Quantum Mechanics

Richard Fitzpatrick

Professor of Physics
The University of Texas at Austin

Contents

1 Introduction

1.1 Intended audience

These lecture notes outline a single semester course on *non-relativistic quantum mechanics* which is primarily intended for *upper-division undergraduate physics majors*. The course assumes some previous knowledge of physics and mathematics. In particular, prospective students should be reasonably familiar with Newtonian dynamics, elementary classical electromagnetism and special relativity, the physics and mathematics of waves (including the representation of waves via complex functions), basic probability theory, ordinary and partial differential equations, linear algebra, vector algebra, and Fourier series and transforms.

1.2 Major Sources

The textbooks which I have consulted most frequently whilst developing course material are:

The Principles of Quantum Mechanics, P.A.M. Dirac, 4th Edition (revised), (Oxford University Press, Oxford UK, 1958).

Quantum Mechanics, E. Merzbacher, 2nd Edition, (John Wiley & Sons, New York NY, 1970).

Introduction to the Quantum Theory, D. Park, 2nd Edition, (McGraw-Hill, New York NY, 1974).

Modern Quantum Mechanics, J.J. Sakurai, (Benjamin/Cummings, Menlo Park CA, 1985).

Quantum Theory, D. Bohm, (Dover, New York NY, 1989).

Problems in Quantum Mechanics, G.L. Squires, (Cambridge University Press, Cambridge UK, 1995).

Quantum Physics, S. Gasiorowicz, 2nd Edition, (John Wiley & Sons, New York NY, 1996).

Nonclassical Physics, R. Harris, (Addison-Wesley, Menlo Park CA, 1998).

Introduction to Quantum Mechanics, D.J. Griffiths, 2nd Edition, (Pearson Prentice Hall, Upper Saddle River NJ, 2005).

1.3 Aim of Course

The aim of this course is to develop non-relativistic quantum mechanics as a complete theory of microscopic dynamics, capable of making detailed predictions, with a minimum of abstract mathematics.

1.4 Outline of Course

The first part of the course is devoted to an in-depth exploration of the basic principles of quantum mechanics. After a brief review of probability theory, in Chapter 2, we shall start, in Chapter 3, by examining how many of the central ideas of quantum mechanics are a direct consequence of wave-particle duality—*i.e.*, the concept that waves sometimes act as particles, and particles as waves. We shall then proceed to investigate the rules of quantum mechanics in a more systematic fashion in Chapter 4. Quantum mechanics is used to examine the motion of a single particle in one dimension, many particles in one dimension, and a single particle in three dimensions, in Chapters 5, 6, and 7, respectively. Chapter 8 is devoted to the investigation of orbital angular momentum, and Chapter 9 to the closely related subject of particle motion in a central potential. Finally, in Chapters 10 and 11, we shall examine spin angular momentum, and the addition of orbital and spin angular momentum, respectively.

The second part of this course describes selected practical applications of quantum mechanics. In Chapter 12, time-independent perturbation theory is used to investigate the Stark effect, the Zeeman effect, fine structure, and hyperfine structure, in the hydrogen atom. Time-dependent perturbation theory is employed to study radiative transitions in the hydrogen atom in Chapter 13. Chapter 14 illustrates the use of variational methods in quantum mechanics. Finally, Chapter 15 contains an introduction to quantum scattering theory.

2 Probability Theory

2.1 Introduction

This section is devoted to a brief, and fairly low level, introduction to a branch of mathematics known as *probability theory*.

2.2 What is Probability?

What is the *scientific* definition of probability? Well, let us consider an observation made on a general system, S. This can result in any one of a number of different possible outcomes. Suppose that we wish to find the probability of some general outcome, X. In order to ascribe a probability, we have to consider the system as a member of a large set, Σ, of similar systems. Mathematicians have a fancy name for a large group of similar systems. They call such a group an *ensemble*, which is just the French for "group." So, let us consider an ensemble, Σ, of similar systems, S. The probability of the outcome X is defined as the ratio of the number of systems in the ensemble which exhibit this outcome to the total number of systems, in the limit that the latter number tends to infinity. We can write this symbolically as

$$P(X) = \lim_{\Omega(\Sigma)\to\infty} \frac{\Omega(X)}{\Omega(\Sigma)}, \tag{2.1}$$

where $\Omega(\Sigma)$ is the total number of systems in the ensemble, and $\Omega(X)$ the number of systems exhibiting the outcome X. We can see that the probability $P(X)$ must be a number between 0 and 1. The probability is *zero* if no systems exhibit the outcome X, even when the number of systems goes to infinity. This is just another way of saying that there is *no chance* of the outcome X. The probability is *unity* if all systems exhibit the outcome X in the limit as the number of systems goes to infinity. This is another way of saying that the outcome X is *bound* to occur.

2.3 Combining Probabilities

Consider two *distinct* possible outcomes, X and Y, of an observation made on the system S, with probabilities of occurrence $P(X)$ and $P(Y)$, respectively. Let us determine the probability of obtaining the outcome X *or* Y, which we shall denote $P(X \mid Y)$. From the basic definition of probability,

$$P(X \mid Y) = \lim_{\Omega(\Sigma)\to\infty} \frac{\Omega(X \mid Y)}{\Omega(\Sigma)}, \tag{2.2}$$

where $\Omega(X \mid Y)$ is the number of systems in the ensemble which exhibit either the outcome X or the outcome Y. Now,

$$\Omega(X \mid Y) = \Omega(X) + \Omega(Y) \tag{2.3}$$

if the outcomes X and Y are mutually exclusive (which must be the case if they are two distinct outcomes). Thus,

$$P(X \mid Y) = P(X) + P(Y). \tag{2.4}$$

So, the probability of the outcome X *or* the outcome Y is just the *sum* of the individual probabilities of X and Y. For instance, with a six-sided die the probability of throwing any particular number (one to six) is 1/6, because all of the possible outcomes are considered to be equally likely. It follows, from what has just been said, that the probability of throwing either a one or a two is simply $1/6 + 1/6$, which equals 1/3.

Let us denote all of the M, say, possible outcomes of an observation made on the system S by X_i, where i runs from 1 to M. Let us determine the probability of obtaining *any* of these outcomes. This quantity is unity, from the basic definition of probability, because each of the systems in the ensemble must exhibit one of the possible outcomes. But, this quantity is also equal to the sum of the probabilities of all the individual outcomes, by (2.4), so we conclude that this sum is equal to unity: *i.e.*,

$$\sum_{i=1}^{M} P(X_i) = 1. \tag{2.5}$$

The above expression is called the *normalization condition*, and must be satisfied by any complete set of probabilities. This condition is equivalent to the self-evident statement that an observation of a system must definitely result in one of its possible outcomes.

There is another way in which we can combine probabilities. Suppose that we make an observation on a system picked at random from the ensemble, and then pick a second system *completely independently* and make another observation. We are assuming here that the first observation does not influence the second observation in any way. The fancy mathematical way of saying this is that the two observations are *statistically independent*. Let us determine the probability of obtaining the outcome X in the first system *and* the outcome Y in the second system, which we shall denote $P(X \otimes Y)$. In order to determine this probability, we have to form an ensemble of all of the possible pairs of systems which we could choose from the ensemble Σ. Let us denote this ensemble $\Sigma \otimes \Sigma$. The number of pairs of systems in this new ensemble is just the square of the number of systems in the original ensemble, so

$$\Omega(\Sigma \otimes \Sigma) = \Omega(\Sigma)\,\Omega(\Sigma). \tag{2.6}$$

Furthermore, the number of pairs of systems in the ensemble $\Sigma \otimes \Sigma$ which exhibit the outcome X in the first system and Y in the second system is simply the product of the number of systems which exhibit the outcome X and the number of systems which exhibit the outcome Y in the original ensemble, so that

$$\Omega(X \otimes Y) = \Omega(X)\,\Omega(Y). \tag{2.7}$$

It follows from the basic definition of probability that

$$P(X \otimes Y) = \lim_{\Omega(\Sigma)\to\infty} \frac{\Omega(X \otimes Y)}{\Omega(\Sigma \otimes \Sigma)} = P(X)\, P(Y). \tag{2.8}$$

Thus, the probability of obtaining the outcomes X *and* Y in two statistically independent observations is the *product* of the individual probabilities of X and Y. For instance, the probability of throwing a one and then a two on a six-sided die is $1/6 \times 1/6$, which equals $1/36$.

2.4 Mean, Variance, and Standard Deviation

What is meant by the mean or average of a quantity? Well, suppose that we wished to calculate the average age of undergraduates at the University of Texas at Austin. We could go to the central administration building and find out how many eighteen year-olds, nineteen year-olds, *etc.* were currently enrolled. We would then write something like

$$\text{Average Age} \simeq \frac{N_{18} \times 18 + N_{19} \times 19 + N_{20} \times 20 + \cdots}{N_{18} + N_{19} + N_{20} \cdots}, \tag{2.9}$$

where N_{18} is the number of enrolled eighteen year-olds, *etc*. Suppose that we were to pick a student *at random* and then ask “What is the probability of this student being eighteen?” From what we have already discussed, this probability is defined

$$P_{18} \simeq \frac{N_{18}}{N_{\rm students}}, \tag{2.10}$$

where $N_{\rm students}$ is the total number of enrolled students. (Actually, this definition is only accurate in the limit that $N_{\rm students}$ is very large.) We can now see that the average age takes the form

$$\text{Average Age} \simeq P_{18} \times 18 + P_{19} \times 19 + P_{20} \times 20 + \cdots. \tag{2.11}$$

Well, there is nothing special about the age distribution of students at UT Austin. So, for a general variable u, which can take on any one of M possible values $u_1, u_2, \cdots, u_M$, with corresponding probabilities $P(u_1), P(u_2), \cdots, P(u_M)$, the mean or average value of u, which is denoted $\langle u \rangle$, is defined as

$$\langle u \rangle \equiv \sum_{i=1}^{M} P(u_i)\, u_i. \tag{2.12}$$

Suppose that $f(u)$ is some function of u. Then, for each of the M possible values of u there is a corresponding value of $f(u)$ which occurs with the same probability. Thus, $f(u_1)$ corresponds to u_1 and occurs with the probability $P(u_1)$, and so on. It follows from our previous definition that the mean value of $f(u)$ is given by

$$\langle f(u) \rangle \equiv \sum_{i=1}^{M} P(u_i)\, f(u_i). \tag{2.13}$$

Suppose that $f(u)$ and $g(u)$ are two general functions of u. It follows that

$$\langle f(u) + g(u)\rangle = \sum_{i=1}^{M} P(u_i)\,[f(u_i) + g(u_i)] = \sum_{i=1}^{M} P(u_i)\,f(u_i) + \sum_{i=1}^{M} P(u_i)\,g(u_i), \tag{2.14}$$

so

$$\langle f(u) + g(u)\rangle = \langle f(u)\rangle + \langle g(u)\rangle. \tag{2.15}$$

Finally, if c is a general constant then

$$\langle c\, f(u)\rangle = c\,\langle f(u)\rangle. \tag{2.16}$$

We now know how to define the mean value of the general variable u. But, how can we characterize the scatter around the mean value? We could investigate the deviation of u from its mean value $\langle u\rangle$, which is denoted

$$\Delta u \equiv u - \langle u\rangle. \tag{2.17}$$

In fact, this is not a particularly interesting quantity, since its average is zero:

$$\langle \Delta u\rangle = \langle (u - \langle u\rangle)\rangle = \langle u\rangle - \langle u\rangle = 0. \tag{2.18}$$

This is another way of saying that the average deviation from the mean vanishes. A more interesting quantity is the square of the deviation. The average value of this quantity,

$$\left\langle (\Delta u)^2\right\rangle = \sum_{i=1}^{M} P(u_i)\,(u_i - \langle u\rangle)^2, \tag{2.19}$$

is usually called the *variance*. The variance is a *positive* number, unless there is no scatter at all in the distribution, so that all possible values of u correspond to the mean value $\langle u\rangle$, in which case it is *zero*. The following general relation is often useful

$$\left\langle (u - \langle u\rangle)^2\right\rangle = \left\langle (u^2 - 2\,u\,\langle u\rangle + \langle u\rangle^2)\right\rangle = \left\langle u^2\right\rangle - 2\,\langle u\rangle\,\langle u\rangle + \langle u\rangle^2, \tag{2.20}$$

giving

$$\left\langle (\Delta u)^2\right\rangle = \left\langle u^2\right\rangle - \langle u\rangle^2. \tag{2.21}$$

The variance of u is proportional to the square of the scatter of u around its mean value. A more useful measure of the scatter is given by the square root of the variance,

$$\sigma_u = \left[\left\langle (\Delta u)^2\right\rangle\right]^{1/2}, \tag{2.22}$$

which is usually called the *standard deviation* of u. The standard deviation is essentially the width of the range over which u is distributed around its mean value $\langle u\rangle$.

2.5 Continuous Probability Distributions

Suppose, now, that the variable u can take on a *continuous* range of possible values. In general, we expect the probability that u takes on a value in the range u to $u + du$ to be *directly proportional* to du, in the limit that $du \to 0$. In other words,

$$P(u \in u : u + du) = P(u)\, du, \tag{2.23}$$

where $P(u)$ is known as the *probability density*. The earlier results (2.5), (2.12), and (2.19) generalize in a straightforward manner to give

$$1 = \int_{-\infty}^{\infty} P(u)\, du, \tag{2.24}$$

$$\langle u \rangle = \int_{-\infty}^{\infty} P(u)\, u\, du, \tag{2.25}$$

$$\left\langle (\Delta u)^2 \right\rangle = \int_{-\infty}^{\infty} P(u)\, (u - \langle u \rangle)^2\, du = \left\langle u^2 \right\rangle - \langle u \rangle^2, \tag{2.26}$$

respectively.

Exercises

1. In the “game” of Russian roulette, the player inserts a single cartridge into the drum of a revolver, leaving the other five chambers of the drum empty. The player then spins the drum, aims at his/her head, and pulls the trigger.
   (a) What is the probability of the player still being alive after playing the game N times?
   (b) What is the probability of the player surviving $N - 1$ turns in this game, and then being shot the Nth time he/she pulls the trigger?
   (c) What is the mean number of times the player gets to pull the trigger?

2. Suppose that the probability density for the speed s of a car on a road is given by

$$P(s) = A\, s\, \exp\left(-\frac{s}{s_0}\right),$$

   where $0 \leq s \leq \infty$. Here, A and s_0 are positive constants. More explicitly, $P(s)\, ds$ gives the probability that a car has a speed between s and $s + ds$.
   (a) Determine A in terms of s_0.
   (b) What is the mean value of the speed?
   (c) What is the “most probable” speed: *i.e.*, the speed for which the probability density has a maximum?
   (d) What is the probability that a car has a speed more than three times as large as the mean value?

