mathbf{r}, t) = \psi_E(\mathbf{r}) \exp(-iEt/\hbar) \quad (2.58)$$

where E is a constant and where $\psi_E(\mathbf{r})$ satisfies the *time-independent Schrödinger equation*

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \right] \psi_E(\mathbf{r}) = E \psi_E(\mathbf{r}) \quad (2.59)$$

or

$$H\psi_E = E\psi_E \quad (2.60)$$

The significance of the constant E is seen by recognising that E is the expectation value of the total energy $H = T + V$. Indeed, using (2.58) and (2.60), we have

$$E = \langle H \rangle = \int \Psi^*(\mathbf{r}, t) H \Psi(\mathbf{r}, t) d\mathbf{r} \quad (2.61)$$

In writing (2.61) we have assumed that the wave function Ψ given by (2.58) is normalised to unity, so that

$$\int \psi_E^*(\mathbf{r}) \psi_E(\mathbf{r}) d\mathbf{r} = 1 \quad (2.62)$$

where the integral can be taken over all space for square integrable functions, or over a large volume V in the other cases. We shall find that only certain values of E are compatible with normalisable solutions. These values are called the *energy eigenvalues* and the corresponding solutions ψ_E are the *eigenfunctions* of the total energy or Hamiltonian operator H . Equations of the type

$$A\psi_n = a_n\psi_n \quad (2.63)$$

— where A is an operator and a_n is a number — are called *eigenvalue equations*, a_n being called the *eigenvalue* and ψ_n the *eigenfunction* of the operator A . If more than one eigenfunction corresponds to a given eigenvalue, this eigenvalue is said to be *degenerate*.

It is a general postulate of quantum mechanics that each dynamical variable (such as the position, momentum, energy, ...) can be represented by a linear operator [1] and that the result of a precise measurement of the variable can only be one of the eigenvalues of this operator. For example, the only possible result of a measurement of the energy of the system is one of the eigenvalues E of the Hamiltonian operator H .

The spectrum of the Hamiltonian

The set of all eigenvalues of the Hamiltonian is called the *energy spectrum*. It may consist of *discrete* values or a *continuous* range, or both. In general the discrete eigenvalues are associated with *bound states* (analogous to the closed orbits of classical mechanics) and the continuum with *scattering states* (corresponding to open classical orbits).

[1] For a definition of linear operators, see Byron and Fuller (1969) or Mathews and Walker (1973).

It is easy to show by integrating by parts that for any normalised eigenfunction ψ_E of H one has

$$\begin{aligned} E^* &= \langle H \rangle^* = \left[\int \psi_E^* \left[-\frac{\hbar^2}{2m} \nabla^2 + V \right] \psi_E \, d\mathbf{r} \right]^* \\ &= \int \psi_E \left[-\frac{\hbar^2}{2m} \nabla^2 + V \right] \psi_E^* \, d\mathbf{r} = \int \psi_E^* \left[-\frac{\hbar^2}{2m} \nabla^2 + V \right] \psi_E \, d\mathbf{r} \\ &= \langle H \rangle = E \end{aligned} \quad (2.64)$$

so that the eigenvalues E are real.

In general, if A is a linear operator and Φ and Ψ are any two wave functions, the adjoint operator $A^\dagger$ is defined by the relation

$$\int \Phi^* A^\dagger \Psi \, d\mathbf{r} = \int (A\Phi)^* \Psi \, d\mathbf{r} \quad (2.65)$$

If $A^\dagger = A$, then A is said to be *self-adjoint* or *Hermitian*, in which case $\langle A \rangle = \langle A \rangle^*$ and the eigenvalues of A are real. Operators representing physical quantities (such as H , $\mathbf{r}$ or $\mathbf{p}$) are called *observables*, and it is a postulate of quantum mechanics that observables must be Hermitian.

We shall now prove that if ψ_E and $\psi_{E'}$ are two energy eigenfunctions corresponding to unequal eigenvalues E and E' , then

$$\int \psi_E^*(\mathbf{r}) \psi_{E'}(\mathbf{r}) \, d\mathbf{r} = 0 \quad (2.66)$$

in which case ψ_E and $\psi_{E'}$ are said to be *orthogonal*. We first note that

$$\psi_E^*(H - E) \psi_E = 0 \quad (2.67a)$$

and

$$\psi_E[(H - E') \psi_{E'}]^* = 0 \quad (2.67b)$$

Subtracting these two equations and integrating, we have

$$(E - E') \int \psi_E^* \psi_{E'} \, d\mathbf{r} = \int [\psi_E^* H \psi_{E'} - (H \psi_{E'})^* \psi_E] \, d\mathbf{r} = 0 \quad (2.68)$$

where in the last step we have used the fact that H is Hermitian. Since $E \neq E'$, we see that (2.66) follows from (2.68).

When several eigenfunctions ψ_{E_r} , $r = 1, 2, \dots, \alpha$, correspond to a given eigenvalue E (that is, if the eigenvalue E is degenerate) linear combinations of the eigenfunctions ψ_{E_r} can always be constructed so that the resulting α eigenfunctions belonging to the eigenvalue E are mutually orthogonal [see for example Bransden and Joachain (2000)]. Each of these eigenfunctions is evidently orthogonal to every eigenfunction belonging to a different eigenvalue $E' \neq E$. Thus all the energy eigenfunctions can be made orthogonal to each other.

The eigenfunctions ψ_n of an operator A are said to constitute an *orthonormal set* if they are both normalised and mutually orthogonal. That is,

$$\int \psi_n^*(\mathbf{r}) \psi_{n'}(\mathbf{r}) \, d\mathbf{r} = \delta_{nn'} \quad (2.69)$$

where $\delta_{nn'}$ is the Kronecker symbol

$$\delta_{nn'} = \begin{cases} 1, & n = n' \\ 0, & n \neq n' \end{cases} \quad (2.70)$$

The energy eigenfunctions ψ_E constitute such a set, and in the same way so do the eigenfunctions of an Hermitian operator representing any observable [2].

2.3

Expansions, operators and observables

We shall assume that all the orthonormal eigenfunctions $\psi_n(\mathbf{r})$ of a given Hermitian operator A form a *complete set*, in the mathematical sense that an arbitrary normalised wave function $\Psi(\mathbf{r}, t)$ can be expanded in terms of them:

$$\Psi(\mathbf{r}, t) = \sum_n c_n(t) \psi_n(\mathbf{r}) \quad (2.71)$$

Suppose now that at time t a measurement is made of the observable represented by A . Then, as we have seen above, the value obtained must be one of the eigenvalues a_n . A further postulate, due to Born, is that the probability $P_n(t)$ of obtaining the value a_n is given by

$$P_n(t) = |c_n(t)|^2 \quad (2.72)$$

As a consequence, we are certain to obtain a particular eigenvalue a_n only when the wave function that describes the particle at time t is the corresponding eigenfunction ψ_n (apart from a constant phase factor). Immediately after a measurement in which the result a_n is obtained, the system will be found in the state represented by the wave function ψ_n .

We can calculate the *probability amplitudes* c_n from (2.71) using the orthonormal property (2.69) of the expansion functions ψ_n . Multiplying (2.71) by ψ_n^* and integrating, we obtain

$$\begin{aligned} \int \psi_n^*(\mathbf{r}) \Psi(\mathbf{r}, t) \, d\mathbf{r} &= \sum_n c_n(t) \int \psi_n^*(\mathbf{r}) \psi_n(\mathbf{r}) \, d\mathbf{r} \\ &= \sum_n c_n(t) \delta_{nn'} = c_n(t) \end{aligned} \quad (2.73)$$

[2] If n is a continuous index, then a 'delta function' normalisation analogous to (2.35) can be used. Alternatively, a 'box' normalisation of the type (2.27) can be introduced, in which case n remains discrete and the relations (2.69)–(2.70) apply without modification.

It is convenient at this point to introduce the Dirac notation

$$\langle f|g\rangle = \int f^*(\mathbf{r})g(\mathbf{r}) d\mathbf{r} \quad (2.74)$$

so that the result (2.73) reads $c_{n'} = \langle \psi_{n'} | \Psi \rangle$. More generally, we define the *matrix element* of an operator A as

$$\langle f|A|g\rangle = \int f^*(\mathbf{r})Ag(\mathbf{r}) d\mathbf{r} \quad (2.75)$$

so that the expectation, or average, value of A can be written as

$$\begin{aligned} \langle A \rangle &= \langle \Psi | A | \Psi \rangle \\ &= \sum_n \sum_{n'} c_n^* c_{n'} \langle \psi_n | A | \psi_{n'} \rangle \\ &= \sum_n \sum_{n'} c_n^* c_{n'} a_{nn'} \\ &= \sum_n |c_n|^2 a_{nn} \end{aligned} \quad (2.76)$$

where we have used (2.63) and the orthonormality relation (2.69), which reads $\langle \psi_{n'} | \psi_n \rangle = \delta_{nn'}$ in Dirac's notation.

Another important relation can be proved from (2.71) and (2.73). Indeed, since

$$\begin{aligned} \Psi(\mathbf{r}, t) &= \sum_n c_n(t) \psi_n(\mathbf{r}) \\ &= \sum_n \left[\int \psi_n^*(\mathbf{r}') \Psi(\mathbf{r}', t) d\mathbf{r}' \right] \psi_n(\mathbf{r}) \end{aligned} \quad (2.77)$$

we have

$$\sum_n \psi_n^*(\mathbf{r}') \psi_n(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}') \quad (2.78)$$

where $\delta(\mathbf{r} - \mathbf{r}')$ is the Dirac delta function discussed in Section 2.1. The result (2.78) (see also (2.36) for the special case of the plane waves) is known as the *closure* property of the orthonormal set of functions ψ_n .

General solution of the Schrödinger equation for a time-independent potential

Let us assume that the potential is independent of time, so that the time-dependent Schrödinger equation (2.47) admits stationary state solutions of the form (2.58). Expanding the general solution of (2.47) in terms of energy eigenfunctions, we then write (see (2.71) and (2.73))

$$\Psi(\mathbf{r}, t) = \sum_E c_E(t) \psi_E(\mathbf{r}) \quad (2.79)$$

where

$$c_E(t) = \int \psi_E^*(\mathbf{r}) \Psi(\mathbf{r}, t) d\mathbf{r} \quad (2.80)$$

Substituting (2.79) into the Schrödinger equation (2.47), and using the orthonormality of the eigenfunctions ψ_E , together with the fact that they satisfy the time-independent Schrödinger equation (2.59), we find that the coefficients $c_E(t)$ satisfy the equation

$$i\hbar \frac{d}{dt} c_E(t) = E c_E(t) \quad (2.81)$$

which is readily integrated to yield

$$c_E(t) = c_E(t_0) \exp \left[-\frac{i}{\hbar} E(t - t_0) \right] \quad (2.82)$$

We note that the probability P_E of obtaining the energy value E is constant, since from (2.72) and (2.82) we have

$$P_E = |c_E(t)|^2 = |c_E(t_0)|^2 \quad (2.83)$$

Using (2.79) and (2.82), we find that

$$\Psi(\mathbf{r}, t) = \sum_E c_E(t_0) \exp \left[-\frac{i}{\hbar} E(t - t_0) \right] \psi_E(\mathbf{r}) \quad (2.84)$$

or

$$\Psi(\mathbf{r}, t) = \sum_E \tilde{c}_E(t_0) \psi_E(\mathbf{r}) \exp \left(-\frac{i}{\hbar} Et \right); \quad \tilde{c}_E(t_0) = c_E(t_0) \exp \left(\frac{i}{\hbar} Et_0 \right) \quad (2.85)$$

so that the general solution $\Psi(\mathbf{r}, t)$ is a linear superposition of stationary state solutions (2.58), as we expect from the linearity of the Schrödinger equation (2.47).

These results may also be used to write down an expression for $\Psi(\mathbf{r}, t)$ at any time, once it is known at the time $t = t_0$. Indeed, from (2.80) and (2.84), we have

$$\begin{aligned} \Psi(\mathbf{r}, t) &= \sum_E \left[\int \psi_E^*(\mathbf{r}') \Psi(\mathbf{r}', t_0) d\mathbf{r}' \right] \exp \left[-\frac{i}{\hbar} E(t - t_0) \right] \psi_E(\mathbf{r}) \\ &= \int K(\mathbf{r}, t; \mathbf{r}', t_0) \Psi(\mathbf{r}', t_0) d\mathbf{r}' \end{aligned} \quad (2.86)$$

where

$$K(\mathbf{r}, t; \mathbf{r}', t_0) = \sum_E \psi_E^*(\mathbf{r}') \psi_E(\mathbf{r}) \exp \left[-\frac{i}{\hbar} E(t - t_0) \right] \quad (2.87)$$

We see from (2.86) that K may be interpreted as the probability amplitude that a particle originally at $\mathbf{r}'$ will propagate to the point $\mathbf{r}$ in the time interval $t - t_0$. The function K is therefore called a *propagator*.

Matrix representations of wave functions and operators

Let us consider a complete orthonormal set of functions $\{\psi_j\}$. Any physically admissible wave function Ψ can be expanded in terms of them as (see (2.71) and (2.73))

$$\Psi = \sum_j c_j \psi_j, \quad c_j = \langle \psi_j | \Psi \rangle \quad (2.88)$$

The set $\{\psi_j\}$ is called a *basis* and the coefficients c_j are the components of Ψ in that basis. Once a given basis – also called a *representation* – has been chosen, Ψ is completely specified by its components c_j . Thus the set $\{\psi_j\}$ may be thought of as a basis in a vector space [3], and Ψ as being represented by a vector in that vector space, whose components are the coefficients c_j .

The action of a linear operator A on Ψ may be specified in terms of its effect on the components c_j in a given representation. Indeed, starting from (2.88) we have

$$A\Psi = A \sum_j c_j \psi_j = \sum_j c_j A\psi_j \quad (2.89)$$

Let us express the new vectors $\psi'_j = A\psi_j$ in our basis as

$$\psi'_j = A\psi_j = \sum_i A_{ij} \psi_i \quad (2.90)$$

where

$$A_{ij} = \langle \psi_i | A | \psi_j \rangle \quad (2.91)$$

Then, if we write

$$\Psi' = A\Psi = \sum_i c'_i \psi_i \quad (2.92)$$

so that the coefficients c'_i are the components of Ψ' in our basis, we have from (2.89)–(2.92)

$$\Psi' = A\Psi = \sum_i \sum_j A_{ij} c_j \psi_i = \sum_i c'_i \psi_i \quad (2.93)$$

so that

$$c'_i = \sum_j A_{ij} c_j \quad (2.94)$$

This is a *matrix equation* which relates the components c_j of Ψ and c'_i of $\Psi' = A\Psi$ by means of the *matrix elements* A_{ij} of the linear operator A .

Thus far we have adopted a given representation (that is, a given complete orthonormal set of basis vectors $\{\psi_j\}$) in which Ψ is represented by a vector having

[3] Vector spaces are discussed by Byron and Fuller (1969) and Mathews and Walker (1973).

components c_j and the operator A by a matrix whose matrix elements are given by (2.91). This representation, however, is not unique. Indeed, we may develop Ψ in terms of another complete orthonormal set of basis vectors $\{\phi_k\}$ as

$$\Psi = \sum_k d_k \phi_k, \quad d_k = \langle \phi_k | \Psi \rangle \quad (2.95)$$

and the coefficients d_k are the components of a vector representing Ψ in the 'new' basis $\{\phi_k\}$. In particular, any 'old' basis vector may be expressed in the 'new' basis as

$$\psi_j = \sum_k U_{kj} \phi_k \quad (2.96)$$

where

$$U_{kj} = \langle \phi_k | \psi_j \rangle \quad (2.97)$$

In the same way, the reverse expansion is

$$\phi_k = \sum_j \langle \psi_j | \phi_k \rangle \psi_j = \sum_j U_{kj}^* \psi_j \quad (2.98)$$

Since

$$\Psi = \sum_j c_j \psi_j = \sum_j \sum_k c_j U_{kj} \phi_k = \sum_k d_k \phi_k \quad (2.99)$$

we also see that

$$d_k = \sum_j U_{kj} c_j \quad (2.100)$$

The set of numbers U_{kj} can be regarded as the elements of a matrix U and the equation (2.100) written in matrix form as

$$d = Uc \quad (2.101)$$

The matrix U is readily shown to be *unitary*, namely such that

$$UU^\dagger = U^\dagger U = I \quad (2.102)$$

where I is the unit matrix. Indeed, we have

$$\begin{aligned} (UU^\dagger)_{kn} &= \sum_j U_{kj} U_{jn}^* = \sum_j U_{kj} U_{nj}^* \\ &= \sum_j \langle \phi_k | \psi_j \rangle \langle \phi_n | \psi_j \rangle^* \\ &= \int d\mathbf{r} \phi_k^*(\mathbf{r}) \int d\mathbf{r}' \phi_n(\mathbf{r}') \sum_j \psi_j^*(\mathbf{r}') \psi_j(\mathbf{r}) \\ &= \int d\mathbf{r} \phi_k^*(\mathbf{r}) \int d\mathbf{r}' \phi_n(\mathbf{r}') \delta(\mathbf{r} - \mathbf{r}') \\ &= \int d\mathbf{r} \phi_k^*(\mathbf{r}) \phi_n(\mathbf{r}) = \delta_{kn} \end{aligned} \quad (2.103)$$

where we have used the closure relation for the functions ψ_i . Similarly, one has

$$(U^\dagger U)_{kn} = \delta_{kn} \quad (2.104)$$

so that the unitarity property (2.102) follows from (2.103) and (2.104). Therefore the passage from one representation of quantum mechanics to another is effected by *unitary transformations*.

Instead of changing the representation by means of a unitary transformation, we may also apply unitary transformations directly to the wave functions and operators. For example, let Ψ and Φ be two wave functions and A an operator such that

$$A\Psi = \Phi \quad (2.105)$$

Let us now apply a unitary transformation such that

$$\Psi' = U\Psi, \quad \Phi' = U\Phi \quad (2.106)$$

where U is a unitary operator, namely a linear operator satisfying the equations (2.102), where I is the unit operator. Writing

$$A'\Psi' = \Phi' \quad (2.107)$$

we have

$$A'U\Psi = U\Phi = UA\Psi \quad (2.108)$$

Hence, using (2.102), we find that

$$A = U^\dagger A' U, \quad A' = UA U^\dagger \quad (2.109)$$

These equations imply (Problem 2.9) that if A and A' are two operators connected by a unitarity transformation:

1. If A is Hermitian, then A' is also Hermitian.
2. The eigenvalues of A' are the same as those of A .
3. One has

$$\langle \Phi' | A' | \Psi' \rangle = \langle \Phi | A | \Psi \rangle \quad (2.110)$$

where Ψ' and Φ' are defined by (2.106). In particular, the expectation value $\langle \Psi' | A' | \Psi' \rangle$ remains unchanged. By choosing $A = I$ (the unit operator) we also see that the scalar (inner) product $\langle \Phi | \Psi \rangle$ is invariant under a unitary transformation. As a result, the normalisation is also preserved, since $\langle \Psi' | \Psi' \rangle = \langle \Psi | \Psi \rangle$.

Let us assume that the orthonormal functions $\{\psi_i\}$ which we use as a basis set for a matrix representation are the eigenfunctions of some quantum mechanical operator A . Then, using (2.63), we find that

$$A_{ij} = \langle \psi_i | A | \psi_j \rangle = \langle \psi_i | a_j | \psi_j \rangle = a_j \delta_{ij} \quad (2.111)$$

so that the matrix which represents the operator A in this basis is *diagonal*. Solving the eigenvalue equation (2.63) for an observable A is therefore equivalent to finding a unitary transformation which diagonalises the corresponding matrix.

The Heisenberg equations of motion

We have learned in Section 2.2 how to calculate the time rate of change of the expectation value of an operator A (see (2.55)). In a similar way, one can show that

$$\frac{d}{dt} \int \Phi^* A \Psi \, d\mathbf{r} = \int \Phi^* \frac{\partial A}{\partial t} \Psi \, d\mathbf{r} + \frac{1}{i\hbar} \int \Phi^* [A, H] \Psi \, d\mathbf{r} \quad (2.112)$$

where $\Phi(\mathbf{r}, t)$ and $\Psi(\mathbf{r}, t)$ are two arbitrary wave functions. Thus we have the operator (or matrix) equation

$$\frac{dA}{dt} = \frac{\partial A}{\partial t} + \frac{1}{i\hbar} [A, H] \quad (2.113)$$

where dA/dt is the operator whose matrix elements are the time rate of change of the matrix elements of A . The equation (2.113) is known as the *Heisenberg equation of motion* of a dynamical variable.

Commuting observables

Let two different observables be represented by the operators A and B . From Born's postulate, it follows that if there is a state in which a simultaneous measurement of the two observables is certain to yield the values a_i and b_j , then a_i and b_j must be eigenvalues of A and B , respectively. The corresponding wave function, ψ_{ij} , must be an eigenfunction of both A and B . That is,

$$A\psi_{ij} = a_i \psi_{ij}; \quad B\psi_{ij} = b_j \psi_{ij} \quad (2.114)$$

If such eigenfunctions of A and B can be found for all the eigenvalues a_i and b_j , then the two observables are said to be *compatible*, and the ψ_{ij} form a complete set. By taking suitable linear combinations of any degenerate eigenfunctions, this set can be made orthonormal,

$$\langle \psi_{ij} | \psi_{i'j'} \rangle = \delta_{ii'} \delta_{jj'} \quad (2.115)$$

From (2.114), we see that

$$[A, B]\psi_{ij} = (AB - BA)\psi_{ij} = 0 \quad (2.116)$$

and since any wave function Ψ can be expanded in terms of the orthonormal set $\{\psi_{ij}\}$, we must have

$$[A, B] = 0 \quad (2.117)$$

so that the two operators A and B commute with each other. Conversely, if A and B are two operators which commute with each other, there exists a complete set

of eigenfunctions which are simultaneously eigenstates of both A and B . On the other hand, if $[A, B] \neq 0$, a precise simultaneous measurement of both the observables is impossible.

As a first example, let us consider the Cartesian components x, y, z of the position operator and the Cartesian components of the momentum operator, $p_x = -i\hbar \partial/\partial x$, $p_y = -i\hbar \partial/\partial y$ and $p_z = -i\hbar \partial/\partial z$. If $f(x)$ is a function of x , we have

$$[x, p_x]f(x) = -i\hbar \left\{ x \frac{\partial f}{\partial x} - \frac{\partial}{\partial x} (xf) \right\} = i\hbar f(x) \quad (2.118)$$

so that we may write the relation $[x, p_x] = i\hbar$. More generally, we have

$$[x, p_x] = [y, p_y] = [z, p_z] = i\hbar \quad (2.119)$$

while all other commutators – such as $[x, p_y]$ – vanish. It follows that, for example, x and p_y have common eigenfunctions and can be measured simultaneously with arbitrary accuracy. In contrast, since x and p_x do not commute with each other, a precise simultaneous measurement of both of these observables is impossible. Indeed, as we saw in Section 2.1, the accuracy to which both x and p_x can be measured is limited by the uncertainty principle.

As a second example, let us consider a system described by a *time-independent* Hamiltonian H , so that stationary states corresponding to a definite value E of the total energy exist. Now, as we have seen above, an operator A corresponding to a constant of the motion commutes with the Hamiltonian. Hence, in this case a constant of the motion can be measured simultaneously with the energy E . It can also be shown (Problem 2.10) that if two non-commuting operators A and B are constants of the motion, then the energy levels of the system are, in general, degenerate.

The foregoing discussion can be extended to any number of observables. In particular if $A, B, C, \dots$ are a set of commuting observables, then a complete set of simultaneous eigenfunctions of these observables exists. The largest set of commuting observables which can be found for a given system is called a *complete set of commuting observables*.

Commutator algebra

It is convenient to list here some elementary rules for the calculation of commutators. These rules are easily verified from the basic definition (2.54). If A, B and C are three linear operators

$$[A, B] = -[B, A] \quad (2.120a)$$

$$[A, B + C] = [A, B] + [A, C] \quad (2.120b)$$

$$[A, BC] = [A, B]C + B[A, C] \quad (2.120c)$$

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0 \quad (2.120d)$$

2.4 One-dimensional examples

In this section we shall analyse the time-independent Schrödinger equation

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x) \quad (2.121)$$

for two simple one-dimensional potentials: the infinite square well and the linear harmonic oscillator. This will allow us to illustrate the theory and to obtain several results which will be useful in further chapters.

The infinite square well

Let us consider a particle of mass m which is constrained by impenetrable walls to move in a region of width L , where the potential energy is constant. Taking this constant to be zero, and setting $a = L/2$, the potential energy for this problem is

$$V(x) = \begin{cases} 0, & |x| < a \\ \infty, & x < -a, \quad x > a \end{cases} \quad (2.122)$$

and is illustrated in Fig. 2.2(a).

Because the potential energy is infinite at $x = \pm a$, the probability of finding the particle outside the well is zero. Hence the wave function $\psi(x)$ must vanish for $|x| > a$, and we only need to solve the Schrödinger eigenvalue equation (2.121) inside the well. Moreover, since the wave function must be continuous, $\psi(x)$ must vanish at the constraining walls, namely

$$\psi(x) = 0 \quad \text{at} \quad x = \pm a \quad (2.123)$$

We shall see shortly that it is precisely this *boundary condition* which leads to the quantisation of the energy.