3. An radioactive atom has a uniform decay probability per unit time w: *i.e.*, the probability of decay in a time interval dt is $w\,dt$. Let $P(t)$ be the probability of the atom not having decayed at time t, given that it was created at time $t = 0$. Demonstrate that

$$P(t) = e^{-w\,t}.$$

   What is the mean lifetime of the atom?

3 Wave-Particle Duality

3.1 Introduction

In classical mechanics, waves and particles are two completely distinct types of physical entity. Waves are continuous and spatially extended, whereas particles are discrete and have little or no spatial extent. However, in quantum mechanics, waves sometimes act as particles, and particles sometimes act as waves—this strange behaviour is known as *wave-particle duality*. In this chapter, we shall examine how wave-particle duality shapes the general features of quantum mechanics.

3.2 Wavefunctions

A *wave* is defined as a disturbance in some physical system which is *periodic* in both space and time. In one dimension, a wave is generally represented in terms of a *wavefunction*: *e.g.*,

$$\psi(x,t) = A\,\cos(k\,x - \omega\,t + \varphi), \tag{3.1}$$

where x represents position, t represents time, and A, k, $\omega > 0$. For instance, if we are considering a sound wave then $\psi(x,t)$ might correspond to the pressure perturbation associated with the wave at position x and time t. On the other hand, if we are considering a light wave then $\psi(x,t)$ might represent the wave's transverse electric field. As is well-known, the cosine function, $\cos(\theta)$, is *periodic* in its argument, θ, with period 2π: *i.e.*, $\cos(\theta + 2\pi) = \cos\theta$ for all θ. The function also *oscillates* between the minimum and maximum values -1 and $+1$, respectively, as θ varies. It follows that the wavefunction (3.1) is periodic in x with period $\lambda = 2\pi/k$: *i.e.*, $\psi(x+\lambda,t) = \psi(x,t)$ for all x and t. Moreover, the wavefunction is periodic in t with period $T = 2\pi/\omega$: *i.e.*, $\psi(x,t+T) = \psi(x,t)$ for all x and t. Finally, the wavefunction oscillates between the minimum and maximum values $-A$ and $+A$, respectively, as x and t vary. The spatial period of the wave, λ, is known as its *wavelength*, and the temporal period, T, is called its *period*. Furthermore, the quantity A is termed the wave *amplitude*, the quantity k the *wavenumber*, and the quantity ω the wave *angular frequency*. Note that the units of ω are *radians per second*. The conventional wave *frequency*, in *cycles per second* (otherwise known as hertz), is $\nu = 1/T = \omega/2\pi$. Finally, the quantity φ, appearing in expression (3.1), is termed the *phase angle*, and determines the exact positions of the wave maxima and minima at a given time. In fact, the maxima are located at $k\,x - \omega\,t + \varphi = j\,2\pi$, where j is an integer. This follows because the maxima of $\cos(\theta)$ occur at $\theta = j\,2\pi$. Note that a given maximum satisfies $x = (j - \varphi/2\pi)\,\lambda + v\,t$, where $v = \omega/k$. It follows that the maximum, and, by implication, the whole wave, *propagates* in the *positive* x-direction at the velocity ω/k. Analogous

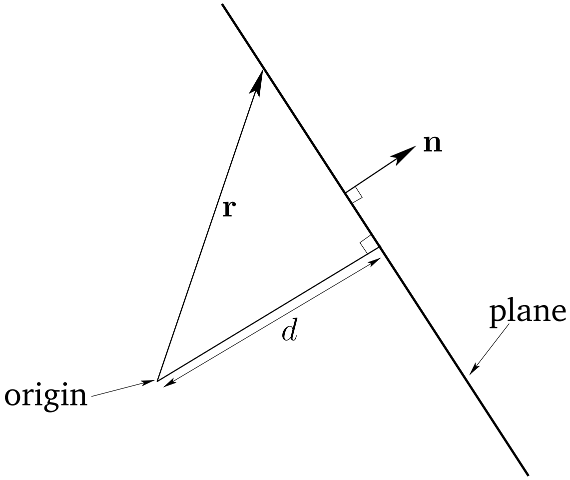

Figure 3.1: *The solution of* $\mathbf{n}\cdot\mathbf{r} = d$ *is a plane.*

reasoning reveals that

$$\psi(x,t) = A\,\cos(-k\,x - \omega\,t + \varphi) = A\,\cos(k\,x + \omega\,t - \varphi), \tag{3.2}$$

is the wavefunction of a wave of amplitude A, wavenumber k, angular frequency ω, and phase angle φ, which propagates in the *negative* x-direction at the velocity ω/k.

3.3 Plane Waves

As we have just seen, a wave of amplitude A, wavenumber k, angular frequency ω, and phase angle φ, propagating in the positive x-direction, is represented by the following wavefunction:

$$\psi(x,t) = A\,\cos(k\,x - \omega\,t + \varphi). \tag{3.3}$$

Now, the type of wave represented above is conventionally termed a *one-dimensional plane wave*. It is *one-dimensional* because its associated wavefunction only depends on the single Cartesian coordinate x. Furthermore, it is a *plane wave* because the wave maxima, which are located at

$$k\,x - \omega\,t + \varphi = j\,2\pi, \tag{3.4}$$

where j is an integer, consist of a series of *parallel planes*, normal to the x-axis, which are equally spaced a distance $\lambda = 2\pi/k$ apart, and propagate along the positive x-axis at the velocity $v = \omega/k$. These conclusions follow because Eq. (3.4) can be re-written in the form

$$x = d, \tag{3.5}$$

where $d = (j - \varphi/2\pi)\,\lambda + v\,t$. Moreover, as is well-known, (3.5) is the equation of a plane, normal to the x-axis, whose distance of closest approach to the origin is d.

The previous equation can also be written in the coordinate-free form

$$\mathbf{n}\cdot\mathbf{r} = d, \tag{3.6}$$

where $\mathbf{n} = (1,\,0,\,0)$ is a unit vector directed along the positive x-axis, and $\mathbf{r} = (x,\,y,\,z)$ represents the vector displacement of a general point from the origin. Since there is nothing special about the x-direction, it follows that if $\mathbf{n}$ is re-interpreted as a unit vector pointing in an *arbitrary* direction then (3.6) can be re-interpreted as the general equation of a plane. As before, the plane is normal to $\mathbf{n}$, and its distance of closest approach to the origin is d. See Fig. 3.1. This observation allows us to write the three-dimensional equivalent to the wavefunction (3.3) as

$$\psi(x,y,z,t) = A\,\cos(\mathbf{k}\cdot\mathbf{r} - \omega\,t + \varphi), \tag{3.7}$$

where the constant vector $\mathbf{k} = (k_x,\,k_y,\,k_z) = k\,\mathbf{n}$ is called the *wavevector*. The wave represented above is conventionally termed a *three-dimensional plane wave*. It is three-dimensional because its wavefunction, $\psi(x,y,z,t)$, depends on all three Cartesian coordinates. Moreover, it is a plane wave because the wave maxima are located at

$$\mathbf{k}\cdot\mathbf{r} - \omega\,t + \varphi = j\,2\pi, \tag{3.8}$$

or

$$\mathbf{n}\cdot\mathbf{r} = (j - \varphi/2\pi)\,\lambda + v\,t, \tag{3.9}$$

where $\lambda = 2\pi/k$, and $v = \omega/k$. Note that the wavenumber, k, is the *magnitude* of the wavevector, $\mathbf{k}$: *i.e.*, $k \equiv |\mathbf{k}|$. It follows, by comparison with Eq. (3.6), that the wave maxima consist of a series of parallel planes, normal to the wavevector, which are equally spaced a distance λ apart, and which propagate in the $\mathbf{k}$-direction at the velocity v. See Fig. 3.2. Hence, the direction of the wavevector specifies the wave propagation direction, whereas its magnitude determines the wavenumber, k, and, thus, the wavelength, $\lambda = 2\pi/k$.

3.4 Representation of Waves via Complex Functions

In mathematics, the symbol i is conventionally used to represent the *square-root of minus one*: *i.e.*, one of the solutions of $i^2 = -1$. Now, a *real number*, x (say), can take any value in a continuum of different values lying between $-\infty$ and $+\infty$. On the other hand, an *imaginary number* takes the general form $i\,y$, where y is a real number. It follows that the square of a real number is a positive real number, whereas the square of an imaginary number is a negative real number. In addition, a general *complex number* is written

$$z = x + i\,y, \tag{3.10}$$

where x and y are real numbers. In fact, x is termed the *real part* of z, and y the *imaginary part* of z. This is written mathematically as $x = \mathrm{Re}(z)$ and $y = \mathrm{Im}(z)$. Finally, the *complex conjugate* of z is defined $z^* = x - i\,y$.

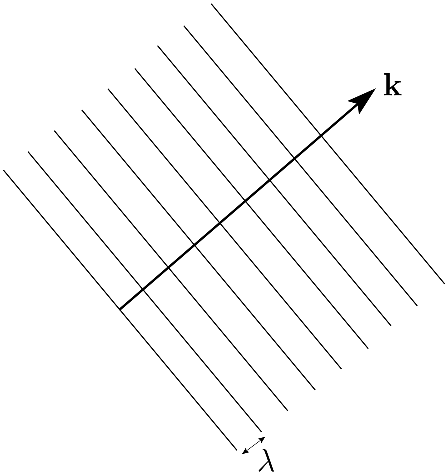

Figure 3.2: *Wave maxima associated with a three-dimensional plane wave.*

Now, just as we can visualize a real number as a point on an infinite straight-line, we can visualize a complex number as a point in an infinite plane. The coordinates of the point in question are the real and imaginary parts of the number: *i.e.*, $z \equiv (x, y)$. This idea is illustrated in Fig. 3.3. The distance, $r = \sqrt{x^2 + y^2}$, of the representative point from the origin is termed the *modulus* of the corresponding complex number, z. This is written mathematically as $|z| = \sqrt{x^2 + y^2}$. Incidentally, it follows that $z\,z^* = x^2 + y^2 = |z|^2$. The angle, $\theta = \tan^{-1}(y/x)$, that the straight-line joining the representative point to the origin subtends with the real axis is termed the *argument* of the corresponding complex number, z. This is written mathematically as $\arg(z) = \tan^{-1}(y/x)$. It follows from standard trigonometry that $x = r\,\cos\theta$, and $y = r\,\sin\theta$. Hence, $z = r\,\cos\theta + i\,r\sin\theta$.

Complex numbers are often used to represent wavefunctions. All such representations depend ultimately on a fundamental mathematical identity, known as *de Moivre's theorem*, which takes the form

$$e^{i\,\phi} \equiv \cos\phi + i\,\sin\phi, \tag{3.11}$$

where ϕ is a real number. Incidentally, given that $z = r\,\cos\theta + i\,r\,\sin\theta = r\,(\cos\theta + i\,\sin\theta)$, where z is a general complex number, $r = |z|$ its modulus, and $\theta = \arg(z)$ its argument, it follows from de Moivre's theorem that any complex number, z, can be written

$$z = r\,e^{i\,\theta}, \tag{3.12}$$

where $r = |z|$ and $\theta = \arg(z)$ are real numbers.

Now, a one-dimensional wavefunction takes the general form

$$\psi(x, t) = A\,\cos(k\,x - \omega\,t + \varphi), \tag{3.13}$$

where A is the wave amplitude, k the wavenumber, ω the angular frequency, and φ the phase angle. Consider the complex wavefunction

$$\psi(x, t) = \psi_0\,e^{\,i\,(k\,x - \omega\,t)}, \tag{3.14}$$

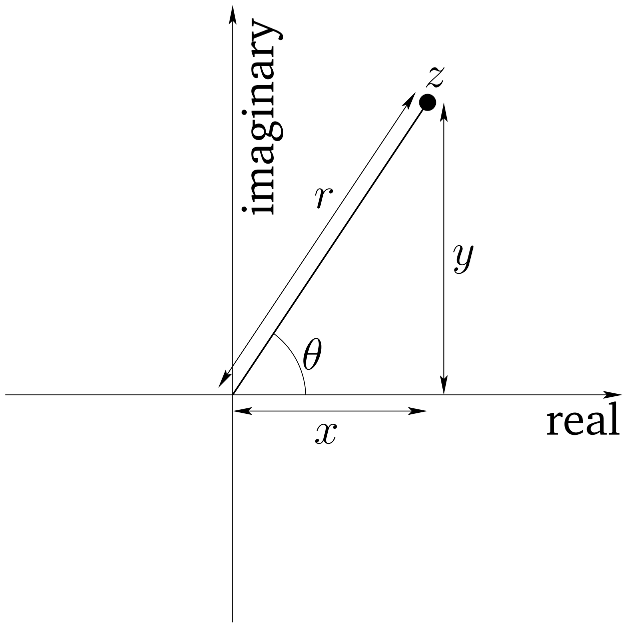

Figure 3.3: *Representation of a complex number as a point in a plane.*

where ψ_0 is a complex constant. We can write

$$\psi_0 = A\,e^{\,i\,\varphi}, \tag{3.15}$$

where A is the modulus, and φ the argument, of ψ_0. Hence, we deduce that

$$\begin{aligned} \text{Re}\left[\psi_0\,e^{\,i\,(k\,x-\omega\,t)}\right] &= \text{Re}\left[A\,e^{\,i\,\varphi}\,e^{\,i\,(k\,x-\omega\,t)}\right] \\ &= \text{Re}\left[A\,e^{\,i\,(k\,x-\omega\,t+\varphi)}\right] \\ &= A\,\text{Re}\left[e^{\,i\,(k\,x-\omega\,t+\varphi)}\right]. \end{aligned} \tag{3.16}$$

Thus, it follows from de Moirve's theorem, and Eq. (3.13), that

$$\text{Re}\left[\psi_0\,e^{\,i\,(k\,x-\omega\,t)}\right] = A\,\cos(k\,x-\omega\,t+\varphi) = \psi(x,t). \tag{3.17}$$

In other words, a general one-dimensional real wavefunction, (3.13), can be represented as the *real part* of a complex wavefunction of the form (3.14). For ease of notation, the "take the real part" aspect of the above expression is usually omitted, and our general one-dimension wavefunction is simply written

$$\psi(x,t) = \psi_0\,e^{\,i\,(k\,x-\omega\,t)}. \tag{3.18}$$

The main advantage of the complex representation, (3.18), over the more straightforward real representation, (3.13), is that the former enables us to combine the amplitude, A, and the phase angle, φ, of the wavefunction into a single complex amplitude, ψ_0. Finally, the three dimensional generalization of the above expression is

$$\psi(\mathbf{r}, t) = \psi_0\, e^{\,i\,(\mathbf{k}\cdot\mathbf{r}-\omega\,t)}, \tag{3.19}$$

where $\mathbf{k}$ is the wavevector.

3.5 Classical Light Waves

Consider a classical, monochromatic, linearly polarized, plane light wave, propagating through a vacuum in the x-direction. It is convenient to characterize a light wave (which is, of course, a type of electromagnetic wave) by specifying its associated electric field. Suppose that the wave is polarized such that this electric field oscillates in the y-direction. (According to standard electromagnetic theory, the magnetic field oscillates in the z-direction, in phase with the electric field, with an amplitude which is that of the electric field divided by the velocity of light in vacuum.) Now, the electric field can be conveniently represented in terms of a *complex wavefunction*:

$$\psi(x, t) = \bar{\psi}\, e^{\,i\,(k\,x-\omega\,t)}. \tag{3.20}$$

Here, $i = \sqrt{-1}$, k and ω are real parameters, and $\bar{\psi}$ is a complex wave amplitude. By convention, the physical electric field is the *real part* of the above expression. Suppose that

$$\bar{\psi} = |\bar{\psi}|\, e^{\,i\,\varphi}, \tag{3.21}$$

where φ is real. It follows that the physical electric field takes the form

$$E_y(x, t) = \mathrm{Re}[\psi(x, t)] = |\bar{\psi}|\, \cos(k\,x - \omega\,t + \varphi), \tag{3.22}$$

where $|\bar{\psi}|$ is the amplitude of the electric oscillation, k the wavenumber, ω the angular frequency, and φ the phase angle. In addition, $\lambda = 2\pi/k$ is the wavelength, and $\nu = \omega/2\pi$ the frequency (in hertz).

According to standard electromagnetic theory, the frequency and wavelength of light waves are related according to the well-known expression

$$c = \nu\,\lambda, \tag{3.23}$$

or, equivalently,

$$\omega = k\,c, \tag{3.24}$$

where $c = 3 \times 10^8\,\mathrm{m/s}$. Equations (3.22) and (3.24) yield

$$E_y(x, t) = |\bar{\psi}|\, \cos\left(k\,[x - (\omega/k)\,t] + \varphi\right) = |\bar{\psi}|\, \cos\left(k\,[x - c\,t] + \varphi\right). \tag{3.25}$$

Note that E_y depends on x and t only via the combination $x - c\,t$. It follows that the wave maxima and minima satisfy

$$x - c\,t = \text{constant}. \tag{3.26}$$

Thus, the wave maxima and minima propagate in the x-direction at the fixed velocity

$$\frac{dx}{dt} = c. \tag{3.27}$$

An expression, such as (3.24), which determines the wave angular frequency as a function of the wavenumber, is generally termed a *dispersion relation*. As we have already seen, and as is apparent from Eq. (3.25), the maxima and minima of a plane wave propagate at the characteristic velocity

$$v_p = \frac{\omega}{k}, \tag{3.28}$$

which is known as the *phase velocity*. Hence, the dispersion relation (3.24) is effectively saying that the phase velocity of a plane light wave propagating through a vacuum always takes the fixed value c, irrespective of its wavelength or frequency.

Now, from standard electromagnetic theory, the *energy density* (*i.e.*, the energy per unit volume) of a light wave is

$$U = \frac{E_y^{\,2}}{\epsilon_0}, \tag{3.29}$$

where $\epsilon_0 = 8.85 \times 10^{-12}\,\text{F/m}$ is the *permittivity of free space*. Hence, it follows from Eqs. (3.20) and (3.22) that

$$U \propto |\psi|^2. \tag{3.30}$$

Furthermore, a light wave possesses linear *momentum*, as well as energy. This momentum is directed along the wave's direction of propagation, and is of density

$$G = \frac{U}{c}. \tag{3.31}$$

3.6 Photoelectric Effect

The so-called *photoelectric effect*, by which a polished metal surface emits electrons when illuminated by visible and ultra-violet light, was discovered by Heinrich Hertz in 1887. The following facts regarding this effect can be established via careful observation. First, a given surface only emits electrons when the *frequency* of the light with which it is illuminated exceeds a certain threshold value, which is a property of the metal. Second, the current of photoelectrons, when it exists, is proportional to the *intensity* of the light falling on the surface. Third, the energy of the photoelectrons is independent of the light intensity, but varies *linearly* with the light frequency. These facts are inexplicable within the framework of classical physics.

In 1905, Albert Einstein proposed a radical new theory of light in order to account for the photoelectric effect. According to this theory, light of fixed frequency ν consists of a collection of indivisible discrete packages, called *quanta*,[1] whose energy is

$$E = h\nu. \tag{3.32}$$

---

[1] Plural of *quantum*: Latin neuter of *quantus*: how much?

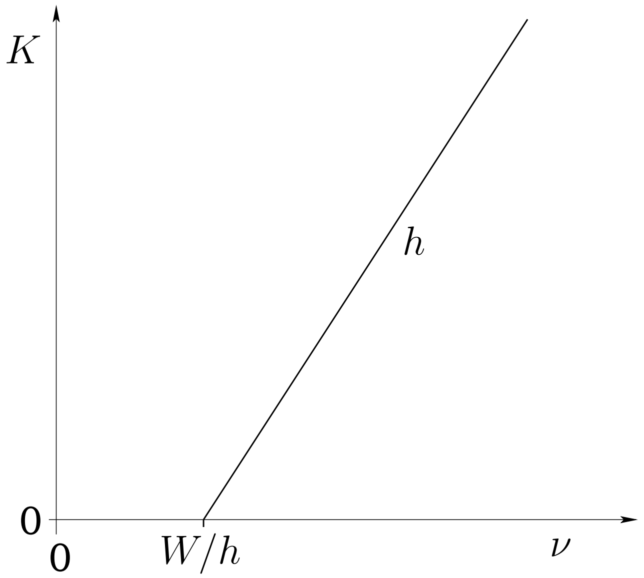

Figure 3.4: *Variation of the kinetic energy* K *of photoelectrons with the wave-frequency* ν.

Here, $h = 6.6261 \times 10^{-34}\,\mathrm{J\,s}$ is a new constant of nature, known as *Planck's constant*. Incidentally, h is called Planck's constant, rather than Einstein's constant, because Max Planck first introduced the concept of the quantization of light, in 1900, whilst trying to account for the electromagnetic spectrum of a black body (*i.e.*, a perfect emitter and absorber of electromagnetic radiation).

Suppose that the electrons at the surface of a metal lie in a potential well of depth W. In other words, the electrons have to acquire an energy W in order to be emitted from the surface. Here, W is generally called the *work function* of the surface, and is a property of the metal. Suppose that an electron absorbs a single quantum of light. Its energy therefore increases by $h\nu$. If $h\nu$ is greater than W then the electron is emitted from the surface with residual kinetic energy

$$K = h\nu - W. \tag{3.33}$$

Otherwise, the electron remains trapped in the potential well, and is not emitted. Here, we are assuming that the probability of an electron simultaneously absorbing two or more light quanta is negligibly small compared to the probability of it absorbing a single light quantum (as is, indeed, the case for low intensity illumination). Incidentally, we can calculate Planck's constant, and the work function of the metal, by simply plotting the kinetic energy of the emitted photoelectrons as a function of the wave frequency, as shown in Fig. 3.4. This plot is a straight-line whose slope is h, and whose intercept with the ν axis is W/h. Finally, the number of emitted electrons increases with the intensity of the light because the more intense the light the larger the flux of light quanta onto the surface. Thus, Einstein's quantum theory is capable of accounting for all three of the previously mentioned observational facts regarding the photoelectric effect.

3.7 Quantum Theory of Light

According to Einstein's quantum theory of light, a monochromatic light wave of angular frequency ω, propagating through a vacuum, can be thought of as a stream of particles, called *photons*, of energy

$$E = \hbar\,\omega, \tag{3.34}$$

where $\hbar = h/2\pi = 1.0546 \times 10^{-34}\,\mathrm{J\,s}$. Since classical light waves propagate at the fixed velocity c, it stands to reason that photons must also move at this velocity. Now, according to Einstein's special theory of relativity, only massless particles can move at the speed of light in vacuum. Hence, photons must be *massless*. Special relativity also gives the following relationship between the energy E and the momentum p of a massless particle,

$$p = \frac{E}{c}. \tag{3.35}$$

Note that the above relation is consistent with Eq. (3.31), since if light is made up of a stream of photons, for which $E/p = c$, then the momentum density of light must be the energy density divided by c. It follows from the previous two equations that photons carry momentum

$$p = \hbar\,k \tag{3.36}$$

along their direction of motion, since $\omega/c = k$ for a light wave [see Eq. (3.24)].

3.8 Classical Interference of Light Waves

Let us now consider the classical interference of light waves. Figure 3.5 shows a standard double-slit interference experiment in which monochromatic plane light waves are normally incident on two narrow parallel slits which are a distance d apart. The light from the two slits is projected onto a screen a distance D behind them, where $D \gg d$.

Consider some point on the screen which is located a distance y from the centre-line, as shown in the figure. Light from the first slit travels a distance x_1 to get to this point, whereas light from the second slit travels a slightly different distance x_2. It is easily demonstrated that

$$\Delta x = x_2 - x_1 \simeq \frac{d}{D}\,y, \tag{3.37}$$

provided $d \ll D$. It follows from Eq. (3.20), and the well-known fact that light waves are *superposible*, that the wavefunction at the point in question can be written

$$\psi(y,t) \propto \psi_1(t)\,e^{\,i\,k\,x_1} + \psi_2(t)\,e^{\,i\,k\,x_2}, \tag{3.38}$$

where ψ_1 and ψ_2 are the wavefunctions at the first and second slits, respectively. However,

$$\psi_1 = \psi_2, \tag{3.39}$$

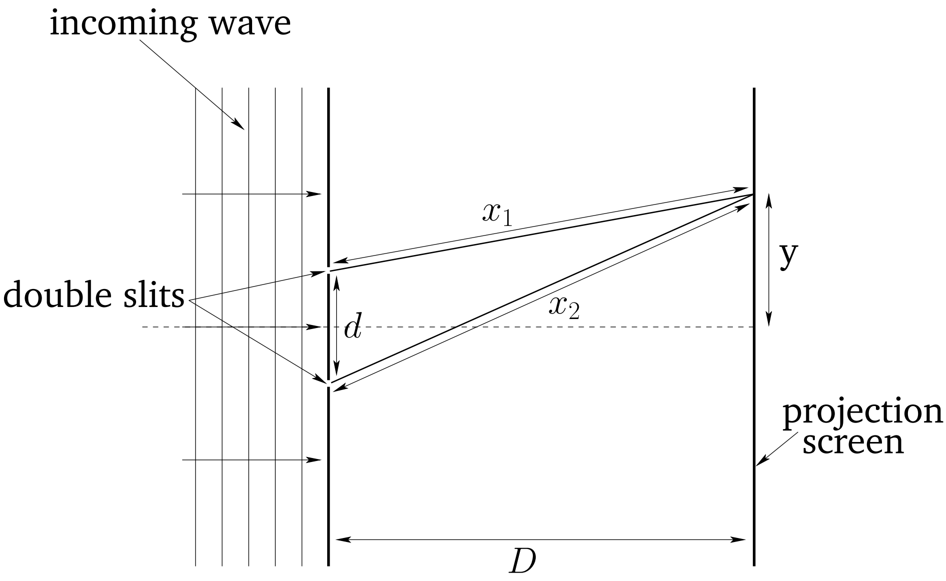

Figure 3.5: *Classical double-slit interference of light.*

since the two slits are assumed to be illuminated by in-phase light waves of equal amplitude. (Note that we are ignoring the difference in amplitude of the waves from the two slits at the screen, due to the slight difference between x_1 and x_2, compared to the difference in their phases. This is reasonable provided $D \gg \lambda$.) Now, the intensity (*i.e.*, the energy flux) of the light at some point on the projection screen is approximately equal to the energy density of the light at this point times the velocity of light (provided that $y \ll D$). Hence, it follows from Eq. (3.30) that the light intensity on the screen a distance y from the center-line is

$$I(y) \propto |\psi(y,t)|^2. \tag{3.40}$$

Using Eqs. (3.37)–(3.40), we obtain

$$I(y) \propto \cos^2\left(\frac{k\,\Delta x}{2}\right) \simeq \cos^2\left(\frac{k\,d}{2\,D}\,y\right). \tag{3.41}$$

Figure 3.6 shows the characteristic interference pattern corresponding to the above expression. This pattern consists of equally spaced light and dark bands of characteristic width

$$\Delta y = \frac{D\,\lambda}{d}. \tag{3.42}$$

3.9 Quantum Interference of Light

Let us now consider double-slit light interference from a quantum mechanical point of view. According to quantum theory, light waves consist of a stream of massless photons

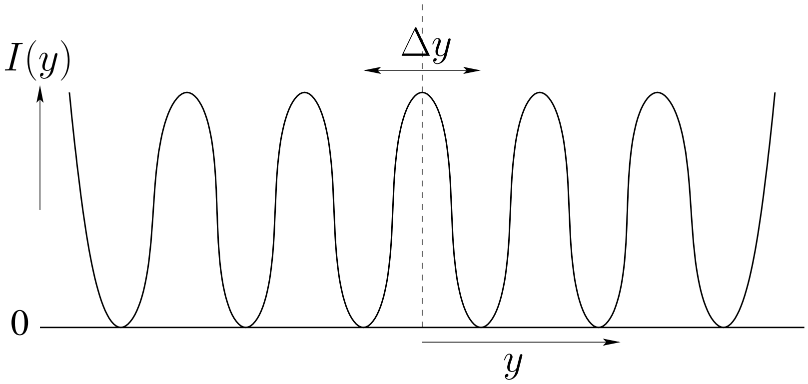

Figure 3.6: *Classical double-slit interference pattern.*

moving at the speed of light. Hence, we expect the two slits in Fig. 3.5 to be spraying photons in all directions at the same rate. Suppose, however, that we reduce the intensity of the light source illuminating the slits until the source is so weak that only a *single* photon is present between the slits and the projection screen at any given time. Let us also replace the projection screen by a photographic film which records the position where it is struck by each photon. So, if we wait a sufficiently long time that a great many photons have passed through the slits and struck the photographic film, and then develop the film, do we see an interference pattern which looks like that shown in Fig. 3.6? The answer to this question, as determined by experiment, is that we see *exactly* the same interference pattern.