The time-independent Schrödinger equation for $|x| < a$ reads

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} = E\psi(x) \quad (2.124)$$

and has the general solution

$$\psi(x) = A \sin kx + B \cos kx, \quad k = \left(\frac{2mE}{\hbar^2} \right)^{1/2} \quad (2.125)$$

Applying the boundary condition (2.123) we find that

$$A \sin ka = 0, \quad B \cos ka = 0 \quad (2.126)$$

As a result, there are two possible classes of solutions. For the first class $A = 0$ and $\cos ka = 0$, so that the only allowed values of k are

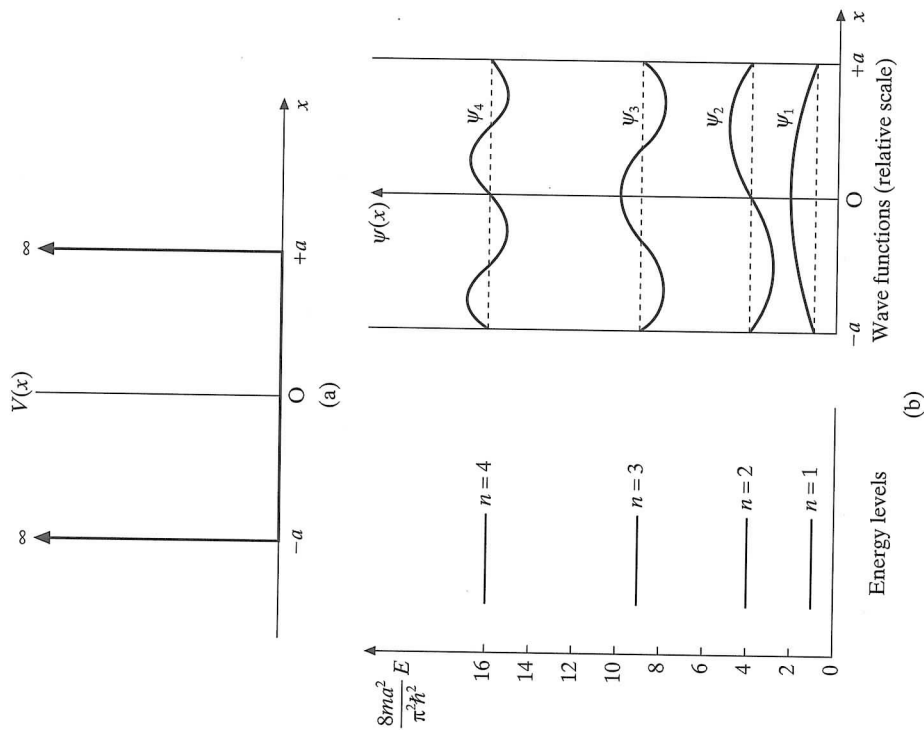


Figure 2.2 (a) The infinite square well
(b) The first four energy eigenvalues and eigenfunctions for the infinite square well.

$$k_n = \frac{n\pi}{2a} = \frac{n\pi}{L} \quad (2.127)$$

with $n = 1, 3, 5, \dots$. The corresponding eigenfunctions $\psi_n(x) = B_n \cos k_n x$ can be normalised so that

$$\int_{-a}^{+a} \psi_n^*(x) \psi_n(x) dx = 1 \quad (2.128)$$

from which the normalisation constants B_n are found (within an arbitrary multiplicative factor of modulus one) to be $B_n = a^{-1/2}$. The normalised eigenfunctions of the first class may therefore be written as

$$\psi_n(x) = \frac{1}{\sqrt{a}} \cos \frac{n\pi}{2a} x, \quad n = 1, 3, 5, \dots \quad (2.129)$$

Similarly, for the *second class* of solutions, such that $B = 0$ and $\sin ka = 0$, the allowed values of k are given by (2.127) with $n = 2, 4, 6, \dots$ and the corresponding normalised eigenfunctions are

$$\psi_n(x) = \frac{1}{\sqrt{a}} \sin \frac{n\pi}{2a} x, \quad n = 2, 4, 6, \dots \quad (2.130)$$

For both classes of eigenfunctions it is unnecessary to consider negative values of n , since these lead to solutions which are not linearly independent of those corresponding to positive n . The *energy eigenvalues* E_n are given by

$$E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{\hbar^2 \pi^2 n^2}{8m a^2} = \frac{\hbar^2 \pi^2 n^2}{2m L^2}, \quad n = 1, 2, 3, \dots \quad (2.131)$$

We see that there is an infinite number of discrete energy levels, with one eigenfunction corresponding to each energy level, so that the energy levels are *non-degenerate*. The energy eigenvalues and the corresponding eigenfunctions are shown in Fig. 2.2(b) for the first few states. It is interesting to note that the lowest energy or *zero-point energy* is $E_1 = \hbar^2 \pi^2 / (8ma^2)$ so that there is no state of zero energy. This is in agreement with the requirements of the uncertainty principle. Indeed, the position uncertainty is roughly given by $\Delta x \approx a$. The corresponding momentum uncertainty is therefore $\Delta p_x \approx \hbar/a$, leading to a minimum kinetic energy of order $\hbar^2 / (ma^2)$, in qualitative agreement with the value of E_1 .

There is an important difference between the two classes of eigenfunctions which we have obtained. That is, the eigenfunctions belonging to the first class are such that $\psi_n(-x) = \psi_n(x)$, and are therefore *even* functions of x , while those of the second class are such that $\psi_n(-x) = -\psi_n(x)$, and hence are *odd*. This division of the eigenfunctions into even and odd types is a consequence of the fact that the potential $V(x)$ is symmetric about $x = 0$, $V(-x) = V(x)$, so that the Hamiltonian is invariant under the *parity* or reflection operation $x \rightarrow -x$.

Finally, we observe that a general solution of the time-dependent Schrödinger equation (2.46) for the present problem can be written as a linear superposition of stationary solutions (see (2.85)), namely

$$\Psi(x, t) = \sum_{n=1}^{\infty} c_n \psi_n(x) \exp(-iE_n t / \hbar) \quad (2.132)$$

where the coefficients c_n can be determined from the knowledge of the wave function Ψ at some particular time, say $t = 0$. Thus, using (2.80), we find that

$$c_n = \int_{-a}^{+a} \psi_n^*(x) \Psi(x, t=0) dx \quad (2.133)$$

The linear harmonic oscillator

We now consider the one-dimensional motion of a particle of mass m which is attracted to a fixed centre by a force proportional to the displacement from that centre. Thus, choosing the origin at the centre of force, the restoring force is given by $F = -kx$ (Hooke's law), where k is the force constant. This force can thus be represented by the potential energy

$$V(x) = \frac{1}{2} kx^2 \quad (2.134)$$

which is shown in Fig. 2.3(a). Such a parabolic potential is of great importance in quantum physics as well as in classical physics, since it can be used to approximate an arbitrary continuous potential $W(x)$ in the neighbourhood of a stable equilibrium position $x = a$ (see Fig. 2.3(b)). Indeed, if we expand $W(x)$ in a Taylor series about $x = a$, we find that

$$W(x) = W(a) + (x-a)W'(a) + \frac{1}{2}(x-a)^2W''(a) + \dots \quad (2.135)$$

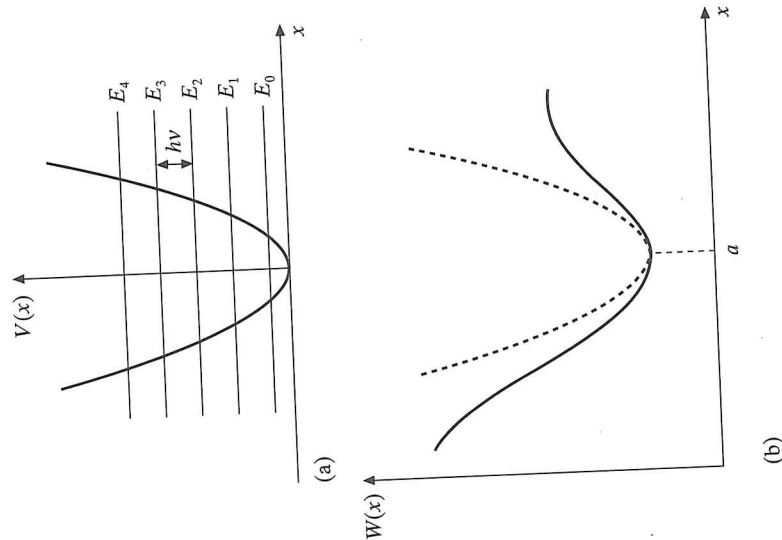


Figure 2.3 (a) The parabolic potential well $V(x) = \frac{1}{2}kx^2$. This is the potential of the linear harmonic oscillator. Also shown are the first few energy eigenvalues (2.143). (b) A continuous potential well $W(x)$, represented by the solid line, can be approximated in the vicinity of a stable equilibrium position at $x = a$ by a linear harmonic oscillator potential, shown as the dashed line.

where $W'(a) = (dW/dx)_{x=a}$ and $W''(a) = (d^2W/dx^2)_{x=a}$. Since $W(x)$ has a minimum at $x = a$, we have $W'(a) = 0$ and $W''(a) > 0$. Choosing a as the coordinate origin and $W(a)$ as the origin of the energy scale, we see that the harmonic oscillator potential (2.134) is the first approximation to $W(x)$. Hence the linear harmonic oscillator is the prototype for systems in which there exist small vibrations about a point of stable equilibrium. This will be illustrated in Chapters 10 and 11, where we shall study the vibrational motion of nuclei in molecules.

The one-dimensional Schrödinger equation (2.121) becomes, for the linear harmonic oscillator,

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + \frac{1}{2} kx^2 \psi(x) = E\psi(x) \quad (2.136)$$

It is convenient to rewrite this equation in terms of dimensionless quantities. To this end we introduce the dimensionless variable

$$\xi = \alpha x, \quad \alpha = \left(\frac{mk}{\hbar^2} \right)^{1/4} \quad (2.137)$$

and the dimensionless eigenvalue

$$\lambda = \frac{2E}{\hbar\omega} \quad (2.138)$$

where

$$\omega = \left(\frac{k}{m} \right)^{1/2} \quad (2.139)$$

is the angular frequency of the corresponding classical oscillator. The Schrödinger equation (2.136) then becomes

$$\frac{d^2\psi}{d\xi^2} + (\lambda - \xi^2)\psi = 0 \quad (2.140)$$

Let us first analyse the behaviour of ψ in the asymptotic region $|\xi| \rightarrow \infty$. We may then neglect the term λ compared to ξ^2 , and the resulting equation is readily seen to have solutions of the form $\xi^n \exp(\pm \xi^2/2)$, where n is finite. The physically acceptable solution must contain only the minus sign in the exponent. This suggests looking for solutions to (2.140) of the form

$$\psi(\xi) = e^{-\xi^2/2} H(\xi) \quad (2.141)$$

where $H(\xi)$ are functions which must not affect the asymptotic behaviour. Substituting (2.141) into (2.140) we obtain for $H(\xi)$ the equation

$$\frac{d^2H}{d\xi^2} - 2\xi \frac{dH}{d\xi} + (\lambda - 1)H = 0 \quad (2.142)$$

Energy levels

Solutions of (2.142) which satisfy the above requirement and are finite everywhere are only found for $\lambda = 2n + 1$, where $n = 0, 1, 2, \dots$. They are called *Hermite polynomials* $H_n(\xi)$. The energy spectrum of the linear harmonic oscillator is therefore given by (see (2.138))

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right) = h\nu \left(n + \frac{1}{2} \right), \quad n = 0, 1, 2, \dots \quad (2.143)$$

where $\nu = \omega/(2\pi)$.

The infinite sequence of discrete, equally spaced energy levels (2.143) is similar to that discovered in 1900 by Planck for the radiation field modes (see Section 1.3). This is due to the fact that a decomposition of the electromagnetic field into normal modes is essentially a decomposition into uncoupled harmonic oscillators. We note, however, that according to (2.143) the linear harmonic oscillator, even in its lowest energy state, $n = 0$, has the energy $\hbar\omega/2$. On the other hand the lowest energy of a classical harmonic oscillator is zero. The finite value $\hbar\omega/2$ of the ground state energy level, which is called the *zero-point energy*, is therefore a purely quantum mechanical effect, and is directly related to the uncertainty principle (Problem 2.12). The eigenvalues (2.143) of the linear harmonic oscillator are *non-degenerate*, since for each eigenvalue there exists only *one* eigenfunction, apart from an arbitrary multiplicative factor.

Hermite polynomials

The Hermite polynomials $H_n(\xi)$, which are the physically acceptable solutions of (2.142) corresponding to the eigenvalues $\lambda = 2n + 1$, are polynomials of order n , having the parity of n , which are uniquely defined except for an arbitrary multiplicative constant. This constant is traditionally chosen so that the highest power of ξ appears with the coefficient 2^n in $H_n(\xi)$. This is consistent with the following definition of the Hermite polynomials:

$$H_n(\xi) = (-1)^n e^{\xi^2} \frac{d^n e^{-\xi^2}}{d\xi^n} \quad (2.144)$$

The first few Hermite polynomials, obtained from (2.144), are

$$\begin{aligned} H_0(\xi) &= 1 \\ H_1(\xi) &= 2\xi \\ H_2(\xi) &= 4\xi^2 - 2 \\ H_3(\xi) &= 8\xi^3 - 12\xi \\ H_4(\xi) &= 16\xi^4 - 48\xi^2 + 12 \\ H_5(\xi) &= 32\xi^5 - 160\xi^3 + 120\xi \end{aligned} \quad (2.145)$$

Another definition of the Hermite polynomials $H_n(\xi)$, which is equivalent to (2.144), involves the use of a *generating function* $G(\xi, s)$. That is,

$$\begin{aligned} G(\xi, s) &= e^{-s^2 + 2s\xi} \\ &= \sum_{n=0}^{\infty} \frac{H_n(\xi)}{n!} s^n \end{aligned} \quad (2.146)$$

The relation (2.146) means that if the function $\exp(-s^2 + 2s\xi)$ is expanded in a power series in s , the coefficients of successive powers of s are just $1/n!$ times the Hermite polynomials $H_n(\xi)$. Using the generating function (2.146) it may be shown that the Hermite polynomials satisfy the recurrence relation

$$H_{n+1}(\xi) - 2\xi H_n(\xi) + 2n H_{n-1}(\xi) = 0 \quad (2.147)$$

The wave functions for the linear harmonic oscillator

Using (2.141), we see that to each of the discrete values E_n of the energy, given by (2.143), there corresponds one, and only one, physically acceptable eigenfunction, namely

$$\psi_n(x) = N_n e^{-\alpha^2 x^2/2} H_n(\alpha x) \quad (2.148)$$

where we have returned to our original variable x , and N_n is a constant which (apart from an arbitrary phase factor) is determined by requiring that the wave function (2.148) be normalised to unity. That is,

$$\int_{-\infty}^{+\infty} |\psi_n(x)|^2 dx = \frac{|N_n|^2}{\alpha} \int_{-\infty}^{+\infty} e^{-\xi^2} H_n^2(\xi) d\xi = 1 \quad (2.149)$$

The integral on the right is evaluated in Appendix 3 by using the generating function (2.146). It is found (see equation (A3.10)) that the normalisation constant N_n can be chosen to be

$$N_n = \left(\frac{\alpha}{\sqrt{\pi} 2^n n!} \right)^{1/2} \quad (2.150)$$

so that the normalised linear harmonic oscillator eigenfunctions are given by

$$\psi_n(x) = \left(\frac{\alpha}{\sqrt{\pi} 2^n n!} \right)^{1/2} e^{-\alpha^2 x^2/2} H_n(\alpha x) \quad (2.151)$$

It is also shown in Appendix 3 that the wave functions $\psi_n(x)$ and $\psi_m(x)$ are orthogonal if $n \neq m$, in agreement with the fact that they correspond to non-degenerate energy eigenvalues.

Because the Hamiltonian of the linear harmonic oscillator,

$$H = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} kx^2 \quad (2.152)$$

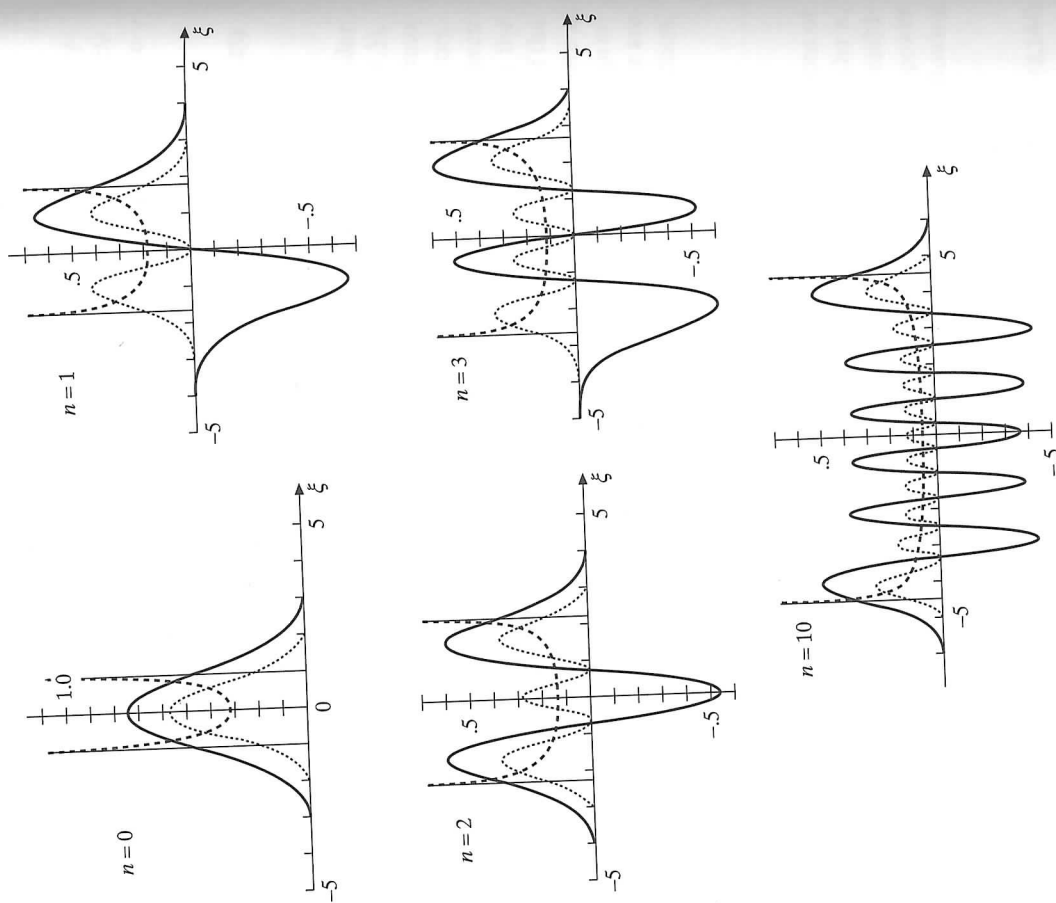


Figure 2.4 Wave functions for the linear harmonic oscillator
 — $\psi_n(\xi)/\sqrt{\alpha}$ with $\xi = \alpha x$, see (2.148) and (2.137)
 $|\psi_n(\xi)|^2/\alpha$
 - - - - The probability distribution for a classical oscillator with energy E_n . The classical motion is confined to the region between the vertical lines.

is invariant under the *parity* operation $x \rightarrow -x$, the eigenfunctions $\psi_n(x)$ are either *even* functions of x (for $n = 0, 2, 4, \dots$) or *odd* functions of x (for $n = 1, 3, 5, \dots$). This is apparent in Fig. 2.4 where plots of the first few harmonic oscillator wave functions are shown, together with the corresponding probability densities $|\psi_n|^2$. It is also clear from this figure that for small values of n the quantum mechanical

probability densities are very different from the corresponding densities for the classical harmonic oscillator. However, the agreement between the classical and quantum mechanical probability densities quickly improves with increasing values of n . This is an illustration of Bohr's *correspondence principle*, according to which the predictions of quantum physics must correspond to those of classical physics in the limit in which the quantum numbers specifying the state of the system become large.

2.5 Angular momentum

The classical orbital angular momentum of a particle is

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \quad (2.153)$$

where $\mathbf{r}$ and $\mathbf{p}$ are the position and momentum vectors of the particle, respectively. Using the fact that $\mathbf{p}$ is represented in wave mechanics by the vector operator $-i\hbar\nabla$, we see that $\mathbf{L}$ is represented by the vector operator $-i\hbar(\mathbf{r} \times \nabla)$. Its Cartesian components are given by

$$L_x = yp_z - zp_y = -i\hbar \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right)$$

$$L_y = zp_x - xp_z = -i\hbar \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) \quad (2.154)$$

$$L_z = xp_y - yp_x = -i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right)$$

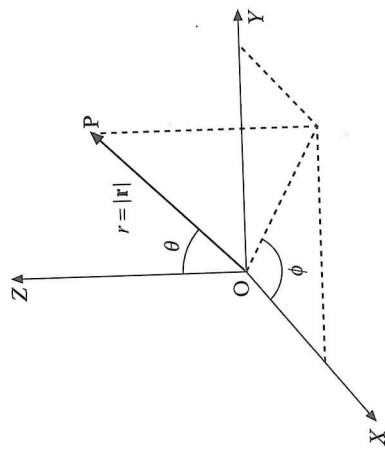
Using the rules (2.120) of commutator algebra and the basic commutation relations (2.119) one finds that

$$[L_x, L_y] = i\hbar L_z, \quad [L_y, L_z] = i\hbar L_x, \quad [L_z, L_x] = i\hbar L_y \quad (2.155)$$

so that the operators L_x , L_y and L_z do not mutually commute. As a consequence, it is impossible to find a representation that diagonalises more than one of them, and the components of the orbital angular momentum cannot in general [4] be assigned definite values simultaneously. However, each of the three components of $\mathbf{L}$ is easily seen to commute with the operator $\mathbf{L}^2 = L_x^2 + L_y^2 + L_z^2$. For example,

$$\begin{aligned} [L_z, \mathbf{L}^2] &= [L_z, L_x^2] + [L_z, L_y^2] \\ &= [L_z, L_x]L_x + L_x[L_z, L_x] + [L_z, L_y]L_y + L_y[L_z, L_y] \\ &= i\hbar(L_y L_x + L_x L_y) - i\hbar(L_x L_y + L_y L_x) = 0 \end{aligned} \quad (2.156)$$

[4] An exception occurs when the angular momentum is zero ($L_x = L_y = L_z = 0$). In this case any function which only depends on the magnitude r of the position vector $\mathbf{r}$ is a simultaneous eigenfunction of L_x , L_y and L_z .

Figure 2.5 Spherical polar coordinates (r, θ, ϕ) of a point P.

Thus it is possible to construct simultaneous eigenfunctions of L^2 and one component of L , so that L^2 and this component can be simultaneously defined precisely. In what follows we shall choose this component to be L_z .