Now, according to the above discussion, the interference pattern is built up one photon at a time: *i.e.*, the pattern is not due to the interaction of different photons. Moreover, the point at which a given photon strikes the film is not influenced by the points at which previous photons struck the film, given that there is only one photon in the apparatus at any given time. Hence, the only way in which the classical interference pattern can be reconstructed, after a great many photons have passed through the apparatus, is if each photon has a greater *probability* of striking the film at points where the classical interference pattern is bright, and a lesser probability of striking the film at points where the interference pattern is dark.

Suppose, then, that we allow N photons to pass through our apparatus, and then count the number of photons which strike the recording film between y and $y + \Delta y$, where Δy is a relatively small division. Let us call this number $n(y)$. Now, the number of photons which strike a region of the film in a given time interval is equivalent to the intensity of the light illuminating that region of the film multiplied by the area of the region, since each photon carries a fixed amount of energy. Hence, in order to reconcile the classical and quantum viewpoints, we need

$$P_y(y) \equiv \lim_{N\to\infty}\left[\frac{n(y)}{N}\right] \propto I(y)\,\Delta y, \tag{3.43}$$

where $I(y)$ is given in Eq. (3.41). Here, $P_y(y)$ is the *probability* that a given photon strikes the film between y and $y + \Delta y$. This probability is simply a number between 0 and 1. A probability of 0 means that there is no chance of a photon striking the film between y and $y + \Delta y$, whereas a probability of 1 means that every photon is certain to strike the film in this interval. Note that $P_y \propto \Delta y$. In other words, the probability of a photon striking a region of the film of width Δy is directly proportional to this width. Actually, this is only true as long as Δy is relatively small. It is convenient to define a quantity known as the *probability density*, $P(y)$, which is such that the probability of a photon striking a region of the film of infinitesimal width dy is $P_y(y) = P(y)\, dy$. Now, Eq. (3.43) yields $P_y(y) \propto I(y)\, dy$, which gives $P(y) \propto I(y)$. However, according to Eq. (3.40), $I(y) \propto |\psi(y,t)|^2$. Thus, we obtain

$$P(y) \propto |\psi(y,t)|^2. \tag{3.44}$$

In other words, the probability density of a photon striking a given point on the film is proportional to the *modulus squared* of the wavefunction at that point. Another way of saying this is that the probability of a measurement of the photon's distance from the centerline, at the location of the film, yielding a result between y and $y+dy$ is proportional to $|\psi(y,t)|^2\, dy$.

Note that, in the quantum mechanical picture, we can only predict the *probability* that a given photon strikes a given point on the film. If photons behaved classically then we could, in principle, solve their equations of motion and predict *exactly* where each photon was going to strike the film, given its initial position and velocity. This loss of determinancy in quantum mechanics is a direct consequence of *wave-particle duality*. In other words, we can only reconcile the wave-like and particle-like properties of light in a *statistical* sense. It is impossible to reconcile them on the individual particle level.

In principle, each photon which passes through our apparatus is equally likely to pass through one of the two slits. So, can we determine which slit a given photon passed through? Well, suppose that our original interference experiment involves sending $N \gg 1$ photons through our apparatus. We know that we get an interference pattern in this experiment. Suppose that we perform a modified interference experiment in which we close off one slit, send $N/2$ photons through the apparatus, and then open the slit and close off the other slit, and send $N/2$ photons through the apparatus. In this second experiment, which is virtually identical to the first on the individual photon level, we know exactly which slit each photon passed through. However, the wave theory of light (which we expect to agree with the quantum theory in the limit $N \gg 1$) tells us that our modified interference experiment will *not* result in the formation of an interference pattern. After all, according to wave theory, it is impossible to obtain a two-slit interference pattern from a single slit. Hence, we conclude that any attempt to measure which slit each photon in our two-slit interference experiment passes through results in the destruction of the interference pattern. It follows that, in the quantum mechanical version of the two-slit interference experiment, we must think of each photon as essentially passing through *both* slits simultaneously.

3.10 Classical Particles

In this course, we are going to concentrate, almost exclusively, on the behaviour of *non-relativistic* particles of *non-zero mass* (*e.g.*, electrons). In the absence of external forces, such particles, of mass m, energy E, and momentum p, move classically in a straight-line with velocity

$$v = \frac{p}{m}, \tag{3.45}$$

and satisfy

$$E = \frac{p^2}{2\,m}. \tag{3.46}$$

3.11 Quantum Particles

Just as light waves sometimes exhibit particle-like properties, it turns out that massive particles sometimes exhibit wave-like properties. For instance, it is possible to obtain a double-slit interference pattern from a stream of mono-energetic electrons passing through two closely spaced narrow slits. Now, the effective wavelength of the electrons can be determined by measuring the width of the light and dark bands in the interference pattern [see Eq. (3.42)]. It is found that

$$\lambda = \frac{h}{p}. \tag{3.47}$$

The same relation is found for other types of particles. The above wavelength is called the *de Broglie wavelength*, after Louis de Broglie who first suggested that particles should have wave-like properties in 1923. Note that the de Broglie wavelength is generally pretty small. For instance, that of an electron is

$$\lambda_e = 1.2\times 10^{-9}\,[E(\mathrm{eV})]^{-1/2}\,\mathrm{m}, \tag{3.48}$$

where the electron energy is conveniently measured in units of electron-volts (eV). (An electron accelerated from rest through a potential difference of 1000 V acquires an energy of 1000 eV, and so on.) The de Broglie wavelength of a proton is

$$\lambda_p = 2.9\times 10^{-11}\,[E(\mathrm{eV})]^{-1/2}\,\mathrm{m}. \tag{3.49}$$

Given the smallness of the de Broglie wavelengths of common particles, it is actually quite difficult to do particle interference experiments. In general, in order to perform an effective interference experiment, the spacing of the slits must not be too much greater than the wavelength of the wave. Hence, particle interference experiments require either very low energy particles (since $\lambda \propto E^{-1/2}$), or very closely spaced slits. Usually the “slits” consist of crystals, which act a bit like diffraction gratings with a characteristic spacing of order the inter-atomic spacing (which is generally about 10^{-9} m).

Equation (3.47) can be rearranged to give

$$p = \hbar\,k, \tag{3.50}$$

which is exactly the same as the relation between momentum and wavenumber that we obtained earlier for photons [see Eq. (3.36)]. For the case of a particle moving the three dimensions, the above relation generalizes to give

$$\mathbf{p} = \hbar\,\mathbf{k}, \tag{3.51}$$

where **p** is the particle's vector momentum, and **k** its wavevector. It follows that the momentum of a quantum particle, and, hence, its velocity, is always parallel to its wavevector.

Since the relation (3.36) between momentum and wavenumber applies to both photons and massive particles, it seems plausible that the closely related relation (3.34) between energy and wave angular frequency should also apply to both photons and particles. If this is the case, and we can write

$$E = \hbar\,\omega \tag{3.52}$$

for particle waves, then Eqs. (3.46) and (3.50) yield the following dispersion relation for such waves:

$$\omega = \frac{\hbar\,k^2}{2\,m}. \tag{3.53}$$

Now, we saw earlier that a plane wave propagates at the so-called phase velocity,

$$v_p = \frac{\omega}{k}. \tag{3.54}$$

However, according to the above dispersion relation, a particle plane wave propagates at

$$v_p = \frac{p}{2\,m}. \tag{3.55}$$

Note, from Eq. (3.45), that this is only *half* of the classical particle velocity. Does this imply that the dispersion relation (3.53) is incorrect? Let us investigate further.

3.12 Wave Packets

The above discussion suggests that the wavefunction of a massive particle of momentum p and energy E, moving in the positive x-direction, can be written

$$\psi(x,t) = \bar{\psi}\,e^{\,i\,(k\,x-\omega\,t)}, \tag{3.56}$$

where $k = p/\hbar > 0$ and $\omega = E/\hbar > 0$. Here, ω and k are linked via the dispersion relation (3.53). Expression (3.56) represents a plane wave whose maxima and minima propagate in the positive x-direction with the phase velocity $v_p = \omega/k$. As we have seen, this phase velocity is only half of the classical velocity of a massive particle.

From before, the most reasonable physical interpretation of the wavefunction is that $|\psi(x,t)|^2$ is proportional to the *probability density* of finding the particle at position x at time t. However, the modulus squared of the wavefunction (3.56) is $|\bar{\psi}|^2$, which depends on neither x nor t. In other words, this wavefunction represents a particle which is equally likely to be found anywhere on the x-axis at all times. Hence, the fact that the maxima and minima of the wavefunction propagate at a phase velocity which does not correspond to the classical particle velocity does not have any real physical consequences.

So, how can we write the wavefunction of a particle which is *localized* in x: *i.e.*, a particle which is more likely to be found at some positions on the x-axis than at others? It turns out that we can achieve this goal by forming a *linear combination* of plane waves of different wavenumbers: *i.e.*,

$$\psi(x,t) = \int_{-\infty}^{\infty} \bar{\psi}(k)\, e^{\,i\,(k\,x-\omega\,t)}\, dk. \tag{3.57}$$

Here, $\bar{\psi}(k)$ represents the complex amplitude of plane waves of wavenumber k in this combination. In writing the above expression, we are relying on the assumption that particle waves are *superposable*: *i.e.*, it is possible to add two valid wave solutions to form a third valid wave solution. The ultimate justification for this assumption is that particle waves satisfy a differential wave equation which is *linear* in ψ. As we shall see, in Sect. 3.15, this is indeed the case. Incidentally, a plane wave which varies as $\exp[i\,(k\,x-\omega\,t)]$ and has a *negative* k (but positive ω) propagates in the *negative* x-direction at the phase velocity $\omega/|k|$. Hence, the superposition (3.57) includes both forward and backward propagating waves.

Now, there is a useful mathematical theorem, known as *Fourier's theorem*, which states that if

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \bar{f}(k)\, e^{\,i\,k\,x}\, dk, \tag{3.58}$$

then

$$\bar{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)\, e^{-i\,k\,x}\, dx. \tag{3.59}$$

Here, $\bar{f}(k)$ is known as the *Fourier transform* of the function $f(x)$. We can use Fourier's theorem to find the k-space function $\bar{\psi}(k)$ which generates any given x-space wavefunction $\psi(x)$ at a given time.

For instance, suppose that at $t=0$ the wavefunction of our particle takes the form

$$\psi(x,0) \propto \exp\left[i\,k_0\,x - \frac{(x-x_0)^{\,2}}{4\,(\Delta x)^{\,2}}\right]. \tag{3.60}$$

Thus, the initial probability density of the particle is written

$$|\psi(x,0)|^{\,2} \propto \exp\left[-\frac{(x-x_0)^{\,2}}{2\,(\Delta x)^{\,2}}\right]. \tag{3.61}$$

This particular probability distribution is called a *Gaussian* distribution, and is plotted in Fig. 3.7. It can be seen that a measurement of the particle's position is most likely to yield the value x_0, and very unlikely to yield a value which differs from x_0 by more than $3\,\Delta x$. Thus, (3.60) is the wavefunction of a particle which is initially localized around $x=x_0$ in

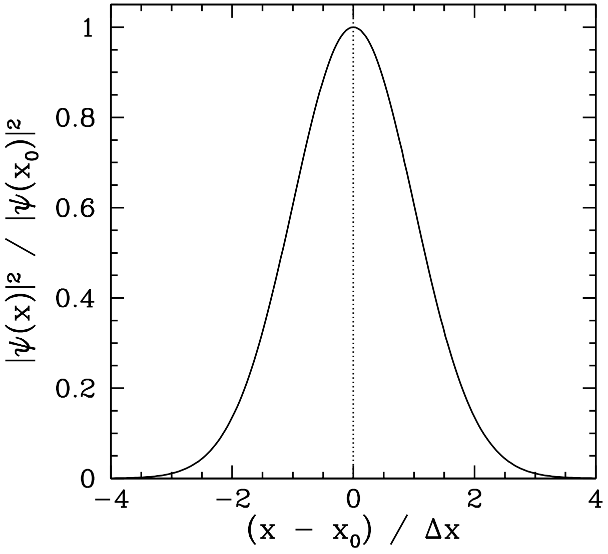

Figure 3.7: *A Gaussian probability distribution in x-space.*

some region whose width is of order Δx. This type of wavefunction is known as a *wave packet*.

Now, according to Eq. (3.57),

$$\psi(x,0) = \int_{-\infty}^{\infty} \bar{\psi}(k)\, e^{\,i\,k\,x}\, dk. \tag{3.62}$$

Hence, we can employ Fourier's theorem to invert this expression to give

$$\bar{\psi}(k) \propto \int_{-\infty}^{\infty} \psi(x,0)\, e^{-i\,k\,x}\, dx. \tag{3.63}$$

Making use of Eq. (3.60), we obtain

$$\bar{\psi}(k) \propto e^{-i\,(k-k_0)\,x_0} \int_{-\infty}^{\infty} \exp\left[-i\,(k-k_0)\,(x-x_0) - \frac{(x-x_0)^2}{4\,(\Delta x)^2}\right] dx. \tag{3.64}$$

Changing the variable of integration to $y = (x - x_0)/(2\,\Delta x)$, this reduces to

$$\bar{\psi}(k) \propto e^{-i\,k\,x_0} \int_{-\infty}^{\infty} \exp\left[-i\,\beta\,y - y^2\right] dy, \tag{3.65}$$

where $\beta = 2\,(k - k_0)\,\Delta x$. The above equation can be rearranged to give

$$\bar{\psi}(k) \propto e^{-i\,k\,x_0-\beta^2/4} \int_{-\infty}^{\infty} e^{-(y-y_0)^2}\, dy, \tag{3.66}$$

where $y_0 = -i\,\beta/2$. The integral now just reduces to a number, as can easily be seen by making the change of variable $z = y - y_0$. Hence, we obtain

$$\bar{\psi}(k) \propto \exp\left[-i\,k\,x_0 - \frac{(k-k_0)^2}{4\,(\Delta k)^2}\right], \tag{3.67}$$

where

$$\Delta k = \frac{1}{2\,\Delta x}. \tag{3.68}$$

Now, if $|\psi(x)|^2$ is proportional to the probability density of a measurement of the particle's position yielding the value x then it stands to reason that $|\bar{\psi}(k)|^2$ is proportional to the probability density of a measurement of the particle's wavenumber yielding the value k. (Recall that $p = \hbar\,k$, so a measurement of the particle's wavenumber, k, is equivalent to a measurement of the particle's momentum, p). According to Eq. (3.67),

$$|\bar{\psi}(k)|^2 \propto \exp\left[-\frac{(k-k_0)^2}{2\,(\Delta k)^2}\right]. \tag{3.69}$$

Note that this probability distribution is a *Gaussian* in k-space. See Eq. (3.61) and Fig. 3.7. Hence, a measurement of k is most likely to yield the value k_0, and very unlikely to yield a value which differs from k_0 by more than $3\,\Delta k$. Incidentally, a Gaussian is the *only* mathematical function in x-space which has the same form as its Fourier transform in k-space.