It is convenient to express L_z and L^2 in spherical polar coordinates (r, θ, ϕ) , which are related to the Cartesian coordinates (x, y, z) of the vector $\mathbf{r}$ by

$$\begin{aligned} x &= r \sin \theta \cos \phi \\ y &= r \sin \theta \sin \phi \\ z &= r \cos \theta \end{aligned} \quad (2.157)$$

with $0 \leq r \leq \infty$, $0 \leq \theta \leq \pi$, $0 \leq \phi \leq 2\pi$ (see Fig. 2.5). It is found that (Problem 2.13)

$$L_z = -i\hbar \frac{\partial}{\partial \phi} \quad (2.158)$$

and

$$L^2 = -\hbar^2 \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right] \quad (2.159)$$

The eigenfunctions $\Phi(\phi)$ of L_z satisfy the eigenvalue equation

$$L_z \Phi(\phi) = m\hbar \Phi(\phi) \quad (2.160)$$

where the eigenvalues have been written as $m\hbar$ for convenience. The normalised solutions of (2.160) are

$$\Phi_m(\phi) = \frac{1}{\sqrt{2\pi}} e^{im\phi} \quad (2.161)$$

Since the functions Φ_m must be single-valued, we have $\Phi_m(2\pi) = \Phi_m(0)$, so that m , which is called the *magnetic quantum number*, is restricted to positive or negative

integers, or zero ($m = 0, \pm 1, \pm 2, \dots$). We also note that the functions (2.161) are orthonormal,

$$\int_0^{2\pi} \Phi_m^*(\phi) \Phi_n(\phi) d\phi = \delta_{mn} \quad (2.162)$$

The simultaneous eigenfunctions of L^2 and L_z are called the *spherical harmonics* and are denoted by $Y_{lm}(\theta, \phi)$. They satisfy the eigenvalue equations

$$L^2 Y_{lm}(\theta, \phi) = l(l+1)\hbar^2 Y_{lm}(\theta, \phi) \quad (2.163)$$

and

$$L_z Y_{lm}(\theta, \phi) = m\hbar Y_{lm}(\theta, \phi) \quad (2.164)$$

where the eigenvalues of L^2 have been written as $l(l+1)\hbar^2$. The quantum number l is known as the *orbital angular momentum quantum number*. Setting

$$Y_{lm}(\theta, \phi) = \Theta_{lm}(\theta) \Phi_m(\phi) \quad (2.165)$$

substituting into (2.163) and using (2.159), we find that the functions $\Theta_{lm}(\theta)$ satisfy the equation

$$\left[-\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{m^2}{\sin^2 \theta} \right] \Theta_{lm}(\theta) = l(l+1) \Theta_{lm}(\theta) \quad (2.166)$$

The physically acceptable solutions of this equation that remain finite over the range $0 \leq \theta \leq \pi$ exist only when $l = 0, 1, 2, \dots$ and $m = -l, -l+1, \dots, +l$. They can be expressed in terms of *associated Legendre functions* $P_l^m(\cos \theta)$, which are defined in the following way. Introducing the variable $w = \cos \theta$, we first define the *Legendre polynomials* $P_l(w)$ of degree l by the relation

$$P_l(w) = \frac{1}{2^l l!} \frac{d^l}{dw^l} (w^2 - 1)^l \quad (2.167)$$

An equivalent definition of $P_l(w)$ can be given in terms of a generating function, namely

$$\begin{aligned} T(w, s) &= (1 - 2sw + s^2)^{-1/2} \\ &= \sum_{l=0}^{\infty} P_l(w) s^l, \quad |s| < 1 \end{aligned} \quad (2.168)$$

The Legendre polynomials satisfy the differential equation

$$\left[(1-w^2) \frac{d^2}{dw^2} - 2w \frac{d}{dw} + l(l+1) \right] P_l(w) = 0 \quad (2.169)$$

which is readily shown to be equivalent to (2.166), with $m = 0$. One has the recurrence relation

$$(2l+1)wP_l - (l+1)P_{l+1} - lP_{l-1} = 0 \quad (2.170)$$

which is also valid for $l=0$ if one defines $P_{-1} = 0$. The orthogonality relations read

$$\int_{-1}^{+1} P_l(w)P_l(w)dw = \frac{2}{2l+1}\delta_{ll} \quad (2.171)$$

One also has the closure relation

$$\frac{1}{2} \sum_{l=0}^{\infty} (2l+1)P_l(w)P_l(w') = \delta(w-w') \quad (2.172)$$

Important particular values of the Legendre polynomials are

$$P_l(1) = 1, \quad P_l(-1) = (-1)^l \quad (2.173)$$

For the lowest values of l one has explicitly

$$P_0(w) = 1$$

$$P_1(w) = w$$

$$P_2(w) = \frac{1}{2}(3w^2 - 1)$$

$$P_3(w) = \frac{1}{2}(5w^3 - 3w)$$

$$P_4(w) = \frac{1}{8}(35w^4 - 30w^2 + 3)$$

$$P_5(w) = \frac{1}{8}(63w^5 - 70w^3 + 15w)$$

The associated Legendre functions $P_l^m(w)$ are now defined by the relations

$$P_l^m(w) = (1-w^2)^{m/2} \frac{d^m}{dw^m} P_l(w), \quad m = 0, 1, 2, \dots, l \quad (2.175)$$

They satisfy the recurrence relations

$$(2l+1)wP_l^m = (l+1-m)P_{l+1}^m + (l+m)P_{l-1}^m \quad (2.176a)$$

$$(2l+1)(1-w^2)^{1/2}P_l^{m-1} = P_{l+1}^m - P_{l-1}^m \quad (2.176b)$$

and the orthogonality relations

$$\int_{-1}^{+1} P_l^m(w)P_l^m(w)dw = \frac{2}{2l+1} \frac{(l+m)!}{(l-m)!} \delta_{ll} \quad (2.177)$$

The first few associated Legendre functions are given explicitly by

$$P_1^1(w) = (1-w^2)^{1/2}$$

$$P_2^1(w) = 3(1-w^2)^{1/2}w$$

$$P_2^2(w) = 3(1-w^2)$$

$$P_3^1(w) = \frac{3}{2}(1-w^2)^{1/2}(5w^2-1)$$

$$P_3^2(w) = 15w(1-w^2)$$

$$P_3^3(w) = 15(1-w^2)^{3/2}$$

(2.178)

The functions $\Theta_{lm}(\theta)$, normalised so that

$$\int_0^\pi \Theta_{lm}^*(\theta)\Theta_{lm}(\theta)\sin\theta d\theta = \delta_{ll} \quad (2.179)$$

are given in terms of the associated Legendre functions P_l^m by

$$\begin{aligned} \Theta_{lm}(\theta) &= (-1)^m \left[\frac{(2l+1)(l-m)!}{2(l+m)!} \right]^{1/2} P_l^m(\cos\theta), & m \geq 0 \\ &= (-1)^m \Theta_{l|m|}(\theta), & m < 0 \end{aligned} \quad (2.180)$$

Using (2.165), (2.161) and (2.180), the spherical harmonics are given by

$$Y_{lm}(\theta, \phi) = (-1)^m \left[\frac{(2l+1)(l-m)!}{4\pi(l+m)!} \right]^{1/2} P_l^m(\cos\theta)e^{im\phi}, \quad m \geq 0 \quad (2.181a)$$

$$Y_{l,-m}(\theta, \phi) = (-1)^m Y_{lm}^*(\theta, \phi) \quad (2.181b)$$

They satisfy the orthonormality relations

$$\begin{aligned} \int Y_{lm}^*(\theta, \phi)Y_{lm}(\theta, \phi)d\Omega &= \int_0^{2\pi} d\phi \int_0^\pi \sin\theta d\theta Y_{lm}^*(\theta, \phi)Y_{lm}(\theta, \phi) \\ &= \delta_{ll}\delta_{mm'} \end{aligned} \quad (2.182)$$

where we have written $d\Omega = \sin\theta d\theta d\phi$. The closure relation reads

$$\sum_{l=0}^{\infty} \sum_{m=-l}^{+l} Y_{lm}^*(\theta, \phi)Y_{lm}(\theta', \phi') = \delta(\Omega - \Omega'),$$

$$\delta(\Omega - \Omega') = \frac{\delta(\theta - \theta')\delta(\phi - \phi')}{\sin\theta} \quad (2.183)$$

The first few spherical harmonics are listed in Table 2.1. Polar plots of the probability distributions

$$|Y_{lm}(\theta, \phi)|^2 = (2\pi)^{-1} |\Theta_{lm}(\theta)|^2 \quad (2.184)$$

are shown in Fig. 2.6. Additional useful formulae involving the Legendre polynomials and the spherical harmonics are given in Appendix 4. In that appendix we also discuss *matrix representations* of angular momentum operators, and define the *raising* and *lowering* operators

$$L_{\pm} = L_x \pm iL_y \quad (2.185)$$

These operators are such that (see (A4.21))

$$L_{\pm}Y_{lm}(\theta, \phi) = \hbar[l(l+1) - m(m \pm 1)]^{1/2}Y_{l, m \pm 1}(\theta, \phi) \quad (2.186a)$$

or

$$L_{\pm}|lm\rangle = \hbar[l(l+1) - m(m \pm 1)]^{1/2}|l, m \pm 1\rangle \quad (2.186b)$$

Table 2.1 The first few spherical harmonics $Y_{lm}(\theta, \phi)$.

l	m	Spherical harmonic $Y_{lm}(\theta, \phi)$
0	0	$Y_{0,0} = \frac{1}{(4\pi)^{1/2}}$
1	0	$Y_{1,0} = \left(\frac{3}{4\pi}\right)^{1/2} \cos \theta$
	± 1	$Y_{1,\pm 1} = \mp \left(\frac{3}{8\pi}\right)^{1/2} \sin \theta e^{\pm i\phi}$
2	0	$Y_{2,0} = \left(\frac{5}{16\pi}\right)^{1/2} (3 \cos^2 \theta - 1)$
	± 1	$Y_{2,\pm 1} = \mp \left(\frac{15}{8\pi}\right)^{1/2} \sin \theta \cos \theta e^{\pm i\phi}$
	± 2	$Y_{2,\pm 2} = \left(\frac{15}{32\pi}\right)^{1/2} \sin^2 \theta e^{\pm 2i\phi}$
3	0	$Y_{3,0} = \left(\frac{7}{16\pi}\right)^{1/2} (5 \cos^3 \theta - 3 \cos \theta)$
	± 1	$Y_{3,\pm 1} = \mp \left(\frac{21}{64\pi}\right)^{1/2} \sin \theta (5 \cos^2 \theta - 1) e^{\pm i\phi}$
	± 2	$Y_{3,\pm 2} = \left(\frac{105}{32\pi}\right)^{1/2} \sin^2 \theta \cos \theta e^{\pm 2i\phi}$
	± 3	$Y_{3,\pm 3} = \mp \left(\frac{35}{64\pi}\right)^{1/2} \sin^3 \theta e^{\pm 3i\phi}$

where in the last line we have used the *Dirac ket notation*, in which the eigenfunctions Y_{lm} are written in the form $|lm\rangle$.

Let us assume that the particle is in the orbital angular momentum state $|lm\rangle$ such that $\mathbf{L}^2|lm\rangle = l(l+1)\hbar^2|lm\rangle$ and $L_z|lm\rangle = m\hbar|lm\rangle$ (see (2.163) and (2.164)). Although two components of the orbital angular momentum cannot in general be assigned precise values simultaneously, it is nevertheless possible to say something about the components L_x and L_y . Indeed, one can readily show (Problem 2.14) that

$$\langle lm|L_x|lm\rangle = \langle lm|L_y|lm\rangle = 0 \quad (2.187)$$

and

$$\langle lm|L_x^2|lm\rangle = \langle lm|L_y^2|lm\rangle = \frac{1}{2}[l(l+1) - m^2]\hbar^2 \quad (2.188)$$

We note that when $m = +l$ or $m = -l$, so that the orbital angular momentum is respectively 'parallel' or 'antiparallel' to the Z axis, its x and y components are

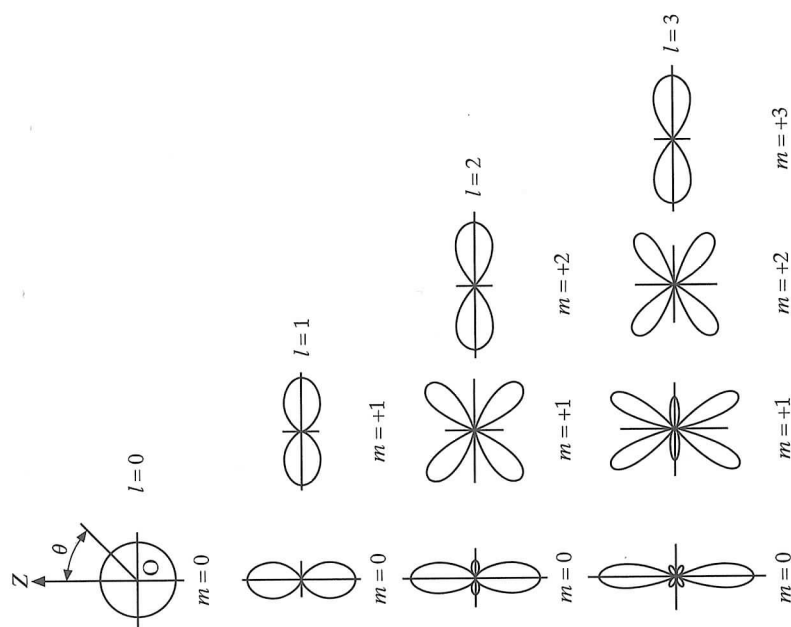


Figure 2.6 Polar plots of the probability distributions $|Y_{lm}(\theta, \phi)|^2 = (2\pi)^{-1} |\Theta_{lm}(\theta)|^2$.

still not zero. It is helpful to visualise these results in terms of the *vector model* introduced in Section 1.8 (see Fig. 1.24). According to this model, the orbital angular momentum vector $\mathbf{L}$, of length $\sqrt{l(l+1)}\hbar$, precesses about the Z axis, the $(2l+1)$ allowed projections of $\mathbf{L}$ on the Z axis being given by $m\hbar$, with $m = -l, -l+1, \dots, +l$.

Spherical harmonics in real form

In some applications it is convenient to use an alternative set of eigenfunctions of $\mathbf{L}^2$, which are the *real forms of the spherical harmonics*. That is,

$$Y_{l,\cos}(\theta, \phi) = N\Theta_{l|m|}(\theta) \cos|m|\phi \quad (2.189)$$

$$Y_{l,\sin}(\theta, \phi) = N\Theta_{l|m|}(\theta) \sin|m|\phi$$

where the normalisation constant N is equal to $\pi^{-1/2}$, except for $m = 0$ for which $N = (2\pi)^{-1/2}$. For $m = 0$, the function $Y_{l,\cos}$ is clearly identical to the spherical harmonic $Y_{l,0}$. On the other hand, for $m \neq 0$ we have

$$Y_{l,\cos} = \frac{1}{\sqrt{2}}(Y_{l|m|} + Y_{l|m|}^*) \quad (2.190)$$

$$Y_{l,\sin} = -\frac{i}{\sqrt{2}}(Y_{l|m|} - Y_{l|m|}^*)$$

The spherical harmonics in real form are eigenfunctions of L^2 and L_z^2 , but not of L_z (except, of course, when $m = 0$). They behave like simple functions of the Cartesian coordinates, and for this reason are well suited for describing the directional properties of chemical bonds (see Chapter 10).

The first few spherical harmonics in real form are listed in Table 2.2. We have used in this table the so-called 'spectroscopic' notation, in which the value of the orbital angular momentum quantum number l is indicated by a letter, according to the correspondence

Value of l	0	1	2	3	4	5
Code letter	s	p	d	f	g	h, ...

Table 2.2 The first few spherical harmonics in real form.

l	$ m $	Spherical harmonic in real form
0	0	$s = \frac{1}{(4\pi)^{1/2}}$
1	0	$p_z = \left(\frac{3}{4\pi}\right)^{1/2} \cos \theta$
	1	$p_x = \left(\frac{3}{4\pi}\right)^{1/2} \sin \theta \cos \phi$ $p_y = \left(\frac{3}{4\pi}\right)^{1/2} \sin \theta \sin \phi$
2	0	$d_{3z^2-r^2} = \left(\frac{5}{16\pi}\right)^{1/2} (3 \cos^2 \theta - 1)$
	1	$d_{xz} = \left(\frac{15}{4\pi}\right)^{1/2} \sin \theta \cos \theta \cos \phi$ $d_{yz} = \left(\frac{15}{4\pi}\right)^{1/2} \sin \theta \cos \theta \sin \phi$
	2	$d_{x^2-y^2} = \left(\frac{15}{16\pi}\right)^{1/2} \sin^2 \theta \cos 2\phi$ $d_{xy} = \left(\frac{15}{16\pi}\right)^{1/2} \sin^2 \theta \sin 2\phi$

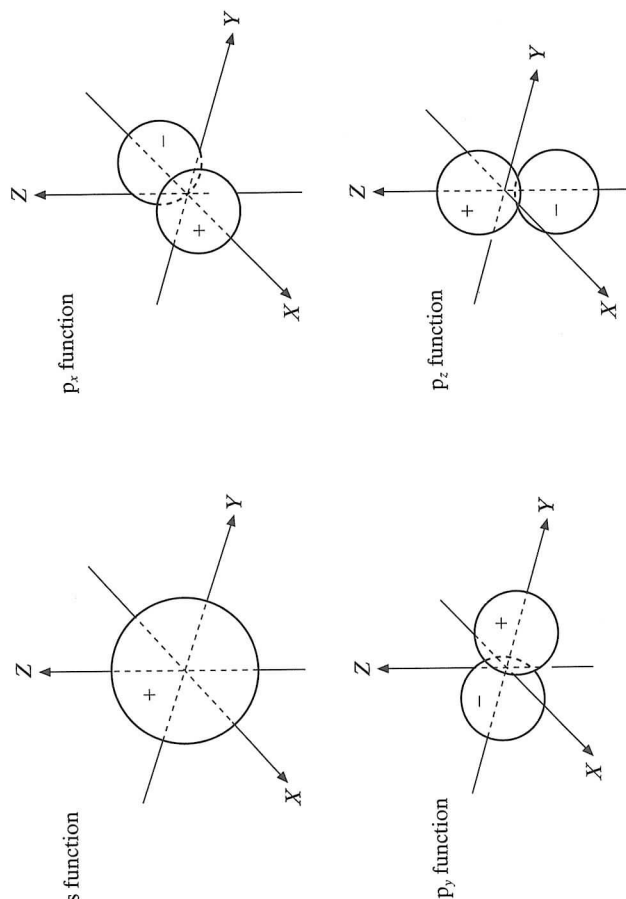


Figure 2.7 Polar diagrams of the real spherical harmonics s , p_x , p_y and p_z .

These code letters are remnants of the spectroscopist's description of various series of spectral lines, the letters s, p, d and f being the first letters of the adjectives 'sharp', 'principal', 'diffuse' and 'fundamental', respectively. For values of l greater than three the letters follow in alphabetical order (g for $l = 4$, h for $l = 5$, and so on). The subscripts x , y , z , etc. used in Table 2.2 indicate the behaviour of the real spherical harmonic in terms of Cartesian coordinates. Polar graphs of the s , p_x , p_y and p_z functions are given in Fig. 2.7. We also remark that while the probability distributions (2.184) corresponding to the 'genuine' spherical harmonics are independent of the azimuthal angle ϕ , those corresponding to the spherical harmonics in real form (that is, $|Y_{l,\cos}|^2$ and $|Y_{l,\sin}|^2$) depend on ϕ in the same way as the functions $\cos^2|m|\phi$ and $\sin^2|m|\phi$, as seen from (2.189).

The rigid rotator

As an example of the use of the operator L^2 and of its eigenfunctions, consider the motion of a particle of mass m , constrained to remain at a given distance R_0 from a fixed point, which we choose as the origin of coordinates. Denoting by $I = mR_0^2$ the moment of inertia, the Hamiltonian of this system, which is known as a (three-dimensional) *rigid rotator*, is

$$H = \frac{L^2}{2I} \quad (2.191)$$

Using the expression (2.159) of L^2 in spherical polar coordinates, the Schrödinger eigenvalue equation reads

$$\frac{L^2}{2I}\psi(\theta, \phi) = -\frac{\hbar^2}{2I}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]\psi(\theta, \phi) = E\psi(\theta, \phi) \quad (2.192)$$

Thus, we see from (2.163) that the eigenfunctions are just the spherical harmonics $Y_{lm}(\theta, \phi)$, and the energy eigenvalues are

$$E = \frac{\hbar^2}{2I}l(l+1), \quad l = 0, 1, 2, \dots \quad (2.193)$$

Spin angular momentum

We saw in Chapter 1 that the Stern-Gerlach experiment can be interpreted as showing that the electron possesses an intrinsic degree of freedom, the *spin*, which behaves like an angular momentum in the way it couples with a magnetic field. The z component of this spin angular momentum can only take on two values $m_s\hbar$, where $m_s = \pm 1/2$. Therefore, the electron spin cannot be described by the orbital angular momentum operator L we have considered thus far, since the z component of L only takes on the values $m\hbar$, with $m = -l, -l+1, \dots, +l$, and $l = 0, 1, 2, \dots$. We shall assume, however, that *all* angular momentum operators, whether orbital or spin, satisfy commutation relations of the form (2.155). Thus, if S_x, S_y and S_z are the three Cartesian components of the electron spin angular momentum operator S , they must satisfy the commutation relations

$$[S_x, S_y] = i\hbar S_z, \quad [S_y, S_z] = i\hbar S_x, \quad [S_z, S_x] = i\hbar S_y \quad (2.194)$$

Using the results of Appendix 4, we may readily obtain the properties of the spin angular momentum and of the spin eigenfunctions. Denoting by χ_{s, m_s} the simultaneous eigenfunctions of S^2 and S_z , we have

$$S^2\chi_{s, m_s} = s(s+1)\hbar^2\chi_{s, m_s} \quad (2.195)$$

and

$$S_z\chi_{s, m_s} = m_s\hbar\chi_{s, m_s} \quad (2.196)$$

Since $m_s = \pm 1/2$ for an electron, we must have $s = 1/2$, and we say that the electron has spin one-half. There are only two different normalised spin eigenfunctions χ_{s, m_s} , namely

$$\alpha \equiv \chi_{1/2, 1/2}; \quad \beta \equiv \chi_{1/2, -1/2} \quad (2.197)$$

and we see from (2.195) and (2.196) that

$$S^2\alpha = \frac{3}{4}\hbar^2\alpha; \quad S^2\beta = \frac{3}{4}\hbar^2\beta \quad (2.198)$$

and

$$S_z\alpha = \frac{\hbar}{2}\alpha; \quad S_z\beta = -\frac{\hbar}{2}\beta \quad (2.199)$$

The spin eigenfunctions α and β are said to correspond respectively to *spin up* ($\uparrow$) and *spin down* ($\downarrow$) states. A general spin-1/2 function χ is an arbitrary linear superposition of the two basic spin states α and β . That is,

$$\chi = \chi_+\alpha + \chi_-\beta \quad (2.200)$$

where χ_+ and χ_- are complex coefficients such that $|\chi_+|^2$ is the probability of finding the electron in the 'spin up' state α , while $|\chi_-|^2$ is the probability of finding it in the 'spin down' state β . The normalisation condition $\langle\chi|\chi\rangle = 1$ gives

$$|\chi_+|^2 + |\chi_-|^2 = 1 \quad (2.201)$$

provided that the basic spin states α and β are orthonormal, namely

$$\langle\alpha|\alpha\rangle = \langle\beta|\beta\rangle = 1 \quad (2.202)$$

$$\langle\alpha|\beta\rangle = \langle\beta|\alpha\rangle = 0$$

We note from (2.198) and (2.200) that for an arbitrary spin-1/2 function χ we have $S^2\chi = (3\hbar^2/4)\chi$, so that S^2 is the purely numerical operator

$$S^2 = \frac{3}{4}\hbar^2 \quad (2.203)$$

Introducing the raising and lowering operators

$$S_{\pm} = S_x \pm iS_y \quad (2.204)$$

and using the general relations (A4.15)–(A4.16) of Appendix 4, with $j = 1/2$ and $m = \pm 1/2$, we have

$$S_+\alpha = 0; \quad S_+\beta = \hbar\alpha \quad (2.205)$$

$$S_-\alpha = \hbar\beta; \quad S_-\beta = 0$$

From (2.199), (2.204) and (2.205) we may construct a table which tells us how the components of S act on α and β . That is,

$$S_x\alpha = \frac{\hbar}{2}\beta; \quad S_x\beta = \frac{\hbar}{2}\alpha$$

$$S_y\alpha = \frac{i\hbar}{2}\beta; \quad S_y\beta = -\frac{i\hbar}{2}\alpha$$

$$S_z\alpha = \frac{\hbar}{2}\alpha; \quad S_z\beta = -\frac{\hbar}{2}\beta \quad (2.206)$$

Using (2.200) and the results of this table, we remark that if χ is an arbitrary spin-1/2 function one has $S_x^2\chi = (\hbar^2/4)\chi$, with a similar result for S_y^2 and S_z^2 .

Thus we have

$$S_x^2 = S_y^2 = S_z^2 = \frac{\hbar^2}{4} \quad (2.207)$$

Since there are only two basic spin states α and β , a *matrix representation* of the spin operators will only require two-by-two matrices. It is apparent from (2.195)–(2.199) that the matrices representing the operators S^2 and S_z may be taken to be the two-by-two diagonal matrices

$$S^2 = \frac{3}{4}\hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \quad S_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (2.208)$$

The normalised spin-1/2 eigenfunctions α and β are given by the two-component column vectors (also called spinors)

$$\alpha = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad \beta = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (2.209)$$

and may be considered as the basis vectors of a two-dimensional 'spin space'. The orthonormality relations (2.202) can then be written in the form

$$\begin{aligned} \alpha^\dagger \alpha &= \beta^\dagger \beta = 1 \\ \alpha^\dagger \beta &= \beta^\dagger \alpha = 0 \end{aligned} \quad (2.210)$$

where the dagger denotes the adjoint. Thus $\alpha^\dagger$ and $\beta^\dagger$ are the row vectors

$$\alpha^\dagger = (1 \ 0); \quad \beta^\dagger = (0 \ 1) \quad (2.211)$$

and according to the rules of matrix multiplication, we have explicitly

$$\alpha^\dagger \alpha = (1 \ 0) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 \quad (2.212)$$

$$\alpha^\dagger \beta = (1 \ 0) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0, \quad \text{etc.}$$

It is also readily verified that the equations (2.205) are satisfied if

$$S_+ = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad S_- = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad (2.213)$$

whence

$$S_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad S_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad (2.214)$$

The results (2.208) and (2.214) can also be written as

$$\mathbf{S} = \frac{\hbar}{2} \boldsymbol{\sigma} \quad (2.215)$$

where

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (2.216)$$

are called the *Pauli spin matrices*. Their principal properties, which can readily be verified (Problem 2.15), are summarised by the equations

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1 \quad (2.217a)$$

$$\sigma_x \sigma_y = -\sigma_y \sigma_x = i\sigma_z; \quad \sigma_y \sigma_z = -\sigma_z \sigma_y = i\sigma_x; \quad \sigma_z \sigma_x = -\sigma_x \sigma_z = i\sigma_y \quad (2.217b)$$

$$\text{Tr } \sigma_x = \text{Tr } \sigma_y = \text{Tr } \sigma_z = 0 \quad (2.217c)$$

$$\det \sigma_x = \det \sigma_y = \det \sigma_z = -1 \quad (2.217d)$$

where Tr means the trace and det the determinant. Moreover, the three Pauli matrices $\sigma_x, \sigma_y, \sigma_z$ and the unit two-by-two matrix form a complete set of 2×2 matrices, in the sense that an arbitrary 2×2 matrix can be expressed in terms of them. Finally, one can prove (Problem 2.15) the identity

$$(\boldsymbol{\sigma} \cdot \mathbf{A})(\boldsymbol{\sigma} \cdot \mathbf{B}) = \mathbf{A} \cdot \mathbf{B} + i\boldsymbol{\sigma} \cdot (\mathbf{A} \times \mathbf{B}) \quad (2.218)$$

where $\mathbf{A}$ and $\mathbf{B}$ are any two vectors, or two vector operators whose components commute with those of the spin $\mathbf{S}$ (that is, with those of $\boldsymbol{\sigma}$). In the latter case the order of $\mathbf{A}$ and $\mathbf{B}$ on both sides of (2.218) must be respected.