We have just seen that a Gaussian probability distribution of characteristic width Δx in x-space [see Eq. (3.61)] transforms to a Gaussian probability distribution of characteristic width Δk in k-space [see Eq. (3.69)], where

$$\Delta x\,\Delta k = \frac{1}{2}. \tag{3.70}$$

This illustrates an important property of wave packets. Namely, if we wish to construct a packet which is very localized in x-space (*i.e.*, if Δx is small) then we need to combine plane waves with a very wide range of different k-values (*i.e.*, Δk will be large). Conversely, if we only combine plane waves whose wavenumbers differ by a small amount (*i.e.*, if Δk is small) then the resulting wave packet will be very extended in x-space (*i.e.*, Δx will be large).

3.13 Evolution of Wave Packets

We have seen, in Eq. (3.60), how to write the wavefunction of a particle which is initially localized in x-space. But, how does this wavefunction evolve in time? Well, according to Eq. (3.57), we have

$$\psi(x,t) = \int_{-\infty}^{\infty} \bar{\psi}(k)\,e^{\,i\,\phi(k)}\,dk, \tag{3.71}$$

where

$$\phi(k) = k\,x - \omega(k)\,t. \tag{3.72}$$

The function $\bar{\psi}(k)$ is obtained by Fourier transforming the wavefunction at $t = 0$. See Eqs. (3.63) and (3.67). Now, according to Eq. (3.69), $|\bar{\psi}(k)|$ is *strongly peaked* around $k = k_0$. Thus, it is a reasonable approximation to Taylor expand $\phi(k)$ about k_0. Keeping terms up to second-order in $k - k_0$, we obtain

$$\psi(x,t) \propto \int_{-\infty}^{\infty} \bar{\psi}(k)\, \exp\left[\,i\left\{\phi_0 + \phi_0'\,(k-k_0) + \frac{1}{2}\,\phi_0''\,(k-k_0)^{\,2}\right\}\right], \tag{3.73}$$

where

$$\begin{aligned} \phi_0 &= \phi(k_0) = k_0\,x - \omega_0\,t, && (3.74)\\ \phi_0' &= \frac{d\phi(k_0)}{dk} = x - v_g\,t, && (3.75)\\ \phi_0'' &= \frac{d^2\phi(k_0)}{dk^2} = -\alpha\,t, && (3.76) \end{aligned}$$

with

$$\begin{aligned} \omega_0 &= \omega(k_0), && (3.77)\\ v_g &= \frac{d\omega(k_0)}{dk}, && (3.78)\\ \alpha &= \frac{d^2\omega(k_0)}{dk^2}. && (3.79) \end{aligned}$$

Substituting from Eq. (3.67), rearranging, and then changing the variable of integration to $y = (k - k_0)/(2\,\Delta k)$, we get

$$\psi(x,t) \propto e^{\,i\,(k_0\,x - \omega_0\,t)} \int_{-\infty}^{\infty} e^{\,i\,\beta_1\,y - (1+i\,\beta_2)\,y^2}\,dy, \tag{3.80}$$

where

$$\begin{aligned} \beta_1 &= 2\,\Delta k\,(x - x_0 - v_g\,t), && (3.81)\\ \beta_2 &= 2\,\alpha\,(\Delta k)^{\,2}\,t. && (3.82) \end{aligned}$$

Incidentally, $\Delta k = 1/(2\,\Delta x)$, where Δx is the initial width of the wave packet. The above expression can be rearranged to give

$$\psi(x,t) \propto e^{\,i\,(k_0\,x - \omega_0\,t) - (1+i\,\beta_2)\,\beta^{\,2}/4} \int_{-\infty}^{\infty} e^{-(1+i\,\beta_2)\,(y-y_0)^{\,2}}\,dy, \tag{3.83}$$

where $y_0 = i\,\beta/2$ and $\beta = \beta_1/(1 + i\,\beta_2)$. Again changing the variable of integration to $z = (1 + i\,\beta_2)^{1/2}\,(y - y_0)$, we get

$$\psi(x,t) \propto (1 + i\,\beta_2)^{-1/2}\,e^{\,i\,(k_0\,x - \omega_0\,t) - (1 + i\,\beta_2)\,\beta^2/4} \int_{-\infty}^{\infty} e^{-z^2}\,dz. \tag{3.84}$$

The integral now just reduces to a number. Hence, we obtain

$$\psi(x,t) \propto \frac{\exp\left[\,i\,(k_0\,x - \omega_0\,t) - (x - x_0 - v_g\,t)^2\{1 - i\,2\,\alpha\,(\Delta k)^2\,t\}/(4\,\sigma^2)\right]}{[1 + i\,2\,\alpha\,(\Delta k)^2\,t]^{1/2}}, \tag{3.85}$$

where

$$\sigma^2(t) = (\Delta x)^2 + \frac{\alpha^2\,t^2}{4\,(\Delta x)^2}. \tag{3.86}$$

Note that the above wavefunction is identical to our original wavefunction (3.60) at $t = 0$. This, justifies the approximation which we made earlier by Taylor expanding the phase factor $\phi(k)$ about $k = k_0$.

According to Eq. (3.85), the probability density of our particle as a function of time is written

$$|\psi(x,t)|^2 \propto \sigma^{-1}(t)\,\exp\left[-\frac{(x - x_0 - v_g\,t)^2}{2\,\sigma^2(t)}\right]. \tag{3.87}$$

Hence, the probability distribution is a Gaussian, of characteristic width σ, which peaks at $x = x_0 + v_g\,t$. Now, the most likely position of our particle coincides with the peak of the distribution function. Thus, the particle's most likely position is given by

$$x = x_0 + v_g\,t. \tag{3.88}$$

It can be seen that the particle effectively moves at the uniform velocity

$$v_g = \frac{d\omega}{dk}, \tag{3.89}$$

which is known as the *group velocity*. In other words, a plane wave travels at the phase velocity, $v_p = \omega/k$, whereas a wave packet travels at the group velocity, $v_g = d\omega/dt$. Now, it follows from the dispersion relation (3.53) for particle waves that

$$v_g = \frac{p}{m}. \tag{3.90}$$

However, it can be seen from Eq. (3.45) that this is identical to the classical particle velocity. Hence, the dispersion relation (3.53) turns out to be consistent with classical physics, after all, as soon as we realize that individual particles must be identified with *wave packets* rather than plane waves. In fact, a plane wave is usually interpreted as a *continuous stream* of particles propagating in the same direction as the wave.

According to Eq. (3.86), the width of our wave packet grows as time progresses. Indeed, it follows from Eqs. (3.53) and (3.79) that the characteristic time for a wave packet of original width Δx to double in spatial extent is

$$t_2 \sim \frac{m\,(\Delta x)^2}{\hbar}. \tag{3.91}$$

For instance, if an electron is originally localized in a region of atomic scale (*i.e.*, $\Delta x \sim 10^{-10}\,\mathrm{m}$) then the doubling time is only about $10^{-16}\,\mathrm{s}$. Evidently, particle wave packets (for freely moving particles) spread very rapidly.

Note, from the previous analysis, that the rate of spreading of a wave packet is ultimately governed by the second derivative of $\omega(k)$ with respect to k. See Eqs. (3.79) and (3.86). This is why a functional relationship between ω and k is generally known as a *dispersion relation*: *i.e.*, because it governs how wave packets disperse as time progresses. However, for the special case where ω is a *linear* function of k, the second derivative of ω with respect to k is zero, and, hence, there is no dispersion of wave packets: *i.e.*, wave packets propagate without changing shape. Now, the dispersion relation (3.24) for light waves is linear in k. It follows that light pulses propagate through a vacuum without spreading. Another property of linear dispersion relations is that the phase velocity, $v_p = \omega/k$, and the group velocity, $v_g = d\omega/dk$, are *identical*. Thus, both plane light waves and light pulses propagate through a vacuum at the characteristic speed $c = 3 \times 10^8\,\mathrm{m/s}$. Of course, the dispersion relation (3.53) for particle waves is *not* linear in k. Hence, particle plane waves and particle wave packets propagate at different velocities, and particle wave packets also gradually disperse as time progresses.

3.14 Heisenberg's Uncertainty Principle

According to the analysis contained in the previous two sections, a particle wave packet which is initially localized in x-space with characteristic width Δx is also localized in k-space with characteristic width $\Delta k = 1/(2\,\Delta x)$. However, as time progresses, the width of the wave packet in x-space increases, whilst that of the wave packet in k-space stays the same. [After all, our previous analysis obtained $\psi(x,t)$ from Eq. (3.71), but assumed that $\bar{\psi}(k)$ was given by Eq. (3.67) at all times.] Hence, in general, we can say that

$$\Delta x\,\Delta k \gtrsim \frac{1}{2}. \tag{3.92}$$

Furthermore, we can think of Δx and Δk as characterizing our *uncertainty* regarding the values of the particle's position and wavenumber, respectively.

Now, a measurement of a particle's wavenumber, k, is equivalent to a measurement of its momentum, p, since $p = \hbar\,k$. Hence, an uncertainty in k of order Δk translates to an uncertainty in p of order $\Delta p = \hbar\,\Delta k$. It follows from the above inequality that

$$\Delta x\,\Delta p \gtrsim \frac{\hbar}{2}. \tag{3.93}$$

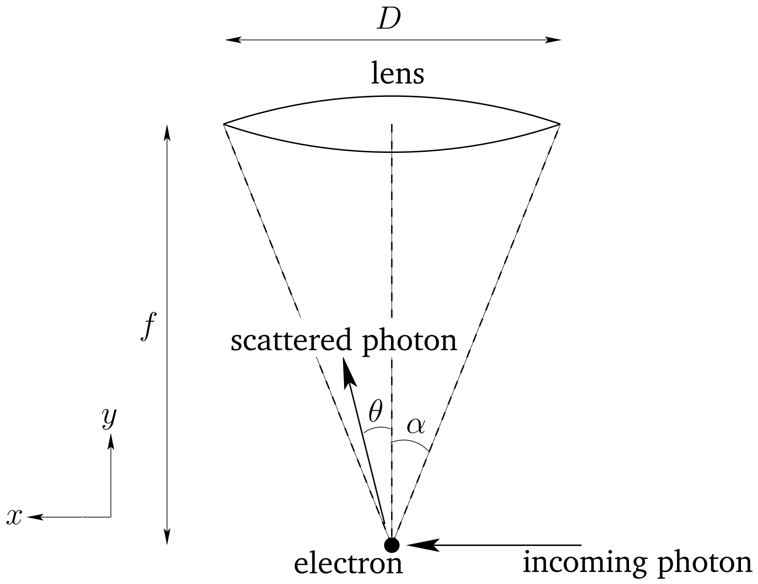

Figure 3.8: *Heisenberg's microscope.*

This is the famous *Heisenberg uncertainty principle*, first proposed by Werner Heisenberg in 1927. According to this principle, it is impossible to simultaneously measure the position and momentum of a particle (exactly). Indeed, a good knowledge of the particle's position implies a poor knowledge of its momentum, and *vice versa*. Note that the uncertainty principle is a direct consequence of representing particles as waves.

It can be seen from Eqs. (3.53), (3.79), and (3.86) that at large t a particle wavefunction of original width Δx (at $t = 0$) spreads out such that its spatial extent becomes

$$\sigma \sim \frac{\hbar\, t}{m\,\Delta x}. \tag{3.94}$$

It is easily demonstrated that this spreading is a consequence of the uncertainty principle. Since the initial uncertainty in the particle's position is Δx, it follows that the uncertainty in its momentum is of order $\hbar/\Delta x$. This translates to an uncertainty in velocity of $\Delta v = \hbar/(m\,\Delta x)$. Thus, if we imagine that parts of the wavefunction propagate at $v_0 + \Delta v/2$, and others at $v_0 - \Delta v/2$, where v_0 is the mean propagation velocity, then the wavefunction will spread as time progresses. Indeed, at large t we expect the width of the wavefunction to be

$$\sigma \sim \Delta v\, t \sim \frac{\hbar\, t}{m\,\Delta x}, \tag{3.95}$$

which is identical to Eq. (3.94). Evidently, the spreading of a particle wavefunction must be interpreted as an increase in our *uncertainty* regarding the particle's position, rather than an increase in the spatial extent of the particle itself.

Figure 3.8 illustrates a famous thought experiment known as *Heisenberg's microscope*. Suppose that we try to image an electron using a simple optical system in which the objective lens is of diameter D and focal-length f. (In practice, this would only be possible using extremely short wavelength light.) It is a well-known result in optics that such a system has a minimum angular resolving power of λ/D, where λ is the wavelength of the light illuminating the electron. If the electron is placed at the focus of the lens, which is where the minimum resolving power is achieved, then this translates to a uncertainty in the electron's transverse position of

$$\Delta x \simeq f\,\frac{\lambda}{D}. \tag{3.96}$$

However,

$$\tan\alpha = \frac{D/2}{f}, \tag{3.97}$$

where α is the half-angle subtended by the lens at the electron. Assuming that α is small, we can write

$$\alpha \simeq \frac{D}{2\,f}, \tag{3.98}$$

so

$$\Delta x \simeq \frac{\lambda}{2\,\alpha}. \tag{3.99}$$

It follows that we can reduce the uncertainty in the electron's position by *minimizing* the ratio λ/α: *i.e.*, by using short wavelength radiation, and a wide-angle lens.