Using the explicit form (2.209) of the basic spinors α and β , an arbitrary spin-1/2 function (2.200) may be written as the spinor

$$\chi = \begin{pmatrix} \chi_+ \\ \chi_- \end{pmatrix} \quad (2.219)$$

and the normalisation condition (2.201) becomes

$$\chi^\dagger \chi = 1 \quad (2.220)$$

where $\chi^\dagger$ denotes the adjoint of the spinor χ , namely

$$\chi^\dagger = (\chi_+^* \ \chi_-^*) \quad (2.221)$$

It is worth noting that if the electron is in a pure 'spin up' state α or a pure 'spin down' state β , the expectation values of S_x and S_y vanish, $\langle S_x \rangle = \langle S_y \rangle = 0$, while $\langle S_x^2 \rangle = \langle S_y^2 \rangle = \hbar^2/4$ (see Problem 2.14). Thus, even when the spin angular momentum $\mathbf{S}$ is said to be 'up' ($m_s = +1/2$) or 'down' ($m_s = -1/2$), its x and y components are still not zero. As in the case of the orbital angular momentum $\mathbf{L}$, these results can be visualised with the help of a vector model. According to this model, the spin vector $\mathbf{S}$, of length $\sqrt{3}/4 \hbar$, precesses about the Z axis, the only allowed projections of $\mathbf{S}$ on the Z axis being $m_s \hbar$, with $m_s = \pm 1/2$. This is illustrated in Fig. 2.8.

Thus far we have only considered the spin of the electron. Other particles, such as the atomic nuclei, may also possess a spin angular momentum $\mathbf{S}$, for which the quantum number s (see (2.195)) can be either integer or half-odd integer.

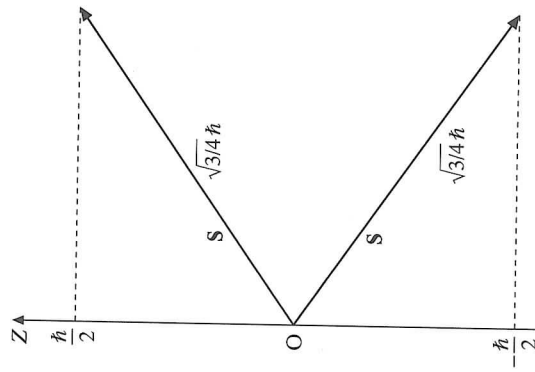


Figure 2.8 The vector model of the spin, for a spin-1/2 particle.

From the results of Appendix 4, with $j = s$ and $m = m_s$, we see that the quantum number m_s takes on the $2s + 1$ values $-s, -s + 1, \dots, s - 1, s$. The matrices representing the spin operators S_x, S_y, S_z and S^2 then have dimensions $2s + 1$; they can be written explicitly by using the methods of Appendix 4 (see Problem 2.16).

Total angular momentum

The total angular momentum of a particle can be written as

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \quad (2.222)$$

The orbital angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ operates only in 'ordinary' space and satisfies the commutation relations (2.155). On the other hand, the spin angular momentum $\mathbf{S}$ satisfies the commutation relations (2.194) and operates only in 'spin space'. All its components therefore commute with those of $\mathbf{r}$ and $\mathbf{p}$, and hence with all those of $\mathbf{L}$. As a result, the total angular momentum $\mathbf{J}$ satisfies the commutation relations

$$[J_x, J_y] = i\hbar J_z, \quad [J_y, J_z] = i\hbar J_x, \quad [J_z, J_x] = i\hbar J_y \quad (2.223)$$

which characterise an angular momentum operator.

It is shown in Appendix 4 that the simultaneous eigenfunctions of $\mathbf{J}^2 = J_x^2 + J_y^2 + J_z^2$ and J_z^2 satisfy the eigenvalue equations

$$\mathbf{J}^2 \psi_{jm_j} = j(j+1)\hbar^2 \psi_{jm_j} \quad (2.224)$$

and

$$J_z \psi_{jm_j} = m_j \hbar \psi_{jm_j} \quad (2.225)$$

where j is an integer (including zero) or a half-odd integer ($j = 0, 1/2, 1, 3/2, \dots$) and $m_j = -j, -j + 1, \dots, j - 1, j$.

Since all the components of $\mathbf{L}$ commute with all those of $\mathbf{S}$, the operators $\mathbf{L}^2, L_z, \mathbf{S}^2$ and S_z mutually commute, and have the simultaneous eigenfunctions

$$\psi_{lmjm_s} = Y_{lm}(\theta, \phi) \chi_{sm_s} \quad (2.226)$$

where we have written $m_l \equiv m$ [5]. It can be shown (Problem 2.17) that the four operators $\mathbf{L}^2, \mathbf{S}^2, \mathbf{J}^2$ and J_z also form a commuting set of operators. Their simultaneous eigenfunctions are linear combinations of the functions ψ_{lmjm_s} , and are often denoted by the symbol ψ_{lsmj} (see Appendix 4). For a given value of l and s the possible values of j are given by

$$|l - s|, |l - s| + 1, \dots, l + s \quad (2.227)$$

For a given j the quantum number m_j can take on the $2j + 1$ values $-j, -j + 1, \dots, j - 1, j$, as we have seen above.

Wave functions for a spin-1/2 particle

Thus far in this section we have focused our attention on the angular and spin parts of the wave functions. In general, the wave functions also depend on the radial coordinate r and on the time t . For example, in the case of a spin-1/2 particle (for example, an electron), a general expression for the wave function is

$$\Psi(\mathbf{r}, t) = \Psi_+(\mathbf{r}, t)\alpha + \Psi_-(\mathbf{r}, t)\beta \quad (2.228)$$

where q denotes the ensemble of the (continuous) spatial variables $\mathbf{r}$ and the (discrete) spin variable ($m_s = \pm 1/2$) of the particle. The probability density for finding at time t the particle at $\mathbf{r}$ with 'spin up' is $|\Psi_+(\mathbf{r}, t)|^2$, and with 'spin down' it is $|\Psi_-(\mathbf{r}, t)|^2$. Using (2.209) we may also write the wave function Ψ as a two-component or spinor wave function, namely

$$\Psi = \begin{pmatrix} \Psi_+ \\ \Psi_- \end{pmatrix} \quad (2.229)$$

Central forces

Let us now return to the Schrödinger equation (2.59) describing the motion of a spinless particle of mass m in a time-independent potential $V(\mathbf{r})$. We shall consider in this section the important case of *central* potentials, that is potentials $V(r)$ which depend only upon the magnitude $r = |\mathbf{r}|$ of the vector $\mathbf{r}$. Since $V(r)$ is spherically symmetric, it is natural to use the spherical polar coordinates defined in (2.157). The Hamiltonian of the system is then given by

[5] In general we shall use the notation m for a magnetic quantum number, but where it is important to distinguish between different kinds of angular momenta, we shall use the notation m_l, m_s, m_j , as necessary.

$$\begin{aligned}
H &= -\frac{\hbar^2}{2m} \nabla^2 + V(r) \\
&= -\frac{\hbar^2}{2m} \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) \right] \\
&\quad + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + V(r)
\end{aligned} \tag{2.230}$$

Using the expression (2.159) of $\mathbf{L}^2$, we may also write

$$H = -\frac{\hbar^2}{2m} \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{\mathbf{L}^2}{\hbar^2 r^2} \right] + V(r) \tag{2.231}$$

so that the Schrödinger equation (2.59) reads

$$\left\{ -\frac{\hbar^2}{2m} \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{\mathbf{L}^2}{\hbar^2 r^2} \right] + V(r) \right\} \psi(\mathbf{r}) = E \psi(\mathbf{r}) \tag{2.232}$$

Now $\mathbf{L}^2$ and L_z only operate on angular variables, and in addition $[\mathbf{L}^2, L_z] = 0$. Therefore, we see from (2.231) that

$$[H, \mathbf{L}^2] = [H, L_z] = 0 \tag{2.233}$$

We may thus look for solutions of the Schrödinger equation (2.232) which are simultaneous eigenfunctions of the operators H , $\mathbf{L}^2$ and L_z . Since the spherical harmonics $Y_{lm}(\theta, \phi)$ are simultaneous eigenfunctions of $\mathbf{L}^2$ and L_z (see (2.163) and (2.164)) we can write a particular solution as

$$\psi_{E,l,m}(r, \theta, \phi) = R_{E,l}(r) Y_{lm}(\theta, \phi) \tag{2.234}$$

Substituting (2.234) into (2.232), and using (2.163), we obtain for the radial function $R_{E,l}(r)$ the equation

$$\left\{ -\frac{\hbar^2}{2m} \left[\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d}{dr} \right) - \frac{l(l+1)}{r^2} \right] + V(r) \right\} R_{E,l}(r) = E R_{E,l}(r) \tag{2.235}$$

which shows that $R_{E,l}$ does not depend on the magnetic quantum number m .

The radial equation (2.235) can be simplified by introducing the new radial function

$$u_{E,l}(r) = r R_{E,l}(r) \tag{2.236}$$

The new radial equation which we obtain for $u_{E,l}(r)$ is then

$$\left[-\frac{\hbar^2}{2m} \frac{d^2}{dr^2} + \frac{l(l+1)\hbar^2}{2mr^2} + V(r) \right] u_{E,l}(r) = E u_{E,l}(r) \tag{2.237}$$

For potentials which are less singular than r^{-2} at the origin, a power series expansion of $u_{E,l}(r)$ can be made for small r , and the examination of the indicial equation shows that

$$u_{E,l}(r) \underset{r \rightarrow 0}{\sim} r^{l+1} \tag{2.238}$$

The equation (2.237) is similar to the one-dimensional Schrödinger equation (2.121) for an effective potential

$$V_{\text{eff}}(r) = V(r) + \frac{l(l+1)\hbar^2}{2mr^2} \tag{2.239}$$

which contains the repulsive centrifugal barrier term $l(l+1)\hbar^2/(2mr^2)$ in addition to $V(r)$. We remark, however, that the variable r is confined to positive values $0 \leq r < \infty$, in contrast to the variable x in (2.121).

Parity

The parity operator $\mathcal{P}$ is defined by the relation

$$\mathcal{P}f(\mathbf{r}) = f(-\mathbf{r}) \tag{2.240}$$

where $f(\mathbf{r})$ is an arbitrary function. Thus the parity operator corresponds to an inversion of the position coordinate $\mathbf{r}$ through the origin. It is a Hermitian operator since, for any two wave functions $\phi(\mathbf{r})$ and $\psi(\mathbf{r})$, we have

$$\begin{aligned}
\int \phi^*(\mathbf{r}) \mathcal{P} \psi(\mathbf{r}) d\mathbf{r} &= \int \phi^*(\mathbf{r}) \psi(-\mathbf{r}) d\mathbf{r} \\
&= \int \phi^*(-\mathbf{r}) \psi(\mathbf{r}) d\mathbf{r} \\
&= \int [\mathcal{P} \phi(\mathbf{r})]^* \psi(\mathbf{r}) d\mathbf{r}
\end{aligned} \tag{2.241}$$

Let us now consider the eigenvalue equation for $\mathcal{P}$, which we write as

$$\mathcal{P} \psi_\alpha(\mathbf{r}) = \alpha \psi_\alpha(\mathbf{r}) \tag{2.242}$$

From the definition (2.240) of $\mathcal{P}$ and the fact that $\mathcal{P}f(-\mathbf{r}) = f(\mathbf{r})$, we deduce that

$$\mathcal{P}^2 = I \tag{2.243}$$

where I is the unit operator. Hence

$$\mathcal{P}^2 \psi_\alpha(\mathbf{r}) = \alpha \mathcal{P} \psi_\alpha(\mathbf{r}) = \alpha^2 \psi_\alpha(\mathbf{r}) = \psi_\alpha(\mathbf{r}) \tag{2.244}$$

so that $\alpha^2 = 1$ and the eigenvalues of $\mathcal{P}$ are $\alpha = \pm 1$. Denoting the corresponding eigenfunctions by ψ_+ and ψ_- , we have

$$\mathcal{P} \psi_+(\mathbf{r}) = \psi_+(\mathbf{r}), \quad \mathcal{P} \psi_-(\mathbf{r}) = -\psi_-(\mathbf{r}) \tag{2.245}$$

or

$$\psi_+(-\mathbf{r}) = \psi_+(\mathbf{r}), \quad \psi_-(-\mathbf{r}) = -\psi_-(\mathbf{r}) \quad (2.246)$$

Thus $\psi_+(\mathbf{r})$ is an *even* function of $\mathbf{r}$, and $\psi_-(\mathbf{r})$ is an *odd* function of $\mathbf{r}$. The eigenfunctions ψ_+ are said to have *even parity*, while the eigenfunctions ψ_- have *odd parity*. We note that ψ_+ and ψ_- are orthogonal. They also form a complete set, since any function $\psi(\mathbf{r})$ can always be written as

$$\psi(\mathbf{r}) = \psi_+(\mathbf{r}) + \psi_-(\mathbf{r}) \quad (2.247)$$

where

$$\psi_+(\mathbf{r}) = \frac{1}{2}[\psi(\mathbf{r}) + \psi(-\mathbf{r})] \quad (2.248)$$

has obviously even parity, while

$$\psi_-(\mathbf{r}) = \frac{1}{2}[\psi(\mathbf{r}) - \psi(-\mathbf{r})] \quad (2.249)$$

has odd parity.

Let us now return to the central force problem. Under the parity operation $\mathbf{r} \rightarrow -\mathbf{r}$, the spherical polar coordinates (r, θ, ϕ) become $(r, \pi - \theta, \phi + \pi)$. The central force Hamiltonian (2.230) is clearly unaffected by this operation, or in other words the parity operator $\mathcal{P}$ commutes with the Hamiltonian (2.230)

$$[\mathcal{P}, H] = 0 \quad (2.250)$$

As a result, simultaneous eigenfunctions of the operators $\mathcal{P}$ and H can be found. Applying the parity operator on the wave function $\psi_{E,l,m}(r, \theta, \phi) = R_{E,l}(r)Y_{lm}(\theta, \phi)$ (see (2.234)), we have

$$\mathcal{P}[R_{E,l}(r)Y_{lm}(\theta, \phi)] = R_{E,l}(r)Y_{lm}(\pi - \theta, \phi + \pi) \quad (2.251)$$

Now, from the definition (2.181) of the spherical harmonics, it can be shown that

$$Y_{lm}(\pi - \theta, \phi + \pi) = (-1)^l Y_{lm}(\theta, \phi) \quad (2.252)$$

so that Y_{lm} has the parity of l . Thus

$$\mathcal{P}[R_{E,l}(r)Y_{lm}(\theta, \phi)] = R_{E,l}(r)(-1)^l Y_{lm}(\theta, \phi) \quad (2.253)$$

and the wave function $\psi_{E,l,m}$ itself has the parity of l (even for even l , odd for odd l).

The free particle

We shall discuss at length various applications involving central potentials in subsequent chapters. For further reference, we consider here the (very) special case of the free particle, for which $V(r) = 0$. Writing $k = \sqrt{2mE}/\hbar$, the radial equation (2.235) becomes

$$\left[\frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} - \frac{l(l+1)}{r^2} + k^2 \right] R_{E,l}(r) = 0 \quad (2.254)$$

It is convenient to set $\rho = kr$ and to write $R_l(\rho) \equiv R_{E,l}(r)$, so that the above equation reads

$$\left[\frac{d^2}{d\rho^2} + \frac{2}{\rho} \frac{d}{d\rho} + \left(1 - \frac{l(l+1)}{\rho^2} \right) \right] R_l(\rho) = 0 \quad (2.255)$$

This equation is known as the *spherical Bessel differential equation*. The solutions of (2.255) which are regular and finite everywhere are given (up to a multiplicative constant) by the spherical Bessel functions $j_l(\rho)$, defined by

$$j_l(\rho) = \left(\frac{\pi}{2\rho} \right)^{1/2} J_{l+\frac{1}{2}}(\rho) \quad (2.256)$$

where $J_\nu(\rho)$ is a Bessel function of order ν . The first few spherical Bessel functions are given explicitly by

$$j_0(\rho) = \frac{\sin \rho}{\rho}$$

$$j_1(\rho) = \frac{\sin \rho}{\rho^2} - \frac{\cos \rho}{\rho} \quad (2.257)$$

$$j_2(\rho) = \left(\frac{3}{\rho^3} - \frac{1}{\rho} \right) \sin \rho - \frac{3}{\rho^2} \cos \rho$$

Additional properties of the functions $j_l(\rho)$ will be discussed in Chapter 12.

The eigenvalues k^2 of (2.254) can take on any positive real value. As a result, the energy $E = \hbar^2 k^2 / (2m)$ can assume any value in the interval $(0, \infty)$, giving an example of a *continuous spectrum*. Thus, returning to (2.234) and using the foregoing results, we see that for every positive value of E there exist eigenfunctions of the free particle Schrödinger equation labelled by the orbital angular momentum quantum numbers (l, m) , namely

$$\psi_{E,l,m}(r, \theta, \phi) = C j_l(kr) Y_{lm}(\theta, \phi) \quad (2.258)$$

where C is a constant. It can be shown that the ensemble of *spherical waves* (2.258) forms a complete set.

On the other hand, we have seen in Section 2.1 that a free particle of momentum $\hbar \mathbf{k}$ and energy $E = \hbar^2 k^2 / (2m)$ is represented by the plane wave $\exp(i\mathbf{k} \cdot \mathbf{r})$. Since the spherical waves (2.258) form a complete set, we may expand the plane wave $\exp(i\mathbf{k} \cdot \mathbf{r})$ in terms of them. That is,

$$\exp(i\mathbf{k} \cdot \mathbf{r}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} c_{lm}(\mathbf{k}) j_l(kr) Y_{lm}(\theta, \phi) \quad (2.259)$$

If we choose the Z axis to be along the wave vector $\mathbf{k}$, then the left-hand side of (2.259) reads $\exp(i\mathbf{k} \cdot \mathbf{r}) = \exp(ikr \cos \theta)$, which is independent of ϕ . Setting $w = \cos \theta$, the expansion of $\exp(i\mathbf{k} \cdot \mathbf{r})$ reduces to an expansion in terms of

Legendre polynomials $P_l(w)$. A straightforward calculation (Problem 2.18) then yields

$$\exp(i\mathbf{k} \cdot \mathbf{r}) = \sum_{l=0}^{\infty} (2l+1) i^l j_l(kr) P_l(\cos \theta) \quad (2.260)$$

Using the addition theorem of the spherical harmonics (see the equation (A4.23) of Appendix 4) we can also write the above formula in the form of equation (2.259). That is,

$$\exp(i\mathbf{k} \cdot \mathbf{r}) = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} i^l j_l(kr) Y_{lm}^*(\hat{\mathbf{k}}) Y_{lm}(\hat{\mathbf{r}}) \quad (2.261)$$

where $\hat{\mathbf{x}}$ denotes the polar angles of a vector $\mathbf{x}$. Upon comparison of (2.259) and (2.261) we see that the coefficients $c_{lm}(\mathbf{k})$ are given by $c_{lm}(\mathbf{k}) = 4\pi i^l Y_{lm}^*(\hat{\mathbf{k}})$.

2.7 Several-particle systems

Until now we have only discussed the motion of a single particle. In this section we shall generalise our results to N -particle systems. We begin by considering a system of N non-relativistic *spinless* particles having masses m_i , position coordinates $\mathbf{r}_i$ and momenta $\mathbf{p}_i$ ($i = 1, 2, \dots, N$). The classical Hamiltonian function of this system is

$$H(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N, t) = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} + V(\mathbf{r}_1, \dots, \mathbf{r}_N, t) \quad (2.262)$$

where V is the potential energy. The total classical energy E_{tot} of the system is

$$E_{\text{tot}} = H(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N, t) \quad (2.263)$$

The Schrödinger equation for the wave function $\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N, t)$ which describes the dynamical state of the system is obtained by considering E_{tot} and $\mathbf{p}_i$ to be the differential operators (compare with (2.11))

$$E_{\text{op}} = i\hbar \frac{\partial}{\partial t}, \quad (\mathbf{p}_i)_{\text{op}} = -i\hbar \nabla_{\mathbf{r}_i} \quad (2.264)$$

and writing that E_{op} and $H_{\text{op}} \equiv H(\mathbf{r}_1, \dots, \mathbf{r}_N, -i\hbar \nabla_1, \dots, -i\hbar \nabla_N, t)$ give identical results when acting on Ψ . That is,

$$i\hbar \frac{\partial}{\partial t} \Psi = H_{\text{op}} \Psi = \left[\sum_{i=1}^N \left(-\frac{\hbar^2}{2m_i} \nabla_{\mathbf{r}_i}^2 \right) + V(\mathbf{r}_1, \dots, \mathbf{r}_N, t) \right] \Psi \quad (2.265)$$

The normalisation condition (2.26) now takes the form

$$\int |\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N, t)|^2 d\mathbf{r}_1 \dots d\mathbf{r}_N = 1 \quad (2.266)$$

The total orbital angular momentum $\mathbf{L}$ of the system is the sum of the individual orbital angular momenta $\mathbf{L}_i = \mathbf{r}_i \times \mathbf{p}_i$:

$$\mathbf{L} = \sum_{i=1}^N \mathbf{L}_i \quad (2.267)$$

It is also worth noting that all the components of the position $\mathbf{r}_i$ and momentum $\mathbf{p}_i$ of particle i commute with all those pertaining to particle j provided that $i \neq j$, so that the fundamental commutation relations read

$$[x_i, p_j] = i\hbar \delta_{ij}, \quad [y_i, p_j] = 0, \text{ etc.} \quad (2.268)$$

Two-body systems

Of particular interest is a system of two particles, of masses m_1 and m_2 , interacting via a time-independent potential $V(\mathbf{r}_1 - \mathbf{r}_2)$ which depends only upon the relative coordinate $\mathbf{r}_1 - \mathbf{r}_2$. The classical Hamiltonian of the system is therefore given by

$$H = \frac{\mathbf{p}_1^2}{2m_1} + \frac{\mathbf{p}_2^2}{2m_2} + V(\mathbf{r}_1 - \mathbf{r}_2) \quad (2.269)$$

Making the substitutions $\mathbf{p}_1 \rightarrow -i\hbar \nabla_1$ and $\mathbf{p}_2 \rightarrow -i\hbar \nabla_2$ in (2.269), we obtain the quantum mechanical Hamiltonian operator, and the corresponding Schrödinger equation reads

$$i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}_1, \mathbf{r}_2, t) = \left[-\frac{\hbar^2}{2m_1} \nabla_1^2 - \frac{\hbar^2}{2m_2} \nabla_2^2 + V(\mathbf{r}_1 - \mathbf{r}_2) \right] \Psi(\mathbf{r}_1, \mathbf{r}_2, t) \quad (2.270)$$

Because the potential V only depends on the difference of coordinates $\mathbf{r}_1 - \mathbf{r}_2$, an important simplification can now be made. We introduce the relative coordinate

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 \quad (2.271)$$

together with the vector

$$\mathbf{R} = \frac{m_1 \mathbf{r}_1 + m_2 \mathbf{r}_2}{m_1 + m_2} \quad (2.272)$$

which determines the position of the centre of mass (CM) of the system. Changing variables from the coordinates $(\mathbf{r}_1, \mathbf{r}_2)$ to the new coordinates $(\mathbf{r}, \mathbf{R})$ one finds that the Schrödinger equation (2.270) becomes

$$i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{R}, \mathbf{r}, t) = \left[-\frac{\hbar^2}{2M} \nabla_{\mathbf{R}}^2 - \frac{\hbar^2}{2\mu} \nabla_{\mathbf{r}}^2 + V(\mathbf{r}) \right] \Psi(\mathbf{R}, \mathbf{r}, t) \quad (2.273)$$

where

$$M = m_1 + m_2 \quad (2.274)$$

is the total mass of the system and

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \quad (2.275)$$

is the reduced mass of the two particles. The Schrödinger equation (2.273) may also be obtained by introducing the relative momentum

$$\mathbf{p} = \frac{m_2 \mathbf{p}_1 - m_1 \mathbf{p}_2}{m_1 + m_2} \quad (2.276)$$

together with the total momentum

$$\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2 \quad (2.277)$$

Since

$$\frac{\mathbf{p}_1^2}{2m_1} + \frac{\mathbf{p}_2^2}{2m_2} = \frac{\mathbf{P}^2}{2M} + \frac{\mathbf{p}^2}{2\mu} \quad (2.278)$$

the classical Hamiltonian (2.269) can be written as

$$H = \frac{\mathbf{P}^2}{2M} + \frac{\mathbf{p}^2}{2\mu} + V(\mathbf{r}) \quad (2.279)$$

Performing the substitutions $\mathbf{P} \rightarrow -i\hbar \nabla_{\mathbf{R}}$ and $\mathbf{p} \rightarrow -i\hbar \nabla_{\mathbf{r}}$ in (2.279), we then obtain the quantum mechanical Hamiltonian operator leading to the Schrödinger equation (2.273).

Two separations of the equation (2.273) can now be made. The time dependence can first be separated as in (2.58) since the potential is time-independent. Secondly, the spatial part of the wave function can be separated into a product of functions of the centre of mass coordinate $\mathbf{R}$ and of the relative coordinate $\mathbf{r}$. Thus the Schrödinger equation (2.273) admits solutions of the form

$$\Psi(\mathbf{R}, \mathbf{r}, t) = \Phi(\mathbf{R}) \psi(\mathbf{r}) \exp[-i(E_{\text{CM}} + E)t/\hbar] \quad (2.280)$$

where the functions $\Phi(\mathbf{R})$ and $\psi(\mathbf{r})$ satisfy respectively the equations

$$-\frac{\hbar^2}{2M} \nabla_{\mathbf{R}}^2 \Phi(\mathbf{R}) = E_{\text{CM}} \Phi(\mathbf{R}) \quad (2.281)$$

and

$$\left[-\frac{\hbar^2}{2\mu} \nabla_{\mathbf{r}}^2 + V(\mathbf{r}) \right] \psi(\mathbf{r}) = E \psi(\mathbf{r}) \quad (2.282)$$

We see that the equation (2.281) is a time-independent Schrödinger equation describing the centre of mass as a free particle of mass M and energy E_{CM} . The second time-independent Schrödinger equation (2.282) describes the relative motion of the two particles; it is the same as the equation corresponding to the

motion of a particle having the reduced mass μ in the potential $V(\mathbf{r})$. The total energy of the system is clearly

$$E_{\text{tot}} = E_{\text{CM}} + E \quad (2.283)$$

We have therefore 'decoupled' the original two-body problem into two one-body problems, that of a free particle (the centre of mass) and that of a single particle of reduced mass μ in a potential $V(\mathbf{r})$. We remark that if we elect to work in the centre of mass system of the two particles, we need not be concerned with the motion of the centre of mass, the coordinates of which are eliminated.

Systems of particles with spin. Addition of angular momenta

Let us now consider a system of particles which may possess spin. We shall denote by q_i the ensemble of space and spin coordinates of the particle i . The angular momentum $\mathbf{J}_i$ of that particle is the sum of its orbital angular momentum $\mathbf{L}_i$ and its spin angular momentum $\mathbf{S}_i$,

$$\mathbf{J}_i = \mathbf{L}_i + \mathbf{S}_i \quad (2.284)$$

The total orbital angular momentum $\mathbf{L}$ of the system is given by (2.267). The total spin angular momentum is

$$\mathbf{S} = \sum_{i=1}^N \mathbf{S}_i \quad (2.285)$$

Finally, the total angular momentum of the system is the sum of the individual angular momenta $\mathbf{J}_i$,

$$\mathbf{J} = \sum_{i=1}^N \mathbf{J}_i \quad (2.286)$$

In many problems of quantum physics the Hamiltonian H of the system is rotationally invariant. As a consequence, it may be shown [6] that H commutes with the components of the total angular momentum $\mathbf{J}$, so that we can look for the eigenfunctions of H among the simultaneous eigenfunctions of $\mathbf{J}^2$ and J_z . On the other hand, we know in general how to obtain the eigenfunctions of the individual angular momentum operators $\mathbf{J}_i^2, J_{iz}$ (see Appendix 4). The problem of the addition of angular momenta consists in obtaining the eigenvalues and eigenfunctions of $\mathbf{J}^2$ and J_z in terms of those of $\mathbf{J}_i^2$ and J_{iz} . The simplest addition problem, namely that of adding two angular momenta, is discussed in Appendix 4.

Indistinguishable particles

Two particles are said to be *identical* when they cannot be distinguished by means of any intrinsic property. Many atomic or molecular systems contain a number of

[6] See Bransden and Joachain (2000).

particles (notably electrons) which are all identical. While in classical physics the existence of sharp trajectories makes it possible, in principle, to distinguish 'classical' particles by their paths, in quantum mechanics there is no way of keeping track of individual particles when the wave functions of identical particles overlap. Thus, it is impossible in quantum mechanics to distinguish between identical particles in regions of space where they may be found simultaneously, such as their interaction region. As we shall now see, this quantum mechanical indistinguishability of identical particles has profound consequences.

Let us consider a quantum mechanical system of N identical particles. Since the N particles are identical, all observables corresponding to this system must be symmetric functions of the basic dynamical variables. In particular, the Hamiltonian H of the system must be symmetric with respect to any interchange of the space and spin coordinates of the particles. Thus an interchange operator P_{ij} that permutes the variables q_i and q_j of the particles i and j commutes with the Hamiltonian:

$$[P_{ij}, H] = 0 \quad (2.287)$$

In general, an exact eigenfunction $\psi(q_1, \dots, q_N)$ of H has no particular symmetry property under the interchange of the variables q_i . However, it is important to recognise that if $\psi(q_1, \dots, q_N)$ is an eigenfunction of H corresponding to the eigenvalue E , then so is $P_{ij}\psi$, where

$$P_{ij}\psi(q_1, \dots, q_i, \dots, q_j, \dots, q_N) = \psi(q_1, \dots, q_j, \dots, q_i, \dots, q_N) \quad (2.288)$$

Since two successive interchanges of q_i and q_j bring the particles back to their initial configurations, we have

$$P_{ij}^2 = I \quad (2.289)$$

so that the eigenvalues of the operator P_{ij} are $\epsilon = \pm 1$. Wave functions corresponding to the eigenvalue $\epsilon = +1$ are such that

$$P_{ij}\psi(q_1, \dots, q_i, \dots, q_j, \dots, q_N) = \psi(q_1, \dots, q_j, \dots, q_i, \dots, q_N) = \psi(q_1, \dots, q_i, \dots, q_j, \dots, q_N) \quad (2.290)$$

and are said to be *symmetric* under the interchange P_{ij} . On the other hand, wave functions which correspond to the eigenvalue $\epsilon = -1$ are such that

$$P_{ij}\psi(q_1, \dots, q_i, \dots, q_j, \dots, q_N) = \psi(q_1, \dots, q_j, \dots, q_i, \dots, q_N) = -\psi(q_1, \dots, q_i, \dots, q_j, \dots, q_N) \quad (2.291)$$

and are said to be *antisymmetric* under the interchange P_{ij} .

More generally, there are $N!$ different permutations of the variables $q_1, \dots, q_N$. Defining P as the permutation that replaces q_n by q_{p_n} , with $n = 1, 2, \dots, N$ and noting that P can be obtained as a succession of interchanges, we have

$$[P, H] = 0 \quad (2.292)$$

A permutation P is said to be *even* or *odd* depending on whether the number of interchanges leading to it is even or odd. If we let the operator P act on a wave function $\psi(q_1, \dots, q_N)$, we have

$$P\psi(q_1, \dots, q_N) = \psi(q_{P1}, \dots, q_{PN}) \quad (2.293)$$

It is worth stressing that except for the case $N = 2$ the $N!$ permutations P do not commute among themselves. This is due to the fact that the interchange operators P_{ij} and P_{ik} ($k \neq j$) do not mutually commute. Therefore, the eigenfunctions $\psi(q_1, \dots, q_N)$ of H are not in general eigenfunctions of all the $N!$ permutation operators P . However, there are *two* exceptional states which are eigenstates of H and of the $N!$ permutation operators P . The first one is the *totally symmetric state* $\psi_S(q_1, \dots, q_N)$ satisfying (2.290) for any interchange P_{ij} , so that for all P

$$\begin{aligned} P\psi_S(q_1, \dots, q_N) &= \psi_S(q_{P1}, \dots, q_{PN}) \\ &= \psi_S(q_1, \dots, q_N) \end{aligned} \quad (2.294)$$

The other one is the *totally antisymmetric state* $\psi_A(q_1, \dots, q_N)$, which satisfies (2.291) for any interchange P_{ij} . Hence, for all P

$$P\psi_A(q_1, \dots, q_N) = \begin{cases} \psi_A(q_1, \dots, q_N) & \text{for an even permutation} \\ -\psi_A(q_1, \dots, q_N) & \text{for an odd permutation} \end{cases} \quad (2.295)$$

We also remark that the equation (2.292) implies that P is a constant of the motion, so that a system of identical particles represented by a wave function of a given symmetry (S or A) will keep that symmetry at all times.

Bosons and fermions

According to our present knowledge of particles occurring in nature, the two types of states ψ_S and ψ_A are thought to be sufficient to describe all systems of identical particles. This is called the *symmetrisation postulate*. Particles having states described by totally symmetric wave functions are called *bosons*; they obey Bose-Einstein statistics [7]. Experiment shows that particles of *zero* or *integer* spin ($s = 0, 1, 2, \dots$) are bosons. For example, all the *mesons* (such as the π and K mesons, which have spin $s = 0$, the ρ meson, which has spin $s = 1$, etc.) are bosons. The photon (which has spin $s = 1$), is also a boson, as are the newly discovered W and Z particles, which also have spin $s = 1$, and are called 'intermediate vector bosons'. On the other hand, particles having states described by totally antisymmetric wave functions are called *fermions*; they satisfy Fermi-Dirac statistics [7]. Experiment shows that particles having half-odd integer spin ($s = 1/2, 3/2, \dots$) are fermions. For example, all the *leptons* (such as the electron, the muon and the

[7] See Bransden and Joachain (2000).

neutrinos, which have spin $s = 1/2$) are fermions, as are the *baryons* (such as the proton and the neutron, which have spin $s = 1/2$, and the hyperons). A composite particle, such as an atom, is a boson if the sum of its protons, neutrons and electrons is an even number; it is a fermion if this sum is an odd number.

If a system is composed of different kinds of bosons (fermions), then its wave function must be separately totally symmetric (antisymmetric) with respect to permutations of each kind of identical particle. For example, the wave function of an alpha particle (that is, a ${}^4\text{He}$ nucleus which contains two protons and two neutrons) must be antisymmetric under the interchange of the two protons and also antisymmetric under the interchange of the two neutrons. It should be noted, however, that in some physical situations one can ignore the possible composite structure of particles. For example, in studying the scattering of two alpha particles at low energies (where the effect of nuclear forces is negligible), the symmetrisation process must be accomplished only with respect to the permutation of the two alpha particles, each of them being considered as a spin zero boson [8]. Similarly, in studying the properties of atomic gases at low temperature, one can treat the atoms as 'elementary' bosons or fermions.

2.8

Approximation methods

As in the case of classical mechanics, there are relatively few physically interesting problems in quantum mechanics which can be solved exactly. Approximation methods are therefore of great importance in discussing the application of quantum theory to specific systems, such as the atomic and molecular ones considered in this book. In this section we shall review several approximation methods which will be used in further chapters.

Time-independent perturbation theory

Perturbation theory deals with the changes induced in a system by a 'small' disturbance. Although we shall also apply perturbation methods to scattering problems at a later stage (see Chapters 12–14) we shall start here by discussing the *Rayleigh–Schrödinger perturbation theory*, which is concerned with the modifications in the *discrete* energy levels and corresponding eigenfunctions of a system when a perturbation is applied.

Let us assume that the time-independent Hamiltonian H of a system may be separated into two parts,

$$H = H_0 + \lambda H' \quad (2.296)$$

[8] On the other hand, if one wants to study the scattering of two alpha particles at high energies (where nuclear forces are important), the symmetrisation of the total wave function must be performed separately with respect to the constituent protons and neutrons.

where the 'unperturbed' Hamiltonian H_0 is sufficiently simple so that the corresponding Schrödinger eigenvalue equation

$$H_0 \psi_k^{(0)} = E_k^{(0)} \psi_k^{(0)} \quad (2.297)$$

may be solved, and the term $\lambda H'$ is a small perturbation. The parameter λ will be used below to distinguish between the various orders of the perturbation calculation; we shall take $\lambda = 1$ for the actual physical problem. We assume that the (known) eigenfunctions $\psi_k^{(0)}$ corresponding to the (known) eigenvalues $E_k^{(0)}$ of H_0 form a complete orthonormal set (which may be partly continuous). Thus, if $\psi_i^{(0)}$ and $\psi_j^{(0)}$ are two members of that set, we have

$$\langle \psi_i^{(0)} | \psi_j^{(0)} \rangle = \delta_{ij} \quad \text{or} \quad \delta(i-j) \quad (2.298)$$

where the symbol $\delta(i-j)$ should be used when both $\psi_i^{(0)}$ and $\psi_j^{(0)}$ correspond to continuous states. In what follows we shall simplify the notation by extending the meaning of δ_{ij} to cover both possibilities in (2.298). The eigenvalue problem which we want to solve is

$$H \psi_k = E_k \psi_k \quad (2.299)$$

where we have used the notation E_k and ψ_k to denote the perturbed energy levels and eigenfunctions, respectively.

Non-degenerate case

Let us focus our attention on a particular unperturbed, discrete energy level $E_k^{(0)}$, which we assume to be *non-degenerate*. We suppose that the effect of the perturbation $\lambda H'$ is small enough so that the perturbed energy level E_k is much closer to $E_k^{(0)}$ than to any other unperturbed level. It is then reasonable to expand both ψ_k and E_k in powers of λ , namely

$$\psi_k = \sum_{n=0}^{\infty} \lambda^n \psi_k^{(n)} \quad (2.300)$$

and

$$E_k = \sum_{n=0}^{\infty} \lambda^n E_k^{(n)} \quad (2.301)$$

where the index n refers to the order of the perturbation. Substituting the expansions (2.300) and (2.301) in (2.299), and using (2.296), we have

$$\begin{aligned} & (H_0 + \lambda H') (\psi_k^{(0)} + \lambda \psi_k^{(1)} + \lambda^2 \psi_k^{(2)} + \dots) \\ &= (E_k^{(0)} + \lambda E_k^{(1)} + \lambda^2 E_k^{(2)} + \dots) (\psi_k^{(0)} + \lambda \psi_k^{(1)} + \lambda^2 \psi_k^{(2)} + \dots) \end{aligned} \quad (2.302)$$

Let us now equate the coefficients of equal powers of λ . Beginning with λ^0 , we see that

$$H_0 \psi_k^{(0)} = E_k^{(0)} \psi_k^{(0)} \quad (2.303)$$

as expected. The coefficient of λ then gives

$$H_0\psi_k^{(1)} + H'\psi_k^{(0)} = E_k^{(0)}\psi_k^{(1)} + E_k^{(1)}\psi_k^{(0)} \quad (2.304)$$

while that of λ^2 yields

$$H_0\psi_k^{(2)} + H'\psi_k^{(1)} = E_k^{(0)}\psi_k^{(2)} + E_k^{(1)}\psi_k^{(1)} + E_k^{(2)}\psi_k^{(0)} \quad (2.305)$$

and so on.

In order to obtain the first energy correction $E_k^{(1)}$, we premultiply (2.304) by $\psi_k^{(0)*}$ and integrate over all space. This gives

$$\langle \psi_k^{(0)} | H_0 - E_k^{(0)} | \psi_k^{(1)} \rangle + \langle \psi_k^{(0)} | H' - E_k^{(1)} | \psi_k^{(0)} \rangle = 0 \quad (2.306)$$

Now, using (2.297) and the fact that the operator H_0 is Hermitian, we have

$$\begin{aligned} \langle \psi_k^{(0)} | H_0 | \psi_k^{(1)} \rangle &= \langle H_0 \psi_k^{(0)} | \psi_k^{(1)} \rangle \\ &= E_k^{(0)} \langle \psi_k^{(0)} | \psi_k^{(1)} \rangle \end{aligned} \quad (2.307)$$

so that the first term on the left of (2.306) vanishes. Moreover, since $\langle \psi_k^{(0)} | \psi_k^{(0)} \rangle = 1$, we see that (2.306) reduces to

$$E_k^{(1)} = \langle \psi_k^{(0)} | H' | \psi_k^{(0)} \rangle \equiv H'_{kk} \quad (2.308)$$

This important result tells us that the first-order correction to the energy for a non-degenerate level is just the expectation value of the perturbation H' with respect to the corresponding unperturbed state of the system.

Proceeding in a similar way with equation (2.305), we have

$$\langle \psi_k^{(0)} | H_0 - E_k^{(0)} | \psi_k^{(2)} \rangle + \langle \psi_k^{(0)} | H' - E_k^{(1)} | \psi_k^{(1)} \rangle - E_k^{(2)} \langle \psi_k^{(0)} | \psi_k^{(0)} \rangle = 0 \quad (2.309)$$

and therefore

$$E_k^{(2)} = \langle \psi_k^{(0)} | H' - E_k^{(1)} | \psi_k^{(1)} \rangle \quad (2.310)$$

An equivalent expression for $E_k^{(2)}$ may be obtained by starting from (2.304), and is given (Problem 2.19) by

$$E_k^{(2)} = -\langle \psi_k^{(1)} | H_0 - E_k^{(0)} | \psi_k^{(1)} \rangle \quad (2.311)$$

Expressions for higher order corrections $E_k^{(n)}$, $n \geq 3$, can be obtained in a similar way. For example, one has (Problem 2.19)

$$E_k^{(3)} = \langle \psi_k^{(1)} | H' - E_k^{(1)} | \psi_k^{(2)} \rangle - 2E_k^{(2)} \langle \psi_k^{(0)} | \psi_k^{(1)} \rangle \quad (2.312)$$

Let us now return to (2.304). The Rayleigh-Schrödinger method attempts to obtain the solution $\psi_k^{(1)}$ of that equation in the following way. First, the 'unperturbed' equation (2.297) is solved for all eigenvalues and eigenfunctions (including those belonging to the continuous part of the spectrum, if one exists). The unknown function $\psi_k^{(1)}$ is then expanded in the basis set of the unperturbed eigenfunctions. That is,

$$\psi_k^{(1)} = \sum_m a_m^{(1)} \psi_m^{(0)} \quad (2.313)$$

where the sum over m means a summation over the discrete part of the set and an integration over its continuous part. Substituting (2.313) into (2.304), we obtain

$$(H_0 - E_k^{(0)}) \sum_m a_m^{(1)} \psi_m^{(0)} + (H' - E_k^{(1)}) \psi_k^{(0)} = 0 \quad (2.314)$$

Premultiplying by $\psi_l^{(0)*}$, integrating over all space and using the fact that $H_0\psi_l^{(0)} = E_l^{(0)}\psi_l^{(0)}$ and $\langle \psi_l^{(0)} | \psi_k^{(0)} \rangle = \delta_{kl}$, we find that

$$a_l^{(1)}(E_l^{(0)} - E_k^{(0)}) + \langle \psi_l^{(0)} | H' | \psi_k^{(0)} \rangle - E_k^{(1)} \delta_{kl} = 0 \quad (2.315)$$

For $l = k$ this reduces to our basic result (2.308). On the other hand, for $l \neq k$ we have

$$a_l^{(1)} = \frac{H'_{lk}}{E_k^{(0)} - E_l^{(0)}}, \quad l \neq k \quad (2.316)$$

where we have set $H'_{lk} \equiv \langle \psi_l^{(0)} | H' | \psi_k^{(0)} \rangle$. We note that the equation (2.304) does not determine the coefficient $a_k^{(1)}$, which is the 'component' of $\psi_k^{(1)}$ along $\psi_k^{(0)}$. We can thus require without loss of generality that

$$a_k^{(1)} = \langle \psi_k^{(0)} | \psi_k^{(1)} \rangle = 0 \quad (2.317)$$

and rewrite (2.313) as

$$\begin{aligned} \psi_k^{(1)} &= \sum_{m \neq k} a_m^{(1)} \psi_m^{(0)} \\ &= \sum_{m \neq k} \frac{H'_{mk}}{E_k^{(0)} - E_m^{(0)}} \psi_m^{(0)} \end{aligned} \quad (2.318)$$

Substituting this result in (2.310) we obtain

$$E_k^{(2)} = \sum_{m \neq k} \frac{H'_{km} H'_{mk}}{E_k^{(0)} - E_m^{(0)}} = \sum_{m \neq k} \frac{|H'_{mk}|^2}{E_k^{(0)} - E_m^{(0)}} \quad (2.319)$$

The third-order correction $E_k^{(3)}$ may be obtained in a similar way from (2.312) and (2.318).

Degenerate case

Thus far we have assumed that the perturbed eigenfunction ψ_k differs slightly from a given function $\psi_k^{(0)}$, the solution of the 'unperturbed' equation (2.297). When the level $E_k^{(0)}$ is α -fold degenerate, there are several 'unperturbed' wave functions $\psi_{kr}^{(0)}$ ($r = 1, 2, \dots, \alpha$) corresponding to this level and we do not know *a priori* to which functions the perturbed eigenfunctions tend when $\lambda \rightarrow 0$. This means that the above treatment – and in particular the basic expansion (2.300) – must be modified to deal with the degenerate case.

The α unperturbed wave functions $\psi_{kr}^{(0)}$ corresponding to the level $E_k^{(0)}$ are orthogonal to the unperturbed wave functions $\psi_l^{(0)}$ corresponding to *other* energy levels $E_l^{(0)} \neq E_k^{(0)}$. Although they need not be orthogonal among themselves, it is always possible to construct from linear combinations of them a new set of α unperturbed wave functions which are mutually orthogonal and normalised to unity. We may therefore assume without loss of generality that this has already been done, so that

$$\langle \psi_{kr}^{(0)} | \psi_{ks}^{(0)} \rangle = \delta_{rs} \quad (r, s = 1, 2, \dots, \alpha) \quad (2.320)$$

Let us now introduce the correct zero-order functions $\chi_{kr}^{(0)}$ which yield the first term in the expansion of the exact wave function ψ_{kr} in powers of λ . That is,

$$\psi_{kr} = \chi_{kr}^{(0)} + \lambda \psi_{kr}^{(1)} + \lambda^2 \psi_{kr}^{(2)} + \dots \quad (2.321)$$

We shall also write the perturbed energy E_{kr} as

$$E_{kr} = E_k^{(0)} + \lambda E_{kr}^{(1)} + \lambda^2 E_{kr}^{(2)} + \dots \quad (2.322)$$

with $E_k^{(0)} \equiv E_{kr}^{(0)}$ ($r = 1, 2, \dots, \alpha$) since the level $E_k^{(0)}$ is α -fold degenerate. Using the above expansions in (2.299) and equating the coefficients of λ we find that

$$H_0 \psi_{kr}^{(1)} + H' \chi_{kr}^{(0)} = E_k^{(0)} \psi_{kr}^{(1)} + E_{kr}^{(1)} \chi_{kr}^{(0)} \quad (2.323)$$

Since the functions $\chi_{kr}^{(0)}$ are linear combinations of the unperturbed wave functions $\psi_{kr}^{(0)}$, we may write

$$\chi_{kr}^{(0)} = \sum_{s=1}^{\alpha} c_{rs} \psi_{ks}^{(0)} \quad (r = 1, 2, \dots, \alpha) \quad (2.324)$$

where the coefficients c_{rs} are to be determined. Similarly, expanding $\psi_{kr}^{(1)}$ in the basis set of the unperturbed wave functions, we have

$$\psi_{kr}^{(1)} = \sum_m \sum_s a_{krms}^{(1)} \psi_{ms}^{(0)} \quad (2.325)$$

where the indices r and s refer explicitly to the degeneracy. Substituting the above expressions of $\chi_{kr}^{(0)}$ and $\psi_{kr}^{(1)}$ in (2.323) and using the fact that $H_0 \psi_{ms}^{(0)} = E_m^{(0)} \psi_{ms}^{(0)}$, we find that

$$\sum_m \sum_s a_{krms}^{(1)} (E_m^{(0)} - E_k^{(0)}) \psi_{ms}^{(0)} + \sum_s c_{rs} (H' - E_{kr}^{(1)}) \psi_{ks}^{(0)} = 0 \quad (2.326)$$

Premultiplying by $\psi_{ku}^{(0)*}$ and integrating over all space, we obtain

$$\sum_m \sum_s a_{krms}^{(1)} (E_m^{(0)} - E_k^{(0)}) \langle \psi_{ku}^{(0)} | \psi_{ms}^{(0)} \rangle + \sum_s c_{rs} [\langle \psi_{ku}^{(0)} | H' | \psi_{ks}^{(0)} \rangle - E_{kr}^{(1)} \delta_{us}] = 0 \quad (2.327)$$

($u = 1, 2, \dots, \alpha$)

where we have used (2.320). Since $\langle \psi_{ku}^{(0)} | \psi_{ms}^{(0)} \rangle = 0$ when $k \neq m$ and $E_k^{(0)} = E_m^{(0)}$ if $k = m$, we see that (2.327) reduces to

$$\sum_{s=1}^{\alpha} c_{rs} [\langle \psi_{ku}^{(0)} | H' | \psi_{ks}^{(0)} \rangle - E_{kr}^{(1)} \delta_{us}] = 0 \quad (u = 1, 2, \dots, \alpha) \quad (2.328)$$

This is a linear, homogeneous system of equations for the α unknown quantities $c_{r1}, c_{r2}, \dots, c_{r\alpha}$. A non-trivial solution is obtained if the determinant of the quantity in square brackets vanishes,

$$\det |\langle \psi_{ku}^{(0)} | H' | \psi_{ks}^{(0)} \rangle - E_{kr}^{(1)} \delta_{us}| = 0 \quad (s, u = 1, 2, \dots, \alpha) \quad (2.329)$$

This equation yields α real roots $E_{kr}^{(1)}, E_{kr}^{(1)}, \dots, E_{kr}^{(1)}$. If all of these roots are *distinct* the degeneracy is *completely removed* to first order in the perturbation. On the other hand, if some or all roots of (2.329) are identical the degeneracy is only *partially* (or *not at all*) *removed*. The residual degeneracy may then either be removed in higher order of perturbation theory, or persist to all orders. The latter case occurs when the operators H_0 and H share symmetry properties.