Let us now examine Heisenberg's microscope from a quantum mechanical point of view. According to quantum mechanics, the electron is imaged when it scatters an incoming photon towards the objective lens. Let the wavevector of the incoming photon have the (x, y) components $(k, 0)$. See Fig. 3.8. If the scattered photon subtends an angle θ with the center-line of the optical system, as shown in the figure, then its wavevector is written $(k\,\sin\theta, k\cos\theta)$. Here, we are ignoring any wavelength shift of the photon on scattering—*i.e.*, the magnitude of the **k**-vector is assumed to be the same before and after scattering. Thus, the change in the x-component of the photon's wavevector is $\Delta k_x = k\,(\sin\theta - 1)$. This translates to a change in the photon's x-component of momentum of $\Delta p_x = \hbar\,k\,(\sin\theta - 1)$. By momentum conservation, the electron's x-momentum will change by an equal and opposite amount. However, θ can range all the way from $-\alpha$ to $+\alpha$, and the scattered photon will still be collected by the imaging system. It follows that the uncertainty in the electron's momentum is

$$\Delta p \simeq 2\,\hbar\,k\,\sin\alpha \simeq \frac{4\pi\,\hbar\,\alpha}{\lambda}. \tag{3.100}$$

Note that in order to reduce the uncertainty in the momentum we need to *maximize* the ratio λ/α. This is exactly the opposite of what we need to do to reduce the uncertainty in the position. Multiplying the previous two equations, we obtain

$$\Delta x\,\Delta p \sim h, \tag{3.101}$$

which is essentially the uncertainty principle.

According to Heisenberg's microscope, the uncertainty principle follows from two facts. First, it is impossible to measure any property of a microscopic dynamical system without *disturbing* the system somewhat. Second, particle and light energy and momentum are *quantized*. Hence, there is a limit to how small we can make the aforementioned disturbance. Thus, there is an irreducible uncertainty in certain measurements which is a consequence of the act of measurement itself.

3.15 Schrödinger's Equation

We have seen that the wavefunction of a free particle of mass m satisfies

$$\psi(x,t) = \int_{-\infty}^{\infty} \bar{\psi}(k)\, e^{\,i\,(k\,x-\omega\,t)}\, dk, \tag{3.102}$$

where $\bar{\psi}(k)$ is determined by $\psi(x,0)$, and

$$\omega(k) = \frac{\hbar\,k^2}{2\,m}. \tag{3.103}$$

Now, it follows from Eq. (3.102) that

$$\frac{\partial\psi}{\partial x} = \int_{-\infty}^{\infty} (i\,k)\, \bar{\psi}(k)\, e^{\,i\,(k\,x-\omega\,t)}\, dk, \tag{3.104}$$

and

$$\frac{\partial^2\psi}{\partial x^2} = \int_{-\infty}^{\infty} (-k^2)\, \bar{\psi}(k)\, e^{\,i\,(k\,x-\omega\,t)}\, dk, \tag{3.105}$$

whereas

$$\frac{\partial\psi}{\partial t} = \int_{-\infty}^{\infty} (-i\,\omega)\, \bar{\psi}(k)\, e^{\,i\,(k\,x-\omega\,t)}\, dk. \tag{3.106}$$

Thus,

$$i\,\frac{\partial\psi}{\partial t} + \frac{\hbar}{2\,m}\,\frac{\partial^2\psi}{\partial x^2} = \int_{-\infty}^{\infty} \left(\omega - \frac{\hbar\,k^2}{2\,m}\right) \bar{\psi}(k)\, e^{\,i\,(k\,x-\omega\,t)}\, dk = 0, \tag{3.107}$$

where use has been made of the dispersion relation (3.103). Multiplying through by $\hbar$, we obtain

$$i\,\hbar\,\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2\,m}\frac{\partial^2\psi}{\partial x^2}. \tag{3.108}$$

This expression is known as *Schrödinger's equation*, since it was first introduced by Erwin Schrödinger in 1925. Schrödinger's equation is a *linear*, second-order, partial differential equation which governs the time evolution of a particle wavefunction, and is generally easier to solve than the integral equation (3.102).

Of course, Eq. (3.108) is only applicable to *freely moving* particles. Fortunately, it is fairly easy to guess the generalization of this equation for particles moving in some potential $V(x)$. It is plausible, from Eq. (3.104), that we can identify k with the differential operator $-i\,\partial/\partial x$. Hence, the differential operator on the right-hand side of Eq. (3.108) is equivalent to $\hbar^2 k^2/(2\,m)$. But, $p = \hbar\,k$. Thus, the operator is also equivalent to $p^2/(2\,m)$, which is just the energy of a freely moving particle. However, in the presence of a potential $V(x)$, the particle's energy is written $p^2/(2\,m) + V$. Thus, it seems reasonable to make the substitution

$$-\frac{\hbar^2}{2\,m}\frac{\partial^2}{\partial x^2} \rightarrow -\frac{\hbar^2}{2\,m}\frac{\partial^2}{\partial x^2} + V(x). \tag{3.109}$$

This leads to the general form of Schrödinger's equation:

$$i\,\hbar\,\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2\,m}\frac{\partial^2\psi}{\partial x^2} + V(x)\,\psi. \tag{3.110}$$

3.16 Collapse of the Wave Function

Consider an extended wavefunction $\psi(x,t)$. According to our usual interpretation, $|\psi(x,t)|^2$ is proportional to the probability density of a measurement of the particle's position yielding the value x at time t. If the wavefunction is extended then there is a wide range of likely values that this measurement could give. Suppose that we make such a measurement, and obtain the value x_0. We now know that the particle is located at $x = x_0$. If we make another measurement immediately after the first one then what value do we expect to obtain? Well, common sense tells us that we must obtain the same value, x_0, since the particle cannot have shifted position appreciably in an infinitesimal time interval. Thus, immediately after the first measurement, a measurement of the particle's position is certain to give the value x_0, and has no chance of giving any other value. This implies that the wavefunction must have *collapsed* to some sort of "spike" function located at $x = x_0$. This is illustrated in Fig. 3.9. Of course, as soon as the wavefunction has collapsed, it starts to expand again, as discussed in Sect. 3.13. Thus, the second measurement must be made reasonably quickly after the first, in order to guarantee that the same result will be obtained.

The above discussion illustrates an important point in quantum mechanics. Namely, that the wavefunction of a particle changes *discontinuously* (in time) whenever a measurement is made. We conclude that there are two types of time evolution of the wavefunction in quantum mechanics. First, there is a *smooth* evolution which is governed by Schrödinger's equation. This evolution takes place *between* measurements. Second, there is a *discontinuous* evolution which takes place each time a measurement is made.

Exercises

1. A He-Ne laser emits radiation of wavelength $\lambda = 633$ nm. How many photons are emitted per second by a laser with a power of 1 mW? What force does such laser exert on a body which completely absorbs its radiation?

2. The ionization energy of a hydrogen atom in its ground state is $E_{ion} = 13.60$ eV (1 eV is the energy acquired by an electron accelerated through a potential difference of 1 V). Calculate

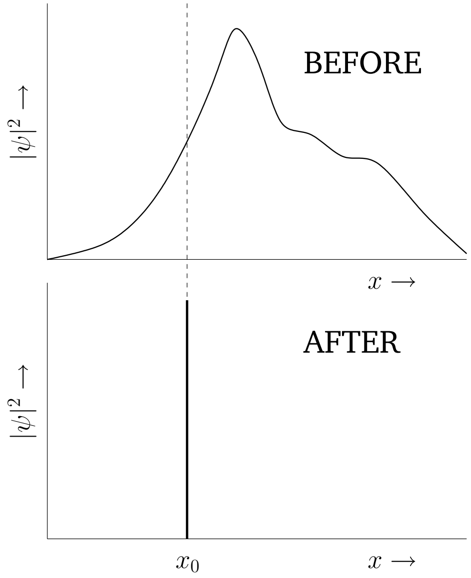

Figure 3.9: *Collapse of the wavefunction upon measurement of* x.

the frequency, wavelength, and wavenumber of the electromagnetic radiation which will just ionize the atom.

3. The maximum energy of photoelectrons from aluminium is 2.3 eV for radiation of wavelength 2000 Å, and 0.90 eV for radiation of wavelength 2580 Å. Use this data to calculate Planck's constant, and the work function of aluminium.

4. Show that the de Broglie wavelength of an electron accelerated from rest across a potential difference V is given by
$$\lambda = 1.29 \times 10^{-9}\, V^{-1/2}\,\mathrm{m},$$
where V is measured in volts.

5. If the atoms in a regular crystal are separated by 3×10^{-10} m demonstrate that an accelerating voltage of about 1.5 kV would be required to produce an electron diffraction pattern from the crystal.

6. The relationship between wavelength and frequency for electromagnetic waves in a waveguide is
$$\lambda = \frac{c}{\sqrt{\nu^2 - \nu_0^2}},$$
where c is the velocity of light in vacuum. What are the group and phase velocities of such waves as functions of ν_0 and λ?

7. Nuclei, typically of size 10^{-14} m, frequently emit electrons with energies of 1–10 MeV. Use the uncertainty principle to show that electrons of energy 1 MeV could not be contained in the nucleus before the decay.

8. A particle of mass m has a wavefunction

$$\psi(x,t) = A\ \exp[-a\,(m\,x^2/\hbar + i\,t)],$$

where A and a are positive real constants. For what potential function $V(x)$ does ψ satisfy the Schrödinger equation?

4 Fundamentals of Quantum Mechanics

4.1 Introduction

The previous chapter serves as a useful introduction to many of the basic concepts of quantum mechanics. In this chapter, we shall examine these concepts in a more systematic fashion. For the sake of simplicity, we shall concentrate on one-dimensional systems.

4.2 Schrödinger's Equation

Consider a dynamical system consisting of a single non-relativistic particle of mass m moving along the x-axis in some real potential $V(x)$. In quantum mechanics, the instantaneous state of the system is represented by a complex wavefunction $\psi(x,t)$. This wavefunction evolves in time according to Schrödinger's equation:

$$i\,\hbar\,\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2\,m}\frac{\partial^2\psi}{\partial x^2} + V(x)\,\psi. \tag{4.1}$$

The wavefunction is interpreted as follows: $|\psi(x,t)|^2$ is the probability density of a measurement of the particle's displacement yielding the value x. Thus, the probability of a measurement of the displacement giving a result between a and b (where $a < b$) is

$$P_{x\,\in\,a:b}(t) = \int_a^b |\psi(x,t)|^2\,dx. \tag{4.2}$$

Note that this quantity is real and positive definite.

4.3 Normalization of the Wavefunction

Now, a probability is a real number between 0 and 1. An outcome of a measurement which has a probability 0 is an impossible outcome, whereas an outcome which has a probability 1 is a certain outcome. According to Eq. (4.2), the probability of a measurement of x yielding a result between $-\infty$ and $+\infty$ is

$$P_{x\,\in\,-\infty:\infty}(t) = \int_{-\infty}^{\infty} |\psi(x,t)|^2\,dx. \tag{4.3}$$

However, a measurement of x *must* yield a value between $-\infty$ and $+\infty$, since the particle has to be located somewhere. It follows that $P_{x\,\in\,-\infty:\infty} = 1$, or

$$\int_{-\infty}^{\infty} |\psi(x,t)|^2\,dx = 1, \tag{4.4}$$

which is generally known as the *normalization condition* for the wavefunction.

For example, suppose that we wish to normalize the wavefunction of a Gaussian wave packet, centered on $x = x_0$, and of characteristic width σ (see Sect. 3.12): *i.e.*,

$$\psi(x) = \psi_0\, e^{-(x-x_0)^2/(4\,\sigma^2)}. \tag{4.5}$$

In order to determine the normalization constant ψ_0, we simply substitute Eq. (4.5) into Eq. (4.4), to obtain

$$|\psi_0|^2 \int_{-\infty}^{\infty} e^{-(x-x_0)^2/(2\,\sigma^2)}\, dx = 1. \tag{4.6}$$

Changing the variable of integration to $y = (x - x_0)/(\sqrt{2}\,\sigma)$, we get

$$|\psi_0|^2 \sqrt{2}\,\sigma \int_{-\infty}^{\infty} e^{-y^2}\, dy = 1. \tag{4.7}$$

However,

$$\int_{-\infty}^{\infty} e^{-y^2}\, dy = \sqrt{\pi}, \tag{4.8}$$

which implies that

$$|\psi_0|^2 = \frac{1}{(2\pi\,\sigma^2)^{1/2}}. \tag{4.9}$$

Hence, a general normalized Gaussian wavefunction takes the form

$$\psi(x) = \frac{e^{\,i\,\varphi}}{(2\pi\,\sigma^2)^{1/4}}\, e^{-(x-x_0)^2/(4\,\sigma^2)}, \tag{4.10}$$

where φ is an arbitrary real phase-angle.

Now, it is important to demonstrate that if a wavefunction is initially normalized then it stays normalized as it evolves in time according to Schrödinger's equation. If this is not the case then the probability interpretation of the wavefunction is untenable, since it does not make sense for the probability that a measurement of x yields *any* possible outcome (which is, manifestly, unity) to change in time. Hence, we require that

$$\frac{d}{dt} \int_{-\infty}^{\infty} |\psi(x,t)|^2\, dx = 0, \tag{4.11}$$

for wavefunctions satisfying Schrödinger's equation. The above equation gives

$$\frac{d}{dt} \int_{-\infty}^{\infty} \psi^*\, \psi\, dx = \int_{-\infty}^{\infty} \left(\frac{\partial \psi^*}{\partial t}\, \psi + \psi^*\, \frac{\partial \psi}{\partial t} \right) dx = 0. \tag{4.12}$$

Now, multiplying Schrödinger's equation by $\psi^*/(i\,\hbar)$, we obtain

$$\psi^*\, \frac{\partial \psi}{\partial t} = \frac{i\,\hbar}{2\,m}\, \psi^*\, \frac{\partial^2 \psi}{\partial x^2} - \frac{i}{\hbar}\, V\, |\psi|^2. \tag{4.13}$$

The complex conjugate of this expression yields

$$\psi\,\frac{\partial\psi^*}{\partial t} = -\frac{i\,\hbar}{2\,m}\,\psi\,\frac{\partial^2\psi^*}{\partial x^2} + \frac{i}{\hbar}\,V\,|\psi|^2 \tag{4.14}$$

[since $(A\,B)^* = A^*\,B^*$, $A^{**} = A$, and $i^* = -i$]. Summing the previous two equations, we get

$$\frac{\partial\psi^*}{\partial t}\,\psi + \psi^*\,\frac{\partial\psi}{\partial t} = \frac{i\,\hbar}{2\,m}\left(\psi^*\,\frac{\partial^2\psi}{\partial x^2} - \psi\,\frac{\partial^2\psi^*}{\partial x^2}\right) = \frac{i\,\hbar}{2\,m}\,\frac{\partial}{\partial x}\left(\psi^*\,\frac{\partial\psi}{\partial x} - \psi\,\frac{\partial\psi^*}{\partial x}\right). \tag{4.15}$$

Equations (4.12) and (4.15) can be combined to produce

$$\frac{d}{dt}\int_{-\infty}^{\infty}|\psi|^2\,dx = \frac{i\,\hbar}{2\,m}\left[\psi^*\,\frac{\partial\psi}{\partial x} - \psi\,\frac{\partial\psi^*}{\partial x}\right]_{-\infty}^{\infty} = 0. \tag{4.16}$$

The above equation is satisfied provided

$$|\psi| \rightarrow 0 \quad \text{as} \quad |x| \rightarrow \infty. \tag{4.17}$$

However, this is a necessary condition for the integral on the left-hand side of Eq. (4.4) to converge. Hence, we conclude that all wavefunctions which are *square-integrable* [*i.e.*, are such that the integral in Eq. (4.4) converges] have the property that if the normalization condition (4.4) is satisfied at one instant in time then it is satisfied at all subsequent times.