For a given value of r , the coefficients c_{rs} ($s = 1, 2, \dots, \alpha$) which determine the 'correct' unperturbed zero-order wave functions $\chi_{kr}^{(0)}$ via (2.324) may be obtained by substituting the value of $E_{kr}^{(1)}$ in the system (2.328) and solving for the coefficients $c_{r1}, c_{r2}, \dots, c_{r\alpha}$ in terms of one of them. The last coefficient is then obtained (up to a phase factor) by requiring the function $\chi_{kr}^{(0)}$ to be normalised to unity. It is clear that this procedure does not lead to a unique result when two or more roots $E_{kr}^{(1)}$ of equation (2.329) coincide, since in this case the degeneracy is not fully removed.

Time-dependent perturbation theory

We shall now discuss the perturbation theory for a system whose total Hamiltonian H may be split as

$$H = H_0 + \lambda H'(t) \quad (2.330)$$

where the unperturbed Hamiltonian H_0 is time-independent and $\lambda H'(t)$ is a small *time-dependent* perturbation. The method which we outline below is known as *Dirac's method of variation of constants*.

Let us suppose that we know the eigenvalues $E_k^{(0)}$ of the unperturbed Hamiltonian H_0 , together with the corresponding stationary eigenfunctions $\psi_k^{(0)}$, which we assume to be orthonormal and to form a complete set. Thus, since $H_0 \psi_k^{(0)} = E_k^{(0)} \psi_k^{(0)}$, the general solution of the time-dependent Schrödinger equation

$$i\hbar \frac{\partial \Psi_0}{\partial t} = H_0 \Psi_0 \quad (2.331)$$

is given by

$$\Psi_0 = \sum_k c_k^{(0)} \psi_k^{(0)} \exp(-iE_k^{(0)}t/\hbar) \quad (2.332)$$

where the coefficients $c_k^{(0)}$ are constants and the sum is over the entire set of eigenfunctions $\psi_k^{(0)}$. Because the functions $\psi_k^{(0)}$ form a complete set, the general solution Ψ of the time-dependent Schrödinger equation

$$i\hbar \frac{\partial \Psi}{\partial t} = H\Psi \quad (2.333)$$

can be expanded as

$$\Psi = \sum_k c_k(t) \psi_k^{(0)} \exp(-iE_k^{(0)}t/\hbar) \quad (2.334)$$

where the unknown coefficients $c_k(t)$ clearly depend on the time. Since the wave functions $\psi_k^{(0)}$ are orthonormal, and provided Ψ is normalised to unity, we can interpret the quantity $|c_k(t)|^2$ as the probability of finding the system in the state labelled k at the time t , and $c_k(t)$ as the corresponding probability amplitude. Upon comparison of (2.332) and (2.334) we see that if $H'(t) = 0$ the coefficients $c_k(t)$ reduce to the constants $c_k^{(0)}$ which are therefore the initial values of the c_k . Thus, as we expect from (2.332) the quantity $|c_k^{(0)}|^2$ gives the probability of finding the system in the stationary state $\psi_k^{(0)}$ before the perturbation is applied.

To find equations for the coefficients $c_k(t)$ the expansion (2.334) is inserted into the Schrödinger equation (2.333). From (2.330) and the fact that $H_0 \psi_k^{(0)} = E_k^{(0)} \psi_k^{(0)}$, we then have

$$i\hbar \sum_k \dot{c}_k(t) \psi_k^{(0)} \exp(-iE_k^{(0)}t/\hbar) = \sum_k c_k(t) \lambda H'(t) \psi_k^{(0)} \exp(-iE_k^{(0)}t/\hbar) \quad (2.335)$$

where the dot indicates a derivative with respect to the time. Taking the scalar product with a particular function $\psi_b^{(0)}$ belonging to the set $\{\psi_k^{(0)}\}$ and using the fact that $\langle \psi_b^{(0)} | \psi_k^{(0)} \rangle = \delta_{bk}$, we then find from (2.335) the set of coupled equations

$$\dot{c}_b(t) = (i\hbar)^{-1} \sum_k \lambda H'_{bk}(t) c_k(t) \exp(i\omega_{bk}t) \quad (2.336)$$

where

$$H'_{bk}(t) = \langle \psi_b^{(0)} | H'(t) | \psi_k^{(0)} \rangle \quad (2.337)$$

and where the Bohr angular frequency ω_{bk} is defined by

$$\omega_{bk} = \frac{E_b^{(0)} - E_k^{(0)}}{\hbar} \quad (2.338)$$

The system of coupled differential equations (2.336) is completely equivalent to the original time-dependent Schrödinger equation (2.333), and no approximation has been made thus far. However, if the perturbation $\lambda H'$ is weak, we can expand the coefficients c_k in powers of the parameter λ as

$$c_k = c_k^{(0)} + \lambda c_k^{(1)} + \lambda^2 c_k^{(2)} + \dots \quad (2.339)$$

Substituting this expansion into the system (2.336) and equating the coefficients of equal powers of λ , we find that

$$\dot{c}_b^{(0)} = 0 \quad (2.340a)$$

$$\dot{c}_b^{(1)} = (i\hbar)^{-1} \sum_k H'_{bk}(t) \exp(i\omega_{bk}t) c_k^{(0)} \quad (2.340b)$$

$\vdots$

$$\dot{c}_b^{(s+1)} = (i\hbar)^{-1} \sum_k H'_{bk}(t) \exp(i\omega_{bk}t) c_k^{(s)}, \quad s = 0, 1, \dots \quad (2.340c)$$

Thus the original system (2.336) has been decoupled in such a way that the equations (2.340) can now in principle be integrated successively to any given order.

The first equation (2.340a) simply confirms that the coefficients $c_k^{(0)}$ are time-independent. As we have seen above, the constants $c_k^{(0)}$ define the initial conditions of the problem. In what follows, we shall assume for the sake of simplicity that the system is initially (that is, for $t \leq t_0$) in a well-defined stationary state $\psi_a^{(0)}$ of energy $E_a^{(0)}$. Thus

$$c_k^{(0)} = \begin{cases} \delta_{ka} & \text{for discrete states} \\ \delta(k-a) & \text{for continuous states.} \end{cases} \quad (2.341)$$

We note that this statement is not in contradiction with the uncertainty relation $\Delta E \Delta t \gtrsim \hbar$ since we have essentially an 'infinite' amount of time available to prepare our initial state. Upon substitution of (2.341) into (2.340b) we then have

$$\dot{c}_b^{(1)}(t) = (i\hbar)^{-1} H'_{ba}(t) \exp(i\omega_{ba}t) \quad (2.342)$$

where $\omega_{ba} = (E_b^{(0)} - E_a^{(0)})/\hbar$. The solution of these first-order equations is

$$c_a^{(1)}(t) = (i\hbar)^{-1} \int_{t_0}^t H'_{aa}(t') dt' \quad (2.343a)$$

for $b = a$, while for $b \neq a$ one has

$$c_b^{(1)}(t) = (i\hbar)^{-1} \int_{t_0}^t H'_{ba}(t') \exp(i\omega_{ba}t') dt', \quad b \neq a \quad (2.343b)$$

where the integration constant has been chosen in such a way that $c_a^{(1)}(t)$ and $c_b^{(1)}(t)$ vanish at $t = t_0$, before the perturbation is applied. To first order in the perturbation, the transition probability corresponding to the transition $a \rightarrow b$ is therefore given by

$$P_{ba}^{(1)}(t) = |c_b^{(1)}(t)|^2 = \hbar^{-2} \left| \int_{t_0}^t H'_{ba}(t') \exp(i\omega_{ba}t') dt' \right|^2, \quad b \neq a \quad (2.344)$$

It is also worth noting that for $t > t_0$ the coefficient c_a of the state a is given to first order in the perturbation by

$$\begin{aligned}
c_a(t) &\approx c_a^{(0)} + c_a^{(1)}(t) \\
&\approx 1 + (i\hbar)^{-1} \int_{t_0}^t H'_{aa}(t') dt' \\
&\approx \exp \left[-\frac{i}{\hbar} \int_{t_0}^t H'_{aa}(t') dt' \right]
\end{aligned} \tag{2.345}$$

so that $|c_a(t)|^2 \approx 1$ and the main effect of the perturbation on the initial state is to change its phase.

Time-independent perturbation

The results (2.343) take a particularly simple form if the perturbation H' is independent of time, except for being switched on suddenly at a given time (say $t_0 = 0$). We then have

$$c_a^{(1)}(t) = (i\hbar)^{-1} H'_{aa} t \tag{2.346a}$$

and

$$c_b^{(1)}(t) = \frac{H'_{ba}}{\hbar\omega_{ba}} [1 - \exp(i\omega_{ba}t)], \quad b \neq a \tag{2.346b}$$

Note that if the perturbation is switched off at time t , the above amplitudes are also those at any subsequent time $t_1 > t$. The first-order transition probability from state a to state $b \neq a$ is given by

$$P_{ba}^{(1)}(t) = |c_b^{(1)}(t)|^2 = \frac{2}{\hbar^2} |H'_{ba}|^2 F(t, \omega_{ba}) \tag{2.347}$$

where

$$F(t, \omega) = \frac{1 - \cos \omega t}{\omega^2} = \frac{2 \sin^2(\omega t/2)}{\omega^2} \tag{2.348}$$

The function $F(t, \omega)$ is shown in Fig. 2.9 for fixed t . We see that it exhibits a sharp peak about the value $\omega = 0$. The height of this peak is proportional to t^2 , while its width is approximately $2\pi/t$. Setting $x = \omega t/2$, we also note that

$$\int_{-\infty}^{+\infty} F(t, \omega) d\omega = t \int_{-\infty}^{+\infty} \frac{\sin^2 x}{x^2} dx = \pi t \tag{2.349}$$

where we have used a standard integral. Using (2.31c) we also deduce that in the limit $t \rightarrow \infty$

$$F(t, \omega) \underset{t \rightarrow \infty}{\sim} \pi t \delta(\omega) \tag{2.350}$$

Let us first analyse (2.347) for a fixed value of t . Since the function $F(t, \omega_{ba})$ has a sharp peak of width $2\pi/t$ about the value $\omega_{ba} = 0$, it is clear from (2.347) that

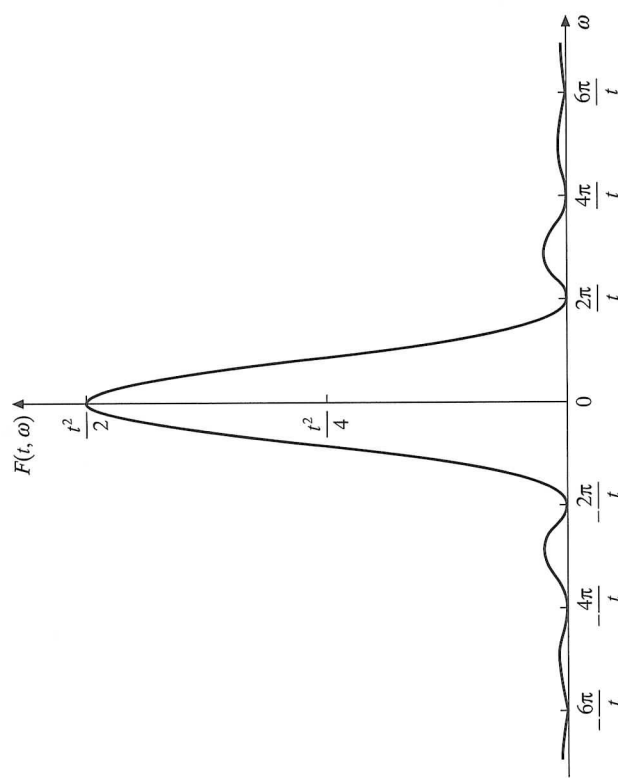


Figure 2.9 The function $F(t, \omega)$ of equation (2.348), for fixed t .

transitions to those final states b for which ω_{ba} does not deviate from zero by more than $\delta\omega_{ba} \approx 2\pi/t$ will be strongly favoured. Therefore the transitions $a \rightarrow b$ will occur mainly towards those final states whose energy $E_b^{(0)}$ is located in a band of width

$$\delta E \approx 2\pi\hbar/t \tag{2.351}$$

about the initial energy $E_a^{(0)}$, so that the unperturbed energy is conserved to within $2\pi\hbar/t$. This result may easily be related to the time-energy uncertainty relation $\Delta E \Delta t \approx \hbar$. Indeed, since the perturbation gives a way of measuring the energy of the system by inducing transitions $a \rightarrow b$, and because this perturbation acts during a time t , the uncertainty related to this energy measurement should be approximately $\hbar/t$, in qualitative agreement with (2.351).

We now study the transition probability (2.347) as a function of t . For a transition to a *given state* b we must distinguish two cases:

1. If the transition is such that $\omega_{ba} = 0$ (so that $E_b^{(0)} = E_a^{(0)}$ and the states a and b are degenerate), then we see from (2.347) that

$$P_{ba}^{(1)}(t) = \frac{|H'_{ba}|^2}{\hbar^2} t^2 \tag{2.352}$$

so that the transition probability increases as t^2 . Thus, after a sufficient length of time the quantity $P_{ba}^{(1)}(t)$ will no longer satisfy the inequality

$$P_{ba}^{(1)}(t) \ll 1 \quad (2.353)$$

required by a perturbative approach. Hence, the present perturbation method cannot be applied to degenerate systems which are perturbed over long periods of time.

2. If on the contrary $\omega_{ba} \neq 0$ (so that the state b is not degenerate with the initial state a), we have

$$P_{ba}^{(1)}(t) = \frac{4|H'_{ba}|^2}{\hbar^2 \omega_{ba}^2} \sin^2(\omega_{ba}t/2) \quad (2.354)$$

and we note that $P_{ba}^{(1)}(t)$ oscillates with a period $2\pi/|\omega_{ba}|$ about the average value $2|H'_{ba}|^2/(\hbar^2 \omega_{ba}^2)$. Since

$$P_{ba}^{(1)}(t) \leq \frac{4|H'_{ba}|^2}{\hbar^2 \omega_{ba}^2} \quad (2.355)$$

the condition (2.353) can always be satisfied if the perturbation H' is sufficiently weak. It is also worth noting that for times t small with respect to the period of oscillation $2\pi/|\omega_{ba}|$ one has $\sin(\omega_{ba}t/2) \approx \omega_{ba}t/2$ so that $P_{ba}^{(1)}(t) \approx |H'_{ba}|^2 t^2 / \hbar^2$ increases quadratically with time.

Instead of considering transitions to a particular state b , it is often necessary to deal with transitions involving a group of states n whose energy E_n lies within a given interval $(E_b^{(0)} - \eta, E_b^{(0)} + \eta)$ centred about the value $E_b^{(0)}$. Let us denote by $\rho_n(E_n)$ the density of levels on the energy scale, so that $\rho_n(E_n) dE_n$ is the number of final states n in an interval dE_n containing the energy E_n . We shall continue to assume that the perturbation H' is constant in time, except that it is switched on at $t = 0$. The first-order transition probability $P_{ba}^{(1)}(t)$ from the initial state a to the group of final states n is then given by

$$P_{ba}^{(1)}(t) = \frac{2}{\hbar^2} \int_{E_b^{(0)}-\eta}^{E_b^{(0)}+\eta} |H'_{na}|^2 F(t, \omega_{na}) \rho_n(E_n) dE_n \quad (2.356)$$

where $\omega_{na} = (E_n - E_a^{(0)})/\hbar$. Assuming that η is small enough so that H'_{na} and $\rho_n(E_n)$ are nearly constant within the integration range, we have

$$P_{ba}^{(1)}(t) = \frac{2}{\hbar^2} |H'_{ba}|^2 \rho_b(E_b^{(0)}) \int_{E_b^{(0)}-\eta}^{E_b^{(0)}+\eta} F(t, \omega_{na}) dE_n \quad (2.357)$$

We shall also assume that t is large enough so that the quantity η satisfies the condition

$$\eta \gg 2\pi\hbar/t. \quad (2.358)$$

It is clear from the examination of the function $F(t, \omega)$ (see (2.348) and Fig. 2.9) that the overwhelming part of the integral on the right of (2.357) arises from

transitions which conserve the energy (within $\delta E = 2\pi\hbar/t$). Since $\eta \gg 2\pi\hbar/t$, we may write

$$\int_{E_b^{(0)}-\eta}^{E_b^{(0)}+\eta} F(t, \omega_{na}) dE_n \approx \hbar \int_{-\infty}^{+\infty} F(t, \omega_{na}) d\omega_{na} = \pi\hbar t \quad (2.359)$$

where we have used (2.349). Thus, (2.357) reduces to

$$P_{ba}^{(1)}(t) = \frac{2\pi}{\hbar} |H'_{ba}|^2 \rho_b(E) t \quad (2.360)$$

with $E = E_a^{(0)} = E_b^{(0)}$. Thus the transition probability increases linearly with time for energy-conserving transitions to a group of states. It is worth stressing that the result (2.360) is soundly based only if the condition (2.353) required by a perturbative approach is obeyed.

Introducing the *transition probability per unit time or transition rate*

$$W_{ba} = \frac{dP_{ba}}{dt} \quad (2.361)$$

we see from (2.360) that to first order in perturbation theory, we have

$$W_{ba} = \frac{2\pi}{\hbar} |H'_{ba}|^2 \rho_b(E) \quad (2.362)$$

This formula, first obtained by P.A.M. Dirac, was later called by E. Fermi 'The Golden Rule' of perturbation theory. Although we have derived it here for a perturbation H' which is constant in time (except for being switched on at $t = 0$), it can be generalised to other perturbations, as we shall now illustrate.

Periodic perturbation

Let us now consider a perturbation $H'(t)$ which is a *periodic* function of time, except for being turned on at $t = 0$. This case is of particular importance for studying the interaction of atoms and molecules with electromagnetic fields, as we shall see in further chapters. We shall assume first that the perturbation is *harmonic* in time, with angular frequency ω . That is,

$$H'(t) = A \exp(i\omega t) + A^\dagger \exp(-i\omega t) \quad (2.363)$$

where A is a time-independent operator and $A^\dagger$ is its adjoint so that the operator H' is Hermitian. It is supposed that the system is initially (for $t \leq 0$) in the unperturbed bound state $\psi_a^{(0)}$, of energy $E_a^{(0)}$, so that the initial conditions are $c_a(t \leq 0) = 1$ and $c_b(t \leq 0) = 0$ for $b \neq a$.

According to (2.345), we have $|c_b(t)|^2 \approx 1$ for $t > 0$. In order to find the coefficient $c_b^{(1)}(t)$ for $t > 0$ and $b \neq a$, we substitute (2.363) into (2.343b) and use the fact that $t_0 = 0$. This gives

$$c_b^{(1)}(t) = (i\hbar)^{-1} \left\{ A_{ba} \int_0^t \exp[i(\omega_{ba} + \omega)t'] dt' + A_{ba}^* \int_0^t \exp[i(\omega_{ba} - \omega)t'] dt' \right\} \quad (2.364)$$

where $A_{ba} = \langle \psi_b^{(0)} | A | \psi_a^{(0)} \rangle$ and $A_{ba}^* = A_{ab}^*$. Performing the integrals, and remembering that $\hbar\omega_{ba} = E_b^{(0)} - E_a^{(0)}$, we find that the corresponding first-order transition probability is given by

$$P_{ba}^{(1)}(t) = \left| A_{ba} \frac{1 - \exp[i(E_b^{(0)} - E_a^{(0)} + \hbar\omega)t/\hbar]}{E_b^{(0)} - E_a^{(0)} + \hbar\omega} + A_{ba}^* \frac{1 - \exp[i(E_b^{(0)} - E_a^{(0)} - \hbar\omega)t/\hbar]}{E_b^{(0)} - E_a^{(0)} - \hbar\omega} \right|^2 \quad (2.365)$$

It is clear from this equation that if t is large enough the probability of finding the system in the state b will only be appreciable if the denominator of one or the other of the two terms on the right-hand side is close to zero. Moreover, assuming that $E_b^{(0)} \neq E_a^{(0)}$ (so that the levels $E_a^{(0)}$ and $E_b^{(0)}$ are not degenerate), both denominators cannot simultaneously be close to zero. A good approximation is therefore to neglect the interference between the two terms in calculating the transition probability. Thus, if the energy $E_b^{(0)}$ lies in a small band about the value

$$E = E_a^{(0)} + \hbar\omega \quad (2.366)$$

only the second term in (2.365) will have an appreciable magnitude, the corresponding transition probability being given by

$$P_{ba}^{(1)}(t) = \frac{2}{\hbar^2} |A_{ba}^*|^2 F(t, \omega_{ba} - \omega) \quad (2.367)$$

The main difference with the expression (2.347), obtained for a time-independent perturbation, is that the angular frequency ω_{ba} has now been replaced by $\omega_{ba} - \omega$. From the properties of the function $F(t, \omega)$ it is apparent that the transition probability (2.367) will only be significant if $E_b^{(0)}$ is located in an interval of width $2\pi\hbar/t$ about the value $E_a^{(0)} + \hbar\omega$. Hence the first-order transition probability (2.367) will be appreciable if the system has *absorbed* an amount of energy given (to within $2\pi\hbar/t$) by $\hbar\omega = E_b^{(0)} - E_a^{(0)}$. This, of course, is nothing but the Bohr frequency rule (1.70). When it is exactly satisfied, so that $\omega = \omega_{ba}$, *resonance* is said to occur and we see from (2.367) and (2.348) that the first-order transition probability increases quadratically with time according to the formula

$$P_{ba}^{(1)}(t) = \frac{|A_{ba}^*|^2}{\hbar^2} t^2 \quad (2.368)$$

In the same way, if the energy $E_b^{(0)}$ lies in a small interval in the neighbourhood of the value

$$E = E_a^{(0)} - \hbar\omega \quad (2.369)$$

only the first term on the right of (2.365) will be significant. The corresponding first-order transition probability will be given by

$$P_{ba}^{(1)}(t) = \frac{2}{\hbar^2} |A_{ba}|^2 F(t, \omega_{ba} + \omega) \quad (2.370)$$

and will only be significant if the system has *emitted* an amount of energy given (to within $2\pi\hbar/t$) by $\hbar\omega = E_b^{(0)} - E_a^{(0)}$. Again *resonance* occurs if this condition is exactly satisfied, in which case $\omega = -\omega_{ba}$ and the transition probability (2.370) increases quadratically with time. It should be noted that in practice t is large enough ($t \gg 2\pi/\omega$) so that the two bands of width $2\pi\hbar/t$ about the values (2.366) and (2.369) do not overlap. Thus our neglect of the interference between the two terms on the right of (2.365) in calculating $P_{ba}^{(1)}$ is indeed justified.

As in the case of a time-independent perturbation, one can consider transitions to a group of final states n whose energy E_n lies within an interval $(E_b^{(0)} - \eta, E_b^{(0)} + \eta)$ about the value $E_b^{(0)} = E_a^{(0)} + \hbar\omega$ (for absorption) or $E_b^{(0)} = E_a^{(0)} - \hbar\omega$ (for emission) with $\eta \gg 2\pi\hbar/t$. Let $\rho_n(E_n)$ be the density of levels E_n on the energy scale. Under conditions similar to those discussed above for a time-independent perturbation, a transition probability per unit time (transition rate) can be defined. For transitions in which the system *absorbs* an energy $\hbar\omega \approx E_b^{(0)} - E_a^{(0)}$, this transition rate is given to first order by

$$W_{ba} = \frac{2\pi}{\hbar} |A_{ba}^*|^2 \rho_b(E) \quad (2.371)$$

where $E = E_a^{(0)} + \hbar\omega$. The above expression is clearly the direct generalisation of the Golden Rule (2.362). Similarly, for transitions in which the system *emits* an energy $\hbar\omega \approx E_a^{(0)} - E_b^{(0)}$ by making transitions to a group of final states, the corresponding transition rate is given (under the same conditions) by

$$W_{ba} = \frac{2\pi}{\hbar} |A_{ba}|^2 \rho_b(E) \quad (2.372)$$

with $E = E_a^{(0)} - \hbar\omega$.

Until now we have considered a perturbation $H'(t)$ which is a harmonic function of time. However, the generalisation of the above results to a perturbation $H'(t)$ which is a general periodic function of time is straightforward. Indeed, we can then develop $H'(t)$ in the Fourier series

$$H'(t) = \sum_{n=-\infty}^{\infty} [A_n \exp(in\omega t) + A_n^* \exp(-in\omega t)] \quad (2.373)$$

where the operators A_n are time-independent. For large enough times ($t \gg 2\pi/\omega$), there is no interference between the contributions of the various terms of this development to the transition probability, because each term corresponds to a different energy transfer. A term of the type $A_n \exp(in\omega t)$ will therefore lead to the *emission* by the system of an energy given (to within $2\pi\hbar/t$) by $\hbar n\omega = E_a^{(0)} - E_b^{(0)}$

while a term of the type $A_a^\dagger \exp(-i\omega t)$ corresponds to the *absorption* by the system of the energy $\hbar\hbar\omega = E_b^{(0)} - E_a^{(0)}$ (to within $2\pi\hbar/t$). The corresponding first-order transition probabilities are readily obtained, either for transitions to a given state b or to a group of final states.

Two-level system with harmonic perturbation

As a simple application of some of the concepts developed above, let us consider a two-level system, with unperturbed energies $E_a^{(0)} < E_b^{(0)}$ and corresponding eigenfunctions $\psi_a^{(0)}$ and $\psi_b^{(0)}$, respectively. The system being initially in the state a , a perturbation of the form (2.363) is switched on at time $t = 0$. Setting $\lambda = 1$ in (2.336), we obtain the two coupled equations

$$i\hbar\dot{c}_a(t) = \{A_{aa} \exp(i\omega t) + A_{aa}^\dagger \exp(-i\omega t)\}c_a + \{A_{ab} \exp[i(\Delta\omega)t] + A_{ab}^\dagger \exp[-i(\omega + \omega_{ba})t]\}c_b \quad (2.374a)$$

and

$$i\hbar\dot{c}_b(t) = \{A_{ba} \exp[i(\omega_{ba} + \omega)t] + A_{ba}^\dagger \exp[-i(\Delta\omega)t]\}c_a + \{A_{bb} \exp(i\omega t) + A_{bb}^\dagger \exp(-i\omega t)\}c_b \quad (2.374b)$$

where $\omega_{ba} = (E_b^{(0)} - E_a^{(0)})/\hbar$ and we have introduced the 'detuning' angular frequency

$$\Delta\omega = \omega - \omega_{ba} \quad (2.375)$$

The system (2.374) must be solved subject to the initial conditions

$$c_a(t \leq 0) = 1, \quad c_b(t \leq 0) = 0 \quad (2.376)$$

The equations (2.374) cannot be solved exactly, but if it is assumed that $|\Delta\omega| \ll \omega$ (so that the angular frequency ω is always close to its resonant value $\omega = \omega_{ba}$) then the terms in $\exp[\pm i(\Delta\omega)t]$ will be much more important than those in $\exp[\pm i(\omega + \omega_{ba})t]$ and $\exp(\pm i\omega t)$. This is because the latter terms oscillate much more rapidly and on the average make little contribution to $\dot{c}_a$ or $\dot{c}_b$. It is therefore reasonable to neglect the higher frequency terms. This is known as the *rotating wave approximation* because the only terms which are kept are those in which the time dependence of the system and of the perturbation are in phase. In this approximation, the pair of equations (2.374) reduces to

$$i\hbar\dot{c}_a(t) = A_{ab} \exp[i(\Delta\omega)t]c_b \quad (2.377a)$$

and

$$i\hbar\dot{c}_b(t) = A_{ba}^\dagger \exp[-i(\Delta\omega)t]c_a \quad (2.377b)$$

This system, which is much simpler than (2.374), can be solved exactly. The solutions $c_a(t)$ and $c_b(t)$ satisfying the initial conditions (2.376) are

$$c_a(t) = \exp[i(\Delta\omega)t/2] \left[\cos(\omega_R t/2) - i \left(\frac{\Delta\omega}{\omega_R} \right) \sin(\omega_R t/2) \right] \quad (2.378a)$$

and

$$c_b(t) = \frac{2A_{ba}^\dagger}{i\hbar\omega_R} \exp[-i(\Delta\omega)t/2] \sin(\omega_R t/2) \quad (2.378b)$$

where

$$\omega_R = \left[(\Delta\omega)^2 + \frac{4|A_{ba}^\dagger|^2}{\hbar^2} \right]^{1/2} \quad (2.379)$$

is called the Rabi 'flopping frequency'.

The probability of finding the system at time $t > 0$ in the state a is therefore given by

$$|c_a(t)|^2 = \cos^2(\omega_R t/2) + \frac{(\Delta\omega)^2}{\omega_R^2} \sin^2(\omega_R t/2) \quad (2.380)$$

while that of finding it in the state b (that is, the probability that the transition $a \rightarrow b$ will take place) is

$$P_{ba}(t) = |c_b(t)|^2 = \frac{4|A_{ba}^\dagger|^2}{\hbar^2 \omega_R^2} \sin^2(\omega_R t/2) \quad (2.381)$$

As expected, the excitation is a typical *resonance* process, since the probability (2.381) rapidly decreases when the absolute value $|\Delta\omega|$ of the detuning angular frequency increases. It is also readily verified from (2.378)–(2.381) that $|c_a(t)|^2 + |c_b(t)|^2 = 1$, and that the system oscillates between the two levels with a period $T = 2\pi/\omega_R$.

Having obtained 'exact' results for this problem (within the framework of the rotating wave approximation), we can compare them with those arising from first-order perturbation theory. Using (2.380), we see that when $\omega_R t \ll 1$ we have $|c_a(t)|^2 \approx 1$, which is in agreement with the perturbative result following from (2.345). From (2.367) and (2.348) we also note that the first-order transition probability is given by

$$P_{ba}^{(1)}(t) = \frac{2}{\hbar^2} |A_{ba}^\dagger|^2 F(t, \Delta\omega) \\ = \frac{4|A_{ba}^\dagger|^2}{\hbar^2 (\Delta\omega)^2} \sin^2[(\Delta\omega)t/2] \quad (2.382)$$

If $\Delta\omega \neq 0$ this result agrees with (2.381) provided that the perturbation is weak enough so that one can write $\omega_R \approx \Delta\omega$ (see (2.379)). At resonance ($\Delta\omega = 0$), we see that $P_{ba}^{(1)}(t)$ increases quadratically with time according to (2.368). As expected, this result is only in agreement with the 'exact' expression (2.381) for small enough times and small enough perturbations.

The variational method

We shall now discuss an approximation method, known as the *variational method*, which is very useful in obtaining the bound state energies and wave functions of a time-independent Hamiltonian H . We denote by E_n the eigenvalues of this Hamiltonian and by ψ_n the corresponding orthonormal eigenfunctions, and assume that H has at least one discrete eigenvalue. Let ϕ be an arbitrary normalisable function, and let $E[\phi]$ be the functional

$$E[\phi] = \frac{\langle \phi | H | \phi \rangle}{\langle \phi | \phi \rangle} = \frac{\int \phi^* H \phi \, d\tau}{\int \phi^* \phi \, d\tau} \quad (2.383)$$

where the integration is extended over the full range of all the coordinates of the system.

It is clear that if the function ϕ is identical to one of the exact eigenfunctions ψ_n of H , then $E[\phi]$ will be identical to the corresponding exact eigenvalue E_n . Moreover, it will now be shown that any function ϕ for which the functional $E[\phi]$ is stationary is an eigenfunction of the discrete spectrum of H . Thus, if ϕ and ψ_n differ by an arbitrary infinitesimal variation $\delta\phi$,

$$\phi = \psi_n + \delta\phi \quad (2.384)$$

then the corresponding first-order variation of $E[\phi]$ vanishes:

$$\delta E = 0 \quad (2.385)$$

and the eigenfunctions of H are solutions of the variational equation (2.385).

To prove this statement, we note that upon clearing the fractions and varying, we have from (2.383)

$$\begin{aligned} \delta E &= \delta \left(\frac{\int \phi^* H \phi \, d\tau}{\int \phi^* \phi \, d\tau} \right) = \frac{\delta \int \phi^* H \phi \, d\tau}{\int \phi^* \phi \, d\tau} - E \frac{\delta \int \phi^* \phi \, d\tau}{\int \phi^* \phi \, d\tau} \\ &= \int \delta\phi^* H \phi \, d\tau + \int \phi^* H \delta\phi \, d\tau \end{aligned} \quad (2.386)$$

Since $\langle \phi | \phi \rangle$ is assumed to be finite and non-vanishing, we see that the equation (2.385) is equivalent to

$$\int \delta\phi^* (H - E) \phi \, d\tau + \int \phi^* (H - E) \delta\phi \, d\tau = 0 \quad (2.387)$$

Although the variations $\delta\phi$ and $\delta\phi^*$ are not independent, they may in fact be treated as such, so that the individual terms in (2.387) can be set equal to zero. To see how this comes about, we replace the arbitrary variation $\delta\phi$ by $i\delta\phi$ in (2.387) so that we have

$$-i \int \delta\phi^* (H - E) \phi \, d\tau + i \int \phi^* (H - E) \delta\phi \, d\tau = 0 \quad (2.388)$$

Upon combining (2.388) with (2.387) we then obtain the two equations

$$\int \delta\phi^* (H - E) \phi \, d\tau = 0 \quad (2.389a)$$

$$\int \phi^* (H - E) \delta\phi \, d\tau = 0 \quad (2.389b)$$

which is the desired result. Using the fact that H is Hermitian, we see that the two equations (2.389) are equivalent to the Schrödinger equation

$$(H - E[\phi])\phi = 0 \quad (2.390)$$

Thus any function $\phi = \psi_n$ for which the functional (2.383) is stationary is an eigenfunction of H corresponding to the eigenvalue $E_n \equiv E[\psi_n]$. Conversely, if ψ_n is an eigenfunction of H and E_n the corresponding energy, we have $E_n = E[\psi_n]$ and the functional $E[\psi_n]$ is stationary because ψ_n satisfies the equations (2.389). It is worth stressing that if ϕ and ψ_n differ by $\delta\phi$, the variational principle (2.385) implies that the leading term of the difference between $E[\phi]$ and the true eigenvalue E_n is *quadratic* in $\delta\phi$. As a result, errors in the approximate energy are of *second order* in $\delta\phi$ when the energy is calculated from the functional (2.383).

We also remark that the functional (2.383) is independent of the normalisation and of the phase of ϕ . In particular, it is often convenient to impose the condition $\langle \phi | \phi \rangle = 1$. The above results may then be retrieved by varying the functional $\langle \phi | H | \phi \rangle$ subject to the condition $\langle \phi | \phi \rangle = 1$, namely

$$\delta \int \phi^* H \phi \, d\tau = 0, \quad \int \phi^* \phi \, d\tau = 1 \quad (2.391)$$

The constraint $\langle \phi | \phi \rangle = 1$ may be taken care of by introducing a Lagrange multiplier [9] which we denote by E , so that the variational equation reads

$$\delta \left[\int \phi^* H \phi \, d\tau - E \int \phi^* \phi \, d\tau \right] = 0 \quad (2.392)$$

or

$$\int \delta\phi^* (H - E) \phi \, d\tau + \int \phi^* (H - E) \delta\phi \, d\tau = 0 \quad (2.393)$$

This equation is identical to (2.387), and we see that the Lagrange multiplier E has the significance of an energy eigenvalue.

[9] Lagrange multipliers are discussed for example in Byron and Fuller (1969).

An important additional property of the functional (2.383) is that it provides an *upper bound* to the exact ground state energy E_0 . To prove this result, we expand the arbitrary, normalisable function ϕ in the complete set of orthonormal eigenfunctions ψ_n of H . That is,

$$\phi = \sum_n a_n \psi_n \quad (2.394)$$

Substituting (2.394) into (2.383), we find that

$$E[\phi] = \frac{\sum_n |a_n|^2 E_n}{\sum_n |a_n|^2} \quad (2.395)$$

where we have used the fact that $H\psi_n = E_n\psi_n$ and $\langle\phi|\phi\rangle = \sum_n |a_n|^2$. If we now subtract E_0 , the lowest energy eigenvalue, from both sides of (2.395) we have

$$E[\phi] - E_0 = \frac{\sum_n |a_n|^2 (E_n - E_0)}{\sum_n |a_n|^2} \quad (2.396)$$

Since $E_n \geq E_0$, the right-hand side of (2.396) is non-negative, so that

$$E_0 \leq E[\phi] \quad (2.397)$$

and the functional $E[\phi]$ gives an upper bound – or in other words a *minimum principle* for the ground state energy.

The property (2.397) constitutes the basis of the *Rayleigh–Ritz variational method* for the approximate calculation of E_0 . This method consists in evaluating the quantity $E[\phi]$ by using *trial functions* ϕ which depend on a certain number of *variational parameters*, and then to minimise $E[\phi]$ with respect to these parameters in order to obtain the best approximation of E_0 allowed by the form chosen for ϕ .

The Rayleigh–Ritz variational method can also be used to obtain an upper bound for the energy of an excited state, provided that the trial function ϕ is made orthogonal to all the energy eigenfunctions corresponding to states having a lower energy than the energy level considered. Indeed, let us arrange the energy levels in an ascending sequence: $E_0, E_1, E_2, \dots$ and let the trial function ϕ be orthogonal to the energy eigenfunctions ψ_n ($n = 0, 1, \dots, i$), namely

$$\langle\psi_n|\phi\rangle = 0, \quad n = 0, 1, \dots, i \quad (2.398)$$

Then, if we expand ϕ in the orthonormal set $\{\psi_n\}$ as in (2.394) we have $a_n = \langle\psi_n|\phi\rangle = 0$ ($n = 0, 1, \dots, i$) and the functional $E[\phi]$ becomes

$$E[\phi] = \frac{\sum_{n=i+1}^{\infty} |a_n|^2 E_n}{\sum_{n=i+1}^{\infty} |a_n|^2} \quad (2.399)$$

so that

$$E_{i+1} \leq E[\phi] \quad (2.400)$$

As an example, suppose that the lowest energy eigenfunction ψ_0 is known, and let ϕ be a trial function. The function

$$\tilde{\phi} = \phi - \psi_0 \langle\psi_0|\phi\rangle \quad (2.401)$$

is orthogonal to ψ_0 (that is, $\langle\psi_0|\tilde{\phi}\rangle = 0$) and can therefore be used to obtain an upper limit of E_1 , the exact energy of the first excited state.

In many practical situations the lower energy eigenfunctions ψ_n ($n = 0, 1, \dots, i$) are not known exactly and one only has approximations (obtained for example from a variational calculation) of these functions. In this case the orthogonality conditions (2.398) cannot be achieved exactly, and the relation (2.400) does not hold. For example, let us suppose that the function ϕ_0 (normalised to unity) is an approximation to the true ground state eigenfunction ψ_0 . If ϕ_1 is a trial function orthogonal to ϕ_0 (that is, if $\langle\phi_0|\phi_1\rangle = 0$) it may be shown (Problem 2.24) that

$$E_1 - \varepsilon_0(E_1 - E_0) \leq E[\phi_1] \quad (2.402)$$

where ε_0 is the positive quantity

$$\varepsilon_0 = 1 - |\langle\psi_0|\phi_0\rangle|^2 \quad (2.403)$$

Thus $E[\phi_1]$ does not provide a rigorous upper bound to E_1 . However, if ϕ_0 is a good approximation to ψ_0 , then ε_0 will be small and the violation of the relation $E_1 \leq E[\phi_1]$ will be mild.

The application of the variational method to excited states is greatly facilitated if the Hamiltonian of the system has certain *symmetry properties*, since in this case the orthogonality conditions (2.398) can be satisfied exactly for certain states. For example, if the excited state in question is of different parity or angular momentum than the lower states, then the orthogonality conditions are automatically satisfied.

Particularly useful trial functions ϕ can be constructed by choosing a certain number (N) of linearly independent functions $\chi_1, \chi_2, \dots, \chi_N$ and forming the linear combination

$$\phi = \sum_{n=1}^N c_n \chi_n \quad (2.404)$$

where the coefficients $c_1, c_2, \dots, c_N$ are linear variational parameters which must be determined by minimising the functional $E[\phi]$ in order to obtain the best approximation to E_0 . Substituting (2.404) in (2.383) we find that

$$E[\phi] = \frac{\sum_{n=1}^N \sum_{n'=1}^N c_n^* c_{n'} H_{nn'}}{\sum_{n=1}^N \sum_{n'=1}^N c_n^* c_{n'} \Delta_{nn'}} \quad (2.405)$$

where we have set

$$\begin{aligned} H_{n'n} &= \langle \chi_{n'} | H | \chi_n \rangle \\ \Delta_{n'n} &= \langle \chi_{n'} | \chi_n \rangle \end{aligned} \quad (2.406)$$

We remark that if the functions χ_n are orthonormal, then $\Delta_{n'n} = \delta_{n'n}$.

In order to find the values of the variational parameters $c_1, c_2, \dots, c_N$ which minimise $E[\phi]$, we first rewrite (2.405) as

$$E[\phi] \sum_{n=1}^N \sum_{n'=1}^N c_n^* c_{n'} \Delta_{n'n} = \sum_{n=1}^N \sum_{n'=1}^N c_n^* c_{n'} H_{n'n} \quad (2.407)$$

Differentiating with respect to each c_n or c_n^* , expressing that $\partial E / \partial c_n = 0$ (or $\partial E / \partial c_n^* = 0$), we obtain a system of N linear and homogeneous equations in the variables $c_1, c_2, \dots, c_N$, namely

$$\sum_{n=1}^N c_n (H_{n'n} - \Delta_{n'n} E) = 0; \quad n' = 1, 2, \dots, N \quad (2.408)$$

The necessary and sufficient condition for this system to have a non-trivial solution is that the determinant of the coefficients vanishes. That is,

$$\det |H_{n'n} - \Delta_{n'n} E| = 0 \quad (2.409)$$

Let $E_0^{(N)}, E_1^{(N)}, \dots, E_{N-1}^{(N)}$ be the N roots of this equation, arranged in an ascending sequence; the superscript (N) indicates that we are dealing with an $N \times N$ determinant. The lowest root $E_0^{(N)}$ is of course an upper bound to the ground state energy E_0 . Upon substituting $E_0^{(N)}$ in the system of equations (2.408) and solving for the coefficients c_n in terms of one of them (for example, c_1 , which may be used as a normalisation factor), we then obtain the corresponding 'optimum' approximation ϕ_0 to the ground state wave function ψ_0 . It may also be shown that the other roots $E_i^{(N)}$ of (2.409), with $i = 1, 2, \dots, N-1$, are upper bounds to excited state energies of the system. In particular, if the Hamiltonian commutes with a Hermitian operator A and the trial function (2.404) has been constructed from eigenfunctions corresponding to a given eigenvalue α of A , then the roots $E_i^{(N)}$, with $i = 1, 2, \dots, N-1$, are upper bounds to the energies E_i associated with excited states belonging to the eigenvalue α of A (for example, a given value of the angular momentum or the parity). The 'optimum' approximations $\phi_0, \phi_1, \dots, \phi_{N-1}$ to the true wave functions, obtained by the above method, may also be shown to be mutually orthogonal. Moreover, if we construct a new trial function $\hat{\phi}$ containing an additional basis function χ_{N+1} , namely

$$\hat{\phi} = \sum_{n=1}^{N+1} c_n \chi_n \quad (2.410)$$

it can be proved that the 'new' $(N+1)$ roots $E_0^{(N+1)}, E_1^{(N+1)}, \dots, E_N^{(N+1)}$ of the determinantal equation (2.409) are separated by the 'old' roots $E_0^{(N)}, E_1^{(N)}, \dots, E_{N-1}^{(N)}$. This property, which is illustrated in Fig. 2.10, is known as the *Hylleraas-Undheim theorem*.

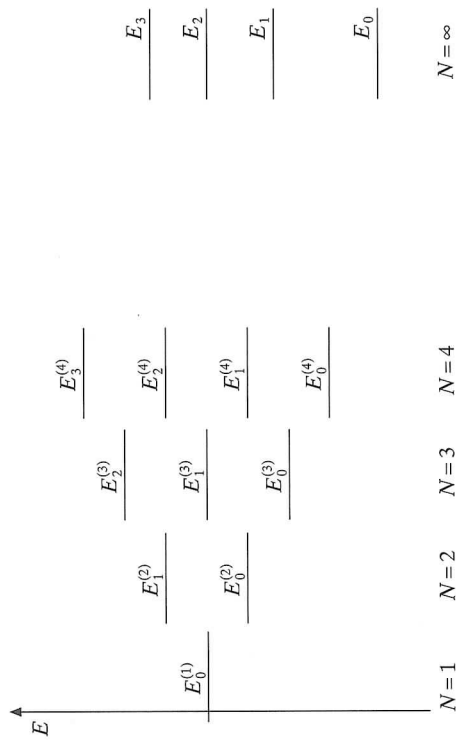


Figure 2.10 Approximate energy eigenvalues given by the Rayleigh-Ritz variational method with linear trial functions. Each root $E_i^{(N)}$ of the determinantal equation (2.409) is an upper bound to the corresponding exact eigenvalue E_i .

The variation-perturbation method

We have seen in our discussion of perturbation theory that according to the Rayleigh-Schrödinger method the solution $\psi_k^{(1)}$ of (2.304) (for the non-degenerate case) or the solution $\psi_{kr}^{(1)}$ of (2.323) (for the degenerate case) is obtained by expanding these functions in the basis set $\{\psi_m^{(0)}\}$ of the unperturbed eigenfunctions. The coefficients of this expansion and the second- and third-order corrections to the energy are then given in terms of matrix elements of H' between unperturbed eigenfunctions (see for example (2.316) and (2.319)). Unfortunately, in many cases the evaluation of all the necessary matrix elements of H' and of the required summations is very difficult. In these cases, however, it is possible to obtain *approximations* to the perturbed energy levels and eigenfunctions by using the *variation-perturbation method*, which we shall now discuss for the non-degenerate case.

Let us assume that the unperturbed eigenfunction $\psi_k^{(0)}$ as well as the corresponding energy $E_k^{(0)}$ and the first-order energy correction $E_k^{(1)} = \langle \psi_k^{(0)} | H' | \psi_k^{(0)} \rangle$ are known for a given state k . Let $\phi_k^{(1)}$ be an arbitrary trial function and $F_1[\phi_k^{(1)}]$ the functional

$$F_1[\phi_k^{(1)}] = \langle \phi_k^{(1)} | H_0 - E_k^{(0)} | \phi_k^{(1)} \rangle + 2 \langle \phi_k^{(1)} | H' - E_k^{(1)} | \psi_k^{(0)} \rangle \quad (2.411)$$

We now express that this functional is *stationary* for variations of $\phi_k^{(1)}$ around the correct first-order wave function $\psi_k^{(1)}$. Proceeding as in the case of the variational method studied above, we find that the variational equation

$$\delta F_1 = 0 \quad (2.412)$$

implies that

$$(H_0 - E_k^{(0)})\phi_k^{(1)} + (H' - E_k^{(1)})\psi_k^{(0)} = 0 \quad (2.413)$$

so that the function $\phi_k^{(1)}$ which makes the functional F_1 stationary is a solution $\psi_k^{(1)}$ of the equation (2.304). Moreover, by comparing (2.304), (2.310) and (2.411) we see that $F_1[\phi_k^{(1)}]$ reduces to the correct value of $E_k^{(2)}$ when $\phi_k^{(1)} \equiv \psi_k^{(1)}$.

Let us now consider the particular case of a state k for which $E_k^{(0)}$ is the *lowest* eigenvalue of H_0 corresponding to a given symmetry. It is then straightforward to show that (compare with (2.397))

$$E_k^{(2)} \leq F_1[\phi_k^{(1)}] \quad (2.414)$$

One can then proceed in a way similar to that followed in the Rayleigh-Ritz method. First, a trial function $\phi_k^{(1)}$ is chosen containing a certain number of variational parameters. One then inserts that trial function in the functional $F_1[\phi_k^{(1)}]$, which is minimised with respect to the variational parameters. Because of (2.414) this minimum value of $F_1[\phi_k^{(1)}]$ gives an *upper bound* to $E_k^{(2)}$. The corresponding 'optimum' function $\phi_k^{(1)}$ can also be used to calculate an approximate value of $E_k^{(3)}$ by replacing $\psi_k^{(1)}$ by $\phi_k^{(1)}$ in (2.312).

The method can be extended to calculate approximate values of $\psi_k^{(2)}$, $E_k^{(4)}$ and $E_k^{(5)}$. To this end, we consider a trial function $\phi_k^{(2)}$ and the functional

$$F_2[\phi_k^{(2)}] = \langle \phi_k^{(2)} | H_0 - E_k^{(0)} | \phi_k^{(2)} \rangle + 2 \langle \phi_k^{(2)} | H' - E_k^{(1)} | \psi_k^{(1)} \rangle - 2 E_k^{(2)} \langle \phi_k^{(2)} | \psi_k^{(0)} \rangle \quad (2.415)$$

where $\psi_k^{(1)}$ is now assumed to be known. Let us express that this functional is stationary with respect to variations of $\phi_k^{(2)}$. The condition

$$\delta F_2 = 0 \quad (2.416)$$

yields the equation

$$(H_0 - E_k^{(0)})\phi_k^{(2)} + (H' - E_k^{(1)})\psi_k^{(1)} - E_k^{(2)}\psi_k^{(0)} = 0 \quad (2.417)$$

so that the function $\phi_k^{(2)}$ which makes the functional F_2 stationary is a solution $\psi_k^{(2)}$ of (2.305). In addition, it is easily shown that when $\phi_k^{(2)} \equiv \psi_k^{(2)}$ the functional F_2 gives the exact value of $E_k^{(4)}$, apart from terms which are independent of $\phi_k^{(2)}$. Moreover, F_2 yields a minimum principle (upper bound) for $E_k^{(4)}$ when the state k is such that $E_k^{(0)}$ is the *lowest* eigenvalue of H_0 corresponding to a given symmetry. The function $\phi_k^{(2)}$ determined from the variational principle (2.416) can also be used to calculate the energy correction $E_k^{(5)}$. By constructing functionals $F_3[\phi_k^{(3)}]$, $F_4[\phi_k^{(4)}]$, and so on, the variation-perturbation method can be used to calculate higher order perturbation corrections to the wave functions and the energy eigenvalues.