It is also possible to demonstrate, via very similar analysis to the above, that

$$\frac{dP_{x\in\,a:b}}{dt} + j(b,t) - j(a,t) = 0, \tag{4.18}$$

where $P_{x\in\,a:b}$ is defined in Eq. (4.2), and

$$j(x,t) = \frac{i\,\hbar}{2\,m}\left(\psi\,\frac{\partial\psi^*}{\partial x} - \psi^*\,\frac{\partial\psi}{\partial x}\right) \tag{4.19}$$

is known as the *probability current*. Note that j is real. Equation (4.18) is a *probability conservation equation*. According to this equation, the probability of a measurement of x lying in the interval a to b evolves in time due to the difference between the flux of probability into the interval [*i.e.*, $j(a,t)$], and that out of the interval [*i.e.*, $j(b,t)$]. Here, we are interpreting $j(x,t)$ as the *flux* of probability in the $+x$-direction at position x and time t.

Note, finally, that not all wavefunctions can be normalized according to the scheme set out in Eq. (4.4). For instance, a plane wave wavefunction

$$\psi(x,t) = \psi_0\,e^{\,i\,(k\,x-\omega\,t)} \tag{4.20}$$

is not square-integrable, and, thus, cannot be normalized. For such wavefunctions, the best we can say is that

$$P_{x\in\,a:b}(t) \propto \int_a^b |\psi(x,t)|^2\,dx. \tag{4.21}$$

In the following, all wavefunctions are assumed to be square-integrable and normalized, unless otherwise stated.

4.4 Expectation Values and Variances

We have seen that $|\psi(x,t)|^2$ is the probability density of a measurement of a particle's displacement yielding the value x at time t. Suppose that we made a large number of independent measurements of the displacement on an equally large number of identical quantum systems. In general, measurements made on different systems will yield different results. However, from the definition of probability, the mean of all these results is simply

$$\langle x\rangle = \int_{-\infty}^{\infty} x\,|\psi|^2\,dx. \tag{4.22}$$

Here, $\langle x\rangle$ is called the *expectation value* of x. Similarly the expectation value of any function of x is

$$\langle f(x)\rangle = \int_{-\infty}^{\infty} f(x)\,|\psi|^2\,dx. \tag{4.23}$$

In general, the results of the various different measurements of x will be scattered around the expectation value $\langle x\rangle$. The degree of scatter is parameterized by the quantity

$$\sigma_x^2 = \int_{-\infty}^{\infty} (x-\langle x\rangle)^2\,|\psi|^2\,dx \equiv \langle x^2\rangle - \langle x\rangle^2, \tag{4.24}$$

which is known as the *variance* of x. The square-root of this quantity, σ_x, is called the *standard deviation* of x. We generally expect the results of measurements of x to lie within a few standard deviations of the expectation value.

For instance, consider the normalized Gaussian wave packet [see Eq. (4.10)]

$$\psi(x) = \frac{e^{i\varphi}}{(2\pi\,\sigma^2)^{1/4}}\,e^{-(x-x_0)^2/(4\,\sigma^2)}. \tag{4.25}$$

The expectation value of x associated with this wavefunction is

$$\langle x\rangle = \frac{1}{\sqrt{2\pi\,\sigma^2}}\int_{-\infty}^{\infty} x\,e^{-(x-x_0)^2/(2\,\sigma^2)}\,dx. \tag{4.26}$$

Let $y = (x-x_0)/(\sqrt{2}\,\sigma)$. It follows that

$$\langle x\rangle = \frac{x_0}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-y^2}\,dy + \frac{\sqrt{2}\,\sigma}{\sqrt{\pi}}\int_{-\infty}^{\infty} y\,e^{-y^2}\,dy. \tag{4.27}$$

However, the second integral on the right-hand side is zero, by symmetry. Hence, making use of Eq. (4.8), we obtain

$$\langle x\rangle = x_0. \tag{4.28}$$

Evidently, the expectation value of x for a Gaussian wave packet is equal to the most likely value of x (*i.e.*, the value of x which maximizes $|\psi|^2$).

The variance of x associated with the Gaussian wave packet (4.25) is

$$\sigma_x^2 = \frac{1}{\sqrt{2\pi\,\sigma^2}}\int_{-\infty}^{\infty}(x-x_0)^2\,e^{-(x-x_0)^2/(2\,\sigma^2)}\,dx. \tag{4.29}$$

Let $y = (x-x_0)/(\sqrt{2}\,\sigma)$. It follows that

$$\sigma_x^2 = \frac{2\,\sigma^2}{\sqrt{\pi}}\int_{-\infty}^{\infty} y^2\,e^{-y^2}\,dy. \tag{4.30}$$

However,

$$\int_{-\infty}^{\infty} y^2\,e^{-y^2}\,dy = \frac{\sqrt{\pi}}{2}, \tag{4.31}$$

giving

$$\sigma_x^2 = \sigma^2. \tag{4.32}$$

This result is consistent with our earlier interpretation of σ as a measure of the spatial extent of the wave packet (see Sect. 3.12). It follows that we can rewrite the Gaussian wave packet (4.25) in the convenient form

$$\psi(x) = \frac{e^{\,i\,\varphi}}{(2\pi\,\sigma_x^2)^{1/4}}\,e^{-(x-\langle x\rangle)^2/(4\,\sigma_x^2)}. \tag{4.33}$$

4.5 Ehrenfest's Theorem

A simple way to calculate the expectation value of momentum is to evaluate the time derivative of $\langle x\rangle$, and then multiply by the mass m: *i.e.*,

$$\langle p\rangle = m\,\frac{d\langle x\rangle}{dt} = m\,\frac{d}{dt}\int_{-\infty}^{\infty} x\,|\psi|^2\,dx = m\int_{-\infty}^{\infty} x\,\frac{\partial|\psi|^2}{\partial t}\,dx. \tag{4.34}$$

However, it is easily demonstrated that

$$\frac{\partial|\psi|^2}{\partial t} + \frac{\partial j}{\partial x} = 0 \tag{4.35}$$

[this is just the differential form of Eq. (4.18)], where j is the probability current defined in Eq. (4.19). Thus,

$$\langle p\rangle = -m\int_{-\infty}^{\infty} x\,\frac{\partial j}{\partial x}\,dx = m\int_{-\infty}^{\infty} j\,dx, \tag{4.36}$$

where we have integrated by parts. It follows from Eq. (4.19) that

$$\langle p\rangle = -\frac{i\,\hbar}{2}\int_{-\infty}^{\infty}\left(\psi^*\,\frac{\partial\psi}{\partial x} - \frac{\partial\psi^*}{\partial x}\,\psi\right)dx = -i\,\hbar\int_{-\infty}^{\infty}\psi^*\,\frac{\partial\psi}{\partial x}\,dx, \tag{4.37}$$

where we have again integrated by parts. Hence, the expectation value of the momentum can be written

$$\langle p\rangle = m\,\frac{d\langle x\rangle}{dt} = -i\,\hbar\int_{-\infty}^{\infty}\psi^*\,\frac{\partial\psi}{\partial x}\,dx. \tag{4.38}$$

It follows from the above that

$$\begin{aligned}\frac{d\langle p\rangle}{dt} &= -i\,\hbar\int_{-\infty}^{\infty}\left(\frac{\partial\psi^*}{\partial t}\,\frac{\partial\psi}{\partial x}+\psi^*\,\frac{\partial^2\psi}{\partial t\partial x}\right)dx\\ &= \int_{-\infty}^{\infty}\left[\left(i\,\hbar\,\frac{\partial\psi}{\partial t}\right)^*\frac{\partial\psi}{\partial x}+\frac{\partial\psi^*}{\partial x}\left(i\,\hbar\,\frac{\partial\psi}{\partial t}\right)\right]dx,\end{aligned} \tag{4.39}$$

where we have integrated by parts. Substituting from Schrödinger's equation (4.1), and simplifying, we obtain

$$\frac{d\langle p\rangle}{dt} = \int_{-\infty}^{\infty}\left[-\frac{\hbar^2}{2\,m}\frac{\partial}{\partial x}\left(\frac{\partial\psi^*}{\partial x}\,\frac{\partial\psi}{\partial x}\right)+V(x)\,\frac{\partial|\psi|^2}{\partial x}\right]dx = \int_{-\infty}^{\infty}V(x)\,\frac{\partial|\psi|^2}{\partial x}\,dx. \tag{4.40}$$

Integration by parts yields

$$\frac{d\langle p\rangle}{dt} = -\int_{-\infty}^{\infty}\frac{dV}{dx}\,|\psi|^2\,dx = -\left\langle\frac{dV}{dx}\right\rangle. \tag{4.41}$$

Hence, according to Eqs. (4.34) and (4.41),

$$m\,\frac{d\langle x\rangle}{dt} = \langle p\rangle, \tag{4.42}$$

$$\frac{d\langle p\rangle}{dt} = -\left\langle\frac{dV}{dx}\right\rangle. \tag{4.43}$$

Evidently, the expectation values of displacement and momentum obey time evolution equations which are analogous to those of classical mechanics. This result is known as *Ehrenfest's theorem*.

Suppose that the potential $V(x)$ is *slowly varying*. In this case, we can expand dV/dx as a Taylor series about $\langle x\rangle$. Keeping terms up to second order, we obtain

$$\frac{dV(x)}{dx} = \frac{dV(\langle x\rangle)}{d\langle x\rangle}+\frac{dV^2(\langle x\rangle)}{d\langle x\rangle^2}\,(x-\langle x\rangle)+\frac{1}{2}\,\frac{dV^3(\langle x\rangle)}{d\langle x\rangle^3}\,(x-\langle x\rangle)^2. \tag{4.44}$$

Substitution of the above expansion into Eq. (4.43) yields

$$\frac{d\langle p\rangle}{dt} = -\frac{dV(\langle x\rangle)}{d\langle x\rangle}-\frac{\sigma_x^2}{2}\,\frac{dV^3(\langle x\rangle)}{d\langle x\rangle^3}, \tag{4.45}$$

since $\langle 1\rangle = 1$, and $\langle x-\langle x\rangle\rangle = 0$, and $\langle(x-\langle x\rangle)^2\rangle = \sigma_x^2$. The final term on the right-hand side of the above equation can be neglected when the spatial extent of the particle wavefunction, σ_x, is much smaller than the variation length-scale of the potential. In this case, Eqs. (4.42) and (4.43) reduce to

$$m\,\frac{d\langle x\rangle}{dt} = \langle p\rangle, \tag{4.46}$$

$$\frac{d\langle p\rangle}{dt} = -\frac{dV(\langle x\rangle)}{d\langle x\rangle}. \tag{4.47}$$

These equations are *exactly equivalent* to the equations of classical mechanics, with $\langle x\rangle$ playing the role of the particle displacement. Of course, if the spatial extent of the wavefunction is negligible then a measurement of x is almost certain to yield a result which lies very close to $\langle x\rangle$. Hence, we conclude that quantum mechanics corresponds to classical mechanics in the limit that the spatial extent of the wavefunction (which is typically of order the de Boglie wavelength) is negligible. This is an important result, since we know that classical mechanics gives the correct answer in this limit.

4.6 Operators

An operator, O (say), is a mathematical entity which transforms one function into another: *i.e.*,

$$O(f(x)) \rightarrow g(x). \tag{4.48}$$

For instance, x is an operator, since $x\,f(x)$ is a different function to $f(x)$, and is fully specified once $f(x)$ is given. Furthermore, d/dx is also an operator, since $df(x)/dx$ is a different function to $f(x)$, and is fully specified once $f(x)$ is given. Now,

$$x\,\frac{df}{dx} \neq \frac{d}{dx}\,(x\,f)\,. \tag{4.49}$$

This can also be written

$$x\,\frac{d}{dx} \neq \frac{d}{dx}\,x, \tag{4.50}$$

where the operators are assumed to act on everything to their right, and a final $f(x)$ is understood [where $f(x)$ is a general function]. The above expression illustrates an important point: *i.e.*, in general, *operators do not commute*. Of course, some operators do commute: *e.g.*,

$$x\,x^2 = x^2\,x. \tag{4.51}$$

Finally, an operator, O, is termed *linear* if

$$O(c\,f(x)) = c\,O(f(x)), \tag{4.52}$$

where f is a general function, and c a general complex number. All of the operators employed in quantum mechanics are linear.

Now, from Eqs. (4.22) and (4.38),

$$\begin{aligned}\langle x\rangle &= \int_{-\infty}^{\infty}\psi^*\,x\,\psi\,dx, &(4.53)\\ \langle p\rangle &= \int_{-\infty}^{\infty}\psi^*\left(-i\,\hbar\,\frac{\partial}{\partial x}\right)\psi\,dx. &(4.54)\end{aligned}$$

These expressions suggest a number of things. First, classical dynamical variables, such as x and p, are represented in quantum mechanics by *linear operators* which act on the wavefunction. Second, displacement is represented by the algebraic operator x, and momentum by the differential operator $-i\,\hbar\,\partial/\partial x$: *i.e.*,

$$p \equiv -i\,\hbar\,\frac{\partial}{\partial x}. \tag{4.55}$$

Finally, the expectation value of some dynamical variable represented by the operator $O(x)$ is simply

$$\langle O\rangle = \int_{-\infty}^{\infty}\psi^*(x,t)\,O(x)\,\psi(x,t)\,dx. \tag{4.56}$$

Clearly, if an operator is to represent a dynamical variable which has physical significance then its expectation value must be *real*. In other words, if the operator O represents a physical variable then we require that $\langle O\rangle = \langle O\rangle^*$, or

$$\int_{-\infty}^{\infty}\psi^*\,(O\,\psi)\,dx = \int_{-\infty}^{\infty}(O\,\psi)^*\,\psi\,dx, \tag{4.57}$$

where O^* is the complex conjugate of O. An operator which satisfies the above constraint is called an *Hermitian* operator. It is easily demonstrated that x and p are both Hermitian. The *Hermitian conjugate*, $O^\dagger$, of a general operator, O, is defined as follows:

$$\int_{-\infty}^{\infty}\psi^*\,(O\,\psi)\,dx = \int_{-\infty}^{\infty}(O^\dagger\,\psi)^*\,\psi\,dx. \tag{4.58}$$

The Hermitian conjugate of an Hermitian operator is the same as the operator itself: *i.e.*, $p^\dagger = p$. For a non-Hermitian operator, O (say), it is easily demonstrated that $(O^\dagger)^\dagger = O$, and that the operator $O + O^\dagger$ is Hermitian. Finally, if A and B are two operators, then $(A\,B)^\dagger = B^\dagger\,A^\dagger$.