The WKB approximation

We shall conclude this chapter with a discussion of an approximation method which can be used when the potential energy is a slowly varying function of position.

The method we shall study is called the *WKB approximation*, after G. Wentzel, H.A. Kramers and L. Brillouin, who independently introduced it in quantum mechanics in 1926. It can only be readily applied to one-dimensional problems, and to problems of higher dimensionality that can be reduced to the solution of a one-dimensional Schrödinger equation. We shall therefore focus our attention on the one-dimensional motion of a particle of mass m in a potential $V(x)$, the corresponding time-independent Schrödinger equation being the equation (2.121). If the potential were constant, $V(x) = V_0$, solutions of (2.121) would be written as linear combinations of the plane waves

$$\psi(x) = A \exp\left(\pm \frac{i}{\hbar} p_0 x\right) \quad (2.418)$$

where

$$p_0 = [2m(E - V_0)]^{1/2} \quad (2.419)$$

and A is a constant. The fact that the potential is 'slowly varying' means that $V(x)$ changes only slightly over the de Broglie wavelength

$$\lambda(x) = \frac{h}{p(x)} \quad (2.420)$$

associated with the particle (having an energy $E > V$) in this potential, where

$$p(x) = \{2m[E - V(x)]\}^{1/2} \quad (2.421)$$

is the classical momentum at the point x . Now, in the classical limit ($\hbar \rightarrow 0$) the de Broglie wavelength λ tends to zero, so that the slow variation of the potential in a de Broglie wavelength is always satisfied for physically realizable potentials. In the quantum case, the fact that the potential varies slowly over a de Broglie wavelength means that we are dealing with a quasi-classical (or semi-classical) situation. For this reason, the WKB method is said to belong to the category of *semi-classical approximations*.

The result (2.418) obtained for constant potentials suggests that in the case of slowly varying potentials, solutions of the Schrödinger equation (2.121) should be sought which have the form

$$\psi(x) = A \exp\left[\frac{i}{\hbar} S(x)\right] \quad (2.422)$$

Substituting (2.422) into (2.121), we obtain for $S(x)$ the equation

$$-\frac{\hbar^2}{2m} \frac{d^2 S(x)}{dx^2} + \frac{1}{2m} \left[\frac{dS(x)}{dx} \right]^2 + V(x) - E = 0 \quad (2.423)$$

So far, no approximation has been made, the above equation being strictly equivalent to the original Schrödinger equation (2.121). Unfortunately, equation (2.423) is a non-linear equation which is in fact more complicated than the Schrödinger equation (2.121) itself! We must therefore try to solve (2.423) approximately. To this end, we first remark that if the potential is constant, then $S(x) = \pm p_0 x$ (see (2.418)) and the first term on the left of (2.423) vanishes. Moreover, this term is proportional to $\hbar$, and hence vanishes in the classical limit. This suggests that we treat $\hbar$ as a parameter of smallness and expand the function $S(x)$ in the power series

$$S(x) = S_0(x) + \hbar S_1(x) + \frac{\hbar^2}{2} S_2(x) + \dots \quad (2.424)$$

Inserting this expansion into (2.423), and equating to zero the coefficients of each power of $\hbar$ separately, we find the set of equations (Problem 2.26)

$$-\frac{1}{2m} \left[\frac{dS_0(x)}{dx} \right]^2 + V(x) - E = 0 \quad (2.425a)$$

$$\frac{dS_0(x)}{dx} \frac{dS_1(x)}{dx} - \frac{i}{2} \frac{d^2 S_0(x)}{dx^2} = 0 \quad (2.425b)$$

$$\frac{dS_0(x)}{dx} \frac{dS_2(x)}{dx} + \left[\frac{dS_1(x)}{dx} \right]^2 - i \frac{d^2 S_1(x)}{dx^2} = 0 \quad (2.425c)$$

which must be solved successively to find $S_0(x)$, $S_1(x)$, $S_2(x)$, and so on. As we shall see below, the WKB method yields the first two terms [10] of the expansion (2.424).

Let us begin by solving the equation (2.425a). Assuming first that $E > V(x)$, so that we are in a classically allowed region of positive kinetic energy, we find that

$$S_0(x) = \pm \int^x p(x') dx' \quad (2.426)$$

where $p(x)$ is the classical momentum given by (2.421) and an arbitrary integration constant, which can be absorbed in the coefficient A on the right of (2.422), has been omitted. Substituting (2.426) into (2.422), we obtain the simple semi-classical wave function

$$\psi(x) = A \exp \left[\pm i \int^x \kappa(x') dx' \right] \quad (2.427)$$

[10] It can be shown that the series (2.424) does not converge, but is an asymptotic series for the function $S(x)$. As a result, the best approximation to $S(x)$ is obtained by keeping a finite number of terms on the right-hand side of (2.424).

where

$$\kappa(x) = \hbar^{-1} p(x) = \hbar^{-1} [2m(E - V(x))]^{1/2} \quad (2.428)$$

is the 'effective wave number'. If we assume that $E \gg |V(x)|$, we may expand the effective wave number as

$$\kappa(x) = k \left[1 - \frac{1}{2} \frac{V(x)}{E} + \dots \right] = k - \frac{U(x)}{2k} + \dots \quad (2.429)$$

where $k = (2mE)^{1/2}/\hbar$ and $U(x) = 2mV(x)/\hbar^2$. The semi-classical wave function (2.427) then becomes

$$\psi(x) = A \exp \left\{ \pm i \left[kx - \frac{1}{2k} \int^x U(x') dx' \right] \right\} \quad (2.430)$$

which is called a (one-dimensional) *eikonal* wave function.

Using the result (2.426) in equation (2.425b), the function $S_1(x)$ is found to be

$$S_1(x) = \frac{i}{2} \log p(x) \quad (2.431)$$

where an arbitrary constant of integration has again been omitted. Next, by using the expressions for $S_0(x)$ and $S_1(x)$ in (2.425c) one finds (Problem 2.27) that

$$S_2(x) = \frac{1}{2} m [p(x)]^{-3} \frac{dV(x)}{dx} - \frac{1}{4} m^2 \int^x [p(x')]^{-5} \left[\frac{dV(x')}{dx'} \right]^2 dx' \quad (2.432)$$

From this expression and that of $p(x)$ (see (2.421)), it is clear that S_2 will be small whenever $dV(x)/dx$ is small, and provided that $E - V$ is not too close to zero. If, in addition, all the higher derivatives of $V(x)$ are small, it can be verified that $S_3, S_4, \dots$ will also be small. Note that in the case of a constant potential $S_0 = \pm p_0 x$ (apart from an irrelevant additive constant) and $S_1, S_2, \dots$ are all zero.

These considerations show that if $V(x)$ is a slowly varying function of position, and provided $E - V$ is not too small, one can retain only the first two terms on the right of the expansion (2.424). Using (2.426) and (2.431), we then obtain the *WKB approximation* to the wave function (2.422) in a classically allowed region, namely

$$\psi(x) = A [p(x)]^{-1/2} \exp \left[\pm \frac{i}{\hbar} \int^x p(x') dx' \right], \quad E > V \quad (2.433)$$

Of course, the general WKB solution in this region is a linear combination of such solutions. That is,

$$\psi(x) = [p(x)]^{-1/2} \left\{ A \exp \left[\frac{i}{\hbar} \int^x p(x') dx' \right] + B \exp \left[-\frac{i}{\hbar} \int^x p(x') dx' \right] \right\}, \quad E > V \quad (2.434)$$

where A and B are arbitrary constants. The first exponential corresponds to a wave moving in the positive direction, the second exponential to a wave moving in the opposite direction. We also remark that if the potential is constant ($V = V_0$) these exponentials reduce to the plane waves $\exp(ip_0 x/\hbar)$ and $\exp(-ip_0 x/\hbar)$, respectively, where p_0 is given by (2.419).

A criterion of validity for the WKB approximation may be obtained by requiring that the contribution to the wave function (2.422) arising from the third term on the right of (2.424) be negligible. This will be the case if

$$|(\hbar/2)S_2(x)| \ll 1 \quad (2.435)$$

Using the expression (2.432) of $S_2(x)$ and taking into account that both terms on the right of (2.432) are of the same order of magnitude, we may rewrite the condition (2.435) more explicitly as

$$\left| \frac{\hbar m dV(x)/dx}{[2m(E - V(x))]^{3/2}} \right| \ll 1 \quad (2.436)$$

where we have used the expression (2.421) of $p(x)$. This criterion is seen to be satisfied if the potential energy varies slowly enough, provided that the kinetic energy $E - V$ is sufficiently large. An alternative way of writing the condition (2.436) is

$$\left| \frac{\hbar}{[p(x)]^2} \frac{dp(x)}{dx} \right| = \left| \frac{1}{p(x)} \frac{dp(x)}{dx} \right| \ll 1 \quad (2.437)$$

where $\lambda(x) = \hbar/p(x) = \lambda(x)/(2\pi)$ is the reduced de Broglie wavelength of the particle. The inequality (2.437) means that the fractional change in the momentum must be small in a wavelength. In other words, the potential must change so slowly that the momentum of the particle is nearly constant over many wavelengths. We remark that the condition (2.437) may equally be written in the form

$$\left| \frac{d\lambda(x)}{dx} \right| \ll 1 \quad (2.438)$$

Now $\delta\lambda = (d\lambda/dx)\delta x$ is the change occurring in the reduced wavelength λ in the distance δx . Upon setting $\delta x = \lambda$, we see that the condition (2.438) is equivalent to

$$|\delta\lambda(x)| = \left| \frac{d\lambda(x)}{dx} \lambda(x) \right| \ll \lambda(x) \quad (2.439)$$

showing that λ must only change by a small fraction of itself over a distance of the order of λ . This condition is well known in wave optics as the one to be satisfied if a wave is to propagate in a medium of varying index of refraction without giving rise to significant reflection.

Until now we have only considered classically allowed regions for which $E > V$. However, we can equally well obtain the WKB approximation to the wave function (2.422) in classically forbidden regions of negative kinetic energy ($E < V$), where $p(x)$ becomes purely imaginary (see (2.421)). Solving equations (2.425a) and (2.425b) in this case, we obtain the WKB solutions (Problem 2.28)

$$\psi(x) = A|p(x)|^{-1/2} \exp \left[\pm \frac{1}{\hbar} \int^x |p(x')| dx' \right], \quad E < V \quad (2.440)$$

where A is a constant. The general WKB solution for $E < V$ is a linear combination of such solutions, namely

$$\psi(x) = |p(x)|^{-1/2} \left\{ C \exp \left[-\frac{1}{\hbar} \int^x |p(x')| dx' \right] + D \exp \left[\frac{1}{\hbar} \int^x |p(x')| dx' \right] \right\}, \quad E < V \quad (2.441)$$

where C and D are arbitrary constants. These WKB solutions are accurate provided the criterion (2.436) is satisfied, which requires that $V(x)$ varies slowly enough and that $V - E$ be sufficiently large.

It is apparent from (2.436) that the WKB approximation breaks down in the vicinity of a *classical turning point*, for which $E = V$. The two kinds of WKB wave functions (2.434) and (2.441) we have obtained are therefore 'asymptotically' valid, because they may only be used sufficiently far away from the nearest turning point. Since the exact wave function is continuous and smooth for all x , it is possible to obtain *connection formulae* which allow one to join smoothly the two types of WKB solutions across a turning point. This interpolation is based on the fact that the original Schrödinger equation (2.121) can be modified slightly so that an 'exact' solution at and near a turning point can be written down. A detailed discussion of the WKB connection formulae may be found in Bransden and Joachain (2000).

Problems

2.1 Consider the momentum space wave function

$$\phi(p_x) = 0, \quad |p_x - p_0| > \gamma \\ = C, \quad |p_x - p_0| \leq \gamma$$

where p_0 , C and γ are constants. Find the corresponding wave function $\psi(x)$ in configuration space and determine the constant C , so that $\psi(x)$ satisfies the normalisation condition

$$\int_{-\infty}^{\infty} dx |\psi(x)|^2 = 1$$

Using a reasonable definition of the 'width' Δx of $|\psi(x)|^2$, show that $\Delta x \Delta p_x \geq \hbar$.

2.2 Consider an electron of momentum p in the Coulomb field of a proton. The total energy is

$$E = \frac{p^2}{2m} - \frac{e^2}{(4\pi\epsilon_0)r}$$

where r is the distance of the electron from the proton. Assuming that the uncertainty Δr of the radial coordinate is $\Delta r \approx r$ and that $\Delta p \approx p$, use Heisenberg's uncertainty principle to obtain estimates of the size and of the energy of the hydrogen atom in the ground state.

2.3 Consider the one-dimensional wave function

$$\psi_{p_x}(x) = C \exp(ip_x x / \hbar)$$

which is an eigenfunction of the operator $(p_x)_{\text{op}} = -i\hbar \partial/\partial x$ corresponding to the eigenvalue $p_x = \hbar k$, and is therefore called a *momentum eigenfunction*.

(a) Determine the constant C (independent of x) so that $\psi_{p_x}(x)$ is 'normalised' according to the *p-normalisation* condition

$$\int \psi_{p'_x}^*(x) \psi_{p_x}(x) dx = \delta(p_x - p'_x)$$

(b) Determine the constant C so that $\psi_{p_x}(x)$ is 'normalised' on the *energy scale* (energy normalisation) according to

$$\int \psi_{p'_x}^*(x) \psi_{p_x}(x) dx = \delta(E - E')$$

where $E = p_x^2/(2m)$ and $E' = p'^2/(2m)$.

(Hint: Use the representation (2.31a) and the properties (2.33d) and (2.33g) of the Dirac delta function.)

2.4 Consider the normalised Gaussian wave packet $\psi(x) = N \exp(-\lambda^2 x^2)$ where $N = (2\lambda^2/\pi)^{1/4}$. Using (2.16), calculate the uncertainties Δx and Δp_x from (2.41) and show that $\Delta x \Delta p_x = \hbar/2$.

2.5 Starting from (2.39) show that

$$\langle \mathbf{p} \rangle = \int \Phi^*(\mathbf{p}, t) \mathbf{p} \Phi(\mathbf{p}, t) d\mathbf{p}$$

where $\Phi(\mathbf{p}, t)$ is the momentum space wave function (normalised to unity) corresponding to $\Psi(\mathbf{r}, t)$ (see (2.20)).

2.6 A particle of mass m moving in a potential $V(\mathbf{r})$ has the Hamiltonian $H = -(\hbar^2/2m)\nabla^2 + V(\mathbf{r})$. Using the result (2.56) prove Ehrenfest's theorem (2.57).

2.7 (a) Using the definition of an adjoint operator (2.65) prove that if A and B are two operators then $(AB)^\dagger = B^\dagger A^\dagger$.

(b) Suppose that A and B are two non-commuting Hermitian operators. Determine which of the following operators are Hermitian: (i) AB (ii) $[A, B]$ (iii) $AB + BA$ (iv) ABA (v) A^n where n is a positive integer.

2.8 Let E_n denote the energy eigenvalues of a one-dimensional system and $\psi_n(x)$ the corresponding energy eigenfunctions. Suppose that the normalised wave function of the system at $t = 0$ is given by

$$\Psi(x, t = 0) = \frac{1}{\sqrt{2}} \exp(i\alpha_1) \psi_1(x) + \frac{1}{\sqrt{3}} \exp(i\alpha_2) \psi_2(x) + \frac{1}{\sqrt{6}} \exp(i\alpha_3) \psi_3(x)$$

where the α_i are constants.

(a) Write down the wave function $\Psi(x, t)$ at time t .

(b) Find the probability that at time t a measurement of the energy of the system gives the value E_2 .

(c) Does $\langle x \rangle$ vary with time? Does $\langle p_x \rangle$ vary with time? Does $E = \langle H \rangle$ vary with time?

2.9 If U is a unitary operator and $A' = UAU^\dagger$ where A is a Hermitian operator, show that:

(a) A' is Hermitian.

(b) The eigenvalues of A' are the same as those of A .

(c) $\langle \Phi' | A' | \Psi' \rangle = \langle \Phi | A | \Psi \rangle$ where Ψ' and Φ' are related to Ψ and Φ by (2.106).

2.10 Consider a system described by a time-independent Hamiltonian H . Let A and B be two non-commuting operators ($[A, B] \neq 0$) which are constants of the motion, so that they commute with H . Prove that in this case the energy levels of the system are, in general, degenerate.

2.11 Consider a particle of mass m moving in one dimension in the infinite square well (2.122). Suppose that at time $t = 0$ its wave function is given by

$$\Psi(x, t = 0) = A(a^2 - x^2)$$

(a) Find the probability P_n that a measurement of the energy will give the value E_n where E_n is given by (2.131).

(b) Obtain the average value of the energy.

2.12 Classically, the linear harmonic oscillator having the total energy $E = p^2/2m + m\omega^2 x^2/2$ could have $x = p = 0$ and hence $E = 0$. In quantum mechanics, there will be uncertainties $x = 0 + \Delta x$, $p_x = 0 + \Delta p_x$ such that $\Delta x \Delta p_x \approx \hbar$. Show that because of this condition the minimum total energy cannot be zero and must be of the order of $\hbar\omega$.

2.13 Starting from the definitions (2.154) show that L_x , L_y and L_z can be expressed in spherical polar coordinates (see (2.157)) as

$$\begin{aligned} L_x &= i\hbar \left(\sin \phi \frac{\partial}{\partial \theta} + \cot \theta \cos \phi \frac{\partial}{\partial \phi} \right) \\ L_y &= i\hbar \left(-\cos \phi \frac{\partial}{\partial \theta} + \cot \theta \sin \phi \frac{\partial}{\partial \phi} \right) \\ L_z &= -i\hbar \frac{\partial}{\partial \phi} \end{aligned}$$

and verify that $\mathbf{L}^2 = L_x^2 + L_y^2 + L_z^2$ is given by (2.159). Using these results show that $[\mathbf{L}, f(r)] = 0$ and $[\mathbf{L}^2, f(r)] = 0$ where $f(r)$ is an arbitrary function of the radial coordinate r . Show also that $[\mathbf{L}, \mathbf{p}^2] = 0$, where $\mathbf{p}^2 = -\hbar^2 \nabla^2$.

2.14 Obtain the results (2.187) and (2.188). Prove also that for a spin-1/2 particle, one has

$$\langle S_x \rangle = \langle S_y \rangle = 0 \text{ while } \langle S_x^2 \rangle = \langle S_y^2 \rangle = \hbar^2/4.$$

2.15 Verify the equations (2.217) and (2.218).

2.16 Write down explicitly the matrices representing the operators S_x , S_y , S_z , S_+ , S_- and S^2 for $s = 1$ and $s = 3/2$.

2.17 Prove that $\mathbf{L}^2$, $\mathbf{S}^2$, $\mathbf{J}^2$ and J_z form a commuting set of operators, and that $[(\mathbf{L} \cdot \mathbf{S}), J_z] = 0$.

2.18 Assuming that the expansion of a plane wave in a series of Legendre polynomials is of the form

$$\exp(ikr \cos \theta) = \sum_{l=0}^{\infty} c_l j_l(kr) P_l(\cos \theta)$$

find the coefficients c_l .

(Hint: Use the fact that $j_l(kr) \rightarrow (kr)^{-1} \sin(kr - l\pi/2)$.)

2.19 Show that $E_k^{(2)}$ and $E_k^{(3)}$ are given by (2.311) and (2.312) respectively.

2.20 Consider a linear harmonic oscillator for which the Hamiltonian is

$$H_0 = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} kx^2$$

If this oscillator is perturbed by an additional potential of the form $H' = \frac{1}{2} k' x^2$, find the first- and second-order corrections to the energy levels using perturbation theory. Also find the first-order correction to the wave function. Compare your results with the exact solution

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right) \left(1 + \frac{k'}{k} \right)^{1/2}$$

where $\omega = (k/m)^{1/2}$. (Hint: Obtain first the matrix elements:

$$\begin{aligned} x_{ij}^2 &= \langle \psi_i | x^2 | \psi_j \rangle \\ &= \begin{cases} (2\alpha^2)^{-1} [(j+1)(j+2)]^{1/2} & i = j+2 \\ (2\alpha^2)^{-1} (2j+1) & i = j \\ (2\alpha^2)^{-1} [j(j-1)]^{1/2} & i = j-2 \end{cases} \end{aligned}$$

where $\alpha = \left(\frac{mk}{\hbar^2} \right)^{1/4}$.)

2.21 A two-dimensional isotropic harmonic oscillator has the Hamiltonian

$$H = -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) + \frac{1}{2} k(x^2 + y^2)$$

(a) Show that the energy levels are given by

$$\begin{aligned} E_{n_x n_y} &= \hbar\omega(n_x + n_y + 1); & n_x &= 0, 1, 2, \dots \\ & & n_y &= 0, 1, 2, \dots, & \omega &= (k/m)^{1/2} \end{aligned}$$

What is the degeneracy of each level?

(b) If this oscillator is perturbed by an interaction of the form $H' = \lambda xy$, where λ is a constant, find the first-order modification of the energy of the first excited level.

2.22 Consider a particle of charge q and mass m , which is in simple harmonic motion along the X axis with a force constant k . An electric field $\mathcal{E}(t)$, directed along the X axis, is switched on at time $t = 0$ so that the system is perturbed by an interaction

$$H'(t) = -qX\mathcal{E}(t)$$

If $\mathcal{E}(t)$ has the form

$$\mathcal{E}(t) = \mathcal{E}_0 \exp(-t/\tau)$$

where $\mathcal{E}_0$ and τ are constants, and if the oscillator is in the ground state for $t \leq 0$, find the probability that the oscillator will be found in an excited state as $t \rightarrow \infty$.

2.23 (a) By varying the parameter a in the trial function

$$\phi_0(x) = (a^2 - x^2)^2, \quad |x| < a \\ = 0 \quad |x| \geq a$$

obtain an upper bound for the ground state energy of a linear harmonic oscillator having the Hamiltonian

$$H = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} m \omega^2 x^2$$

(b) Show that the function $\phi_1(x) = x\phi_0(x)$ is a suitable trial function for the first excited state and obtain a variational estimate of the energy of this level.

2.24 Prove the result given in equation (2.402).

2.25 Let $E_{n,l}$ denote the discrete energy levels of a particle of mass m in a central potential $V(r)$, corresponding to a given orbital angular momentum quantum number l , and let $E_{n,l}^{\min}$ be their minimum value. Prove that $E_{n,l}^{\min} < E_{n,l+1}^{\min}$ (Hint: Write the Hamiltonian of the particle as

$$H = H_r + \frac{\mathbf{L}^2}{2mr^2}, \quad H_r = -\frac{\hbar^2}{2m} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + V(r)$$

and note that H_r is a purely radial operator.)

2.26 Derive the equations (2.425) for $S_0(x)$, $S_1(x)$ and $S_2(x)$.

2.27 Derive the expression (2.432) for $S_2(x)$.

2.28 Solve the equations (2.425a) and (2.425b) in the case $E < V$ to obtain the WKB solutions (2.440).

3 One-electron atoms

In this chapter we begin our quantum mechanical study of atomic structure by considering the simplest atom, namely the hydrogen atom, which consists of a proton and an electron. Apart from small corrections, which we shall discuss in Chapter 5, the hydrogen atom may be considered as a non-relativistic system of two particles interacting by means of an attractive Coulomb potential. Other similar one-electron systems, called hydrogenic atoms, include the isotopes of hydrogen (deuterium, tritium) and the hydrogenic ions (He^+ , Li^{++} , etc.) which we have already encountered in our study of the Bohr model. These hydrogenic systems will also be discussed in the present chapter.

Our starting point is the Schrödinger equation for one-electron atoms. After separating the centre of mass motion, we solve the eigenvalue equation for the relative motion in spherical polar coordinates, and obtain the energy levels and wave functions of the discrete spectrum. We then consider the expectation values of various operators, and prove the virial theorem. This is followed by a study of the bound state wave functions of one-electron atoms in parabolic coordinates. We conclude this chapter with a brief discussion of 'special' hydrogenic systems such as positronium, muonium, antihydrogen, muonic and hadronic atoms, and Rydberg atoms.

3.1 The Schrödinger equation for one-electron atoms

Let us consider a hydrogenic atom containing an atomic nucleus of charge Ze and an electron of charge $-e$ interacting by means of the Coulomb potential

$$V(r) = -\frac{Ze^2}{(4\pi\epsilon_0)r} \quad (3.1)$$

where r is the distance between the two particles. We denote by m the mass of the electron and M the mass of the nucleus. Since the interaction potential (3.1) only depends on the relative coordinate of the two particles, we may use the results of Section 2.7 to separate the motion of the centre of mass. Remembering that $\mathbf{P}$ is the total momentum associated with the motion of the centre of mass and $\mathbf{p}$ the relative momentum, the total energy of the atom can be split into two parts. The first one is the kinetic energy $P^2/[2(M+m)]$ corresponding to the motion of