Suppose that we wish to find the operator which corresponds to the classical dynamical variable $x\,p$. In classical mechanics, there is no difference between $x\,p$ and $p\,x$. However, in quantum mechanics, we have already seen that $x\,p \neq p\,x$. So, should be choose $x\,p$ or $p\,x$? Actually, neither of these combinations is Hermitian. However, $(1/2)\,[x\,p + (x\,p)^\dagger]$ is Hermitian. Moreover, $(1/2)\,[x\,p + (x\,p)^\dagger] = (1/2)\,(x\,p + p^\dagger\,x^\dagger) = (1/2)\,(x\,p + p\,x)$, which neatly resolves our problem of which order to put x and p.

It is a reasonable guess that the operator corresponding to energy (which is called the Hamiltonian, and conventionally denoted H) takes the form

$$H \equiv \frac{p^2}{2\,m} + V(x). \tag{4.59}$$

Note that H is Hermitian. Now, it follows from Eq. (4.55) that

$$H \equiv -\frac{\hbar^2}{2\,m}\,\frac{\partial^2}{\partial x^2} + V(x). \tag{4.60}$$

However, according to Schrödinger's equation, (4.1), we have

$$-\frac{\hbar^2}{2\,m}\,\frac{\partial^2}{\partial x^2} + V(x) = i\,\hbar\,\frac{\partial}{\partial t}, \tag{4.61}$$

so

$$H \equiv i\,\hbar\,\frac{\partial}{\partial t}. \tag{4.62}$$

Thus, the time-dependent Schrödinger equation can be written

$$i\,\hbar\,\frac{\partial\psi}{\partial t} = H\,\psi. \tag{4.63}$$

Finally, if $O(x, p, E)$ is a classical dynamical variable which is a function of displacement, momentum, and energy, then a reasonable guess for the corresponding operator in quantum mechanics is $(1/2)\,[O(x, p, H) + O^\dagger(x, p, H)]$, where $p = -i\,\hbar\,\partial/\partial x$, and $H = i\,\hbar\,\partial/\partial t$.

4.7 Momentum Representation

Fourier's theorerm (see Sect. 3.12), applied to one-dimensional wavefunctions, yields

$$\psi(x, t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\bar{\psi}(k, t)\,e^{+i\,k\,x}\,dk, \tag{4.64}$$

$$\bar{\psi}(k, t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\psi(x, t)\,e^{-i\,k\,x}\,dx, \tag{4.65}$$

where k represents wavenumber. However, $p = \hbar\,k$. Hence, we can also write

$$\psi(x, t) = \frac{1}{\sqrt{2\pi\,\hbar}}\int_{-\infty}^{\infty}\phi(p, t)\,e^{+i\,p\,x/\hbar}\,dp, \tag{4.66}$$

$$\phi(p, t) = \frac{1}{\sqrt{2\pi\,\hbar}}\int_{-\infty}^{\infty}\psi(x, t)\,e^{-i\,p\,x/\hbar}\,dx, \tag{4.67}$$

where $\phi(p, t) = \bar{\psi}(k, t)/\sqrt{\hbar}$ is the momentum-space equivalent to the real-space wavefunction $\psi(x, t)$.

At this stage, it is convenient to introduce a useful function called the *Dirac delta-function*. This function, denoted $\delta(x)$, was first devised by Paul Dirac, and has the following rather unusual properties: $\delta(x)$ is zero for $x \neq 0$, and is infinite at $x = 0$. However, the singularity at $x = 0$ is such that

$$\int_{-\infty}^{\infty} \delta(x)\, dx = 1. \tag{4.68}$$

The delta-function is an example of what is known as a *generalized function*: *i.e.*, its value is not well-defined at all x, but its integral is well-defined. Consider the integral

$$\int_{-\infty}^{\infty} f(x)\, \delta(x)\, dx. \tag{4.69}$$

Since $\delta(x)$ is only non-zero infinitesimally close to $x = 0$, we can safely replace $f(x)$ by $f(0)$ in the above integral (assuming $f(x)$ is well behaved at $x = 0$), to give

$$\int_{-\infty}^{\infty} f(x)\, \delta(x)\, dx = f(0) \int_{-\infty}^{\infty} \delta(x)\, dx = f(0), \tag{4.70}$$

where use has been made of Eq. (4.68). A simple generalization of this result yields

$$\int_{-\infty}^{\infty} f(x)\, \delta(x - x_0)\, dx = f(x_0), \tag{4.71}$$

which can also be thought of as an alternative definition of a delta-function.

Suppose that $\psi(x) = \delta(x - x_0)$. It follows from Eqs. (4.67) and (4.71) that

$$\phi(p) = \frac{e^{-i\,p\,x_0/\hbar}}{\sqrt{2\pi\,\hbar}}. \tag{4.72}$$

Hence, Eq. (4.66) yields the important result

$$\delta(x - x_0) = \frac{1}{2\pi\,\hbar} \int_{-\infty}^{\infty} e^{+i\,p\,(x - x_0)/\hbar}\, dp. \tag{4.73}$$

Similarly,

$$\delta(p - p_0) = \frac{1}{2\pi\,\hbar} \int_{-\infty}^{\infty} e^{+i\,(p - p_0)\,x/\hbar}\, dx. \tag{4.74}$$

It turns out that we can just as well formulate quantum mechanics using momentum-space wavefunctions, $\phi(p, t)$, as real-space wavefunctions, $\psi(x, t)$. The former scheme is known as the *momentum representation* of quantum mechanics. In the momentum representation, wavefunctions are the Fourier transforms of the equivalent real-space wavefunctions, and dynamical variables are represented by *different* operators. Furthermore, by analogy with Eq. (4.56), the expectation value of some operator $O(p)$ takes the form

$$\langle O \rangle = \int_{-\infty}^{\infty} \phi^*(p, t)\, O(p)\, \phi(p, t)\, dp. \tag{4.75}$$

Consider momentum. We can write

$$\begin{aligned}\langle p\rangle &= \int_{-\infty}^{\infty}\psi^*(x,t)\left(-i\,\hbar\,\frac{\partial}{\partial x}\right)\psi(x,t)\,dx\\ &= \frac{1}{2\pi\,\hbar}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi^*(p',t)\,\phi(p,t)\,p\,e^{+i\,(p-p')\,x/\hbar}\,dx\,dp\,dp',\end{aligned}\tag{4.76}$$

where use has been made of Eq. (4.66). However, it follows from Eq. (4.74) that

$$\langle p\rangle = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi^*(p',t)\,\phi(p,t)\,p\,\delta(p-p')\,dp\,dp'.\tag{4.77}$$

Hence, using Eq. (4.71), we obtain

$$\langle p\rangle = \int_{-\infty}^{\infty}\phi^*(p,t)\,p\,\phi(p,t)\,dp = \int_{-\infty}^{\infty}p\,|\phi|^2\,dp.\tag{4.78}$$

Evidently, momentum is represented by the operator p in the momentum representation. The above expression also strongly suggests [by comparison with Eq. (4.22)] that $|\phi(p,t)|^2$ can be interpreted as the probability density of a measurement of momentum yielding the value p at time t. It follows that $\phi(p,t)$ must satisfy an analogous normalization condition to Eq. (4.4): *i.e.*,

$$\int_{-\infty}^{\infty}|\phi(p,t)|^2\,dp = 1.\tag{4.79}$$

Consider displacement. We can write

$$\begin{aligned}\langle x\rangle &= \int_{-\infty}^{\infty}\psi^*(x,t)\,x\,\psi(x,t)\,dx \qquad (4.80)\\ &= \frac{1}{2\pi\,\hbar}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi^*(p',t)\,\phi(p,t)\left(-i\,\hbar\,\frac{\partial}{\partial p}\right)e^{+i\,(p-p')\,x/\hbar}\,dx\,dp\,dp'.\end{aligned}$$

Integration by parts yields

$$\langle x\rangle = \frac{1}{2\pi\,\hbar}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi^*(p',t)\,e^{+i\,(p-p')\,x/\hbar}\left(i\,\hbar\,\frac{\partial}{\partial p}\right)\phi(p,t)\,dx\,dp\,dp'.\tag{4.81}$$

Hence, making use of Eqs. (4.74) and (4.71), we obtain

$$\langle x\rangle = \frac{1}{2\pi\,\hbar}\int_{-\infty}^{\infty}\phi^*(p)\left(i\,\hbar\,\frac{\partial}{\partial p}\right)\phi(p)\,dp.\tag{4.82}$$

Evidently, displacement is represented by the operator

$$x \equiv i\,\hbar\,\frac{\partial}{\partial p}\tag{4.83}$$

in the momentum representation.

Finally, let us consider the normalization of the momentum-space wavefunction $\phi(p,t)$. We have

$$\int_{-\infty}^{\infty} \psi^*(x,t)\,\psi(x,t)\,dx = \frac{1}{2\pi\,\hbar}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi^*(p',t)\,\phi(p,t)\,e^{+i\,(p-p')\,x/\hbar}\,dx\,dp\,dp'. \tag{4.84}$$

Thus, it follows from Eqs. (4.71) and (4.74) that

$$\int_{-\infty}^{\infty} |\psi(x,t)|^2\,dx = \int_{-\infty}^{\infty} |\phi(p,t)|^2\,dp. \tag{4.85}$$

Hence, if $\psi(x,t)$ is properly normalized [see Eq. (4.4)] then $\phi(p,t)$, as defined in Eq. (4.67), is also properly normalized [see Eq. (4.79)].

The existence of the momentum representation illustrates an important point: *i.e.*, that there are many different, but entirely equivalent, ways of mathematically formulating quantum mechanics. For instance, it is also possible to represent wavefunctions as row and column vectors, and dynamical variables as matrices which act upon these vectors.

4.8 Heisenberg's Uncertainty Principle

Consider a real-space Hermitian operator $O(x)$. A straightforward generalization of Eq. (4.57) yields

$$\int_{-\infty}^{\infty} \psi_1^*\,(O\,\psi_2)\,dx = \int_{-\infty}^{\infty} (O\,\psi_1)^*\,\psi_2\,dx, \tag{4.86}$$

where $\psi_1(x)$ and $\psi_2(x)$ are general functions.

Let $f = (A - \langle A\rangle)\,\psi$, where $A(x)$ is an Hermitian operator, and $\psi(x)$ a general wavefunction. We have

$$\int_{-\infty}^{\infty} |f|^2\,dx = \int_{-\infty}^{\infty} f^*\,f\,dx = \int_{-\infty}^{\infty} [(A - \langle A\rangle)\,\psi]^*\,[(A - \langle A\rangle)\,\psi]\,dx. \tag{4.87}$$

Making use of Eq. (4.86), we obtain

$$\int_{-\infty}^{\infty} |f|^2\,dx = \int_{-\infty}^{\infty} \psi^*\,(A - \langle A\rangle)^2\,\psi\,dx = \sigma_A^2, \tag{4.88}$$

where σ_A^2 is the variance of A [see Eq. (4.24)]. Similarly, if $g = (B - \langle B\rangle)\,\psi$, where B is a second Hermitian operator, then

$$\int_{-\infty}^{\infty} |g|^2\,dx = \sigma_B^2, \tag{4.89}$$

Now, there is a standard result in mathematics, known as the *Schwartz inequality*, which states that

$$\left|\int_a^b f^*(x)\,g(x)\,dx\right|^2 \leq \int_a^b |f(x)|^2\,dx \int_a^b |g(x)|^2\,dx, \tag{4.90}$$

where f and g are two general functions. Furthermore, if z is a complex number then

$$|z|^2 = [\mathrm{Re}(z)]^2 + [\mathrm{Im}(z)]^2 \geq [\mathrm{Im}(z)]^2 = \left[\frac{1}{2\,i}\,(z - z^*)\right]^2. \tag{4.91}$$

Hence, if $z = \int_{-\infty}^{\infty} f^*\,g\,dx$ then Eqs. (4.88)–(4.91) yield

$$\sigma_A^2\,\sigma_B^2 \geq \left[\frac{1}{2\,i}\,(z - z^*)\right]^2. \tag{4.92}$$

However,

$$z = \int_{-\infty}^{\infty} [(A - \langle A\rangle)\,\psi]^*\,[(B - \langle B\rangle)\,\psi]\,dx = \int_{-\infty}^{\infty} \psi^*\,(A - \langle A\rangle)\,(B - \langle B\rangle)\,\psi\,dx, \tag{4.93}$$

where use has been made of Eq. (4.86). The above equation reduces to

$$z = \int_{-\infty}^{\infty} \psi^*\,A\,B\,\psi\,dx - \langle A\rangle\,\langle B\rangle. \tag{4.94}$$

Furthermore, it is easily demonstrated that

$$z^* = \int_{-\infty}^{\infty} \psi^*\,B\,A\,\psi\,dx - \langle A\rangle\,\langle B\rangle. \tag{4.95}$$

Hence, Eq. (4.92) gives

$$\sigma_A^2\,\sigma_B^2 \geq \left(\frac{1}{2\,i}\langle[A, B]\rangle\right)^2, \tag{4.96}$$

where

$$[A, B] \equiv A\,B - B\,A. \tag{4.97}$$

Equation (4.96) is the general form of *Heisenberg's uncertainty principle* in quantum mechanics. It states that if two dynamical variables are represented by the two Hermitian operators A and B, and these operators *do not commute* (*i.e.*, $A\,B \neq B\,A$), then it is *impossible* to simultaneously (exactly) measure the two variables. Instead, the product of the variances in the measurements is always greater than some critical value, which depends on the extent to which the two operators do not commute.

For instance, displacement and momentum are represented (in real-space) by the operators x and $p \equiv -i\,\hbar\,\partial/\partial x$, respectively. Now, it is easily demonstrated that

$$[x, p] = i\,\hbar. \tag{4.98}$$

Thus,

$$\sigma_x\,\sigma_p \geq \frac{\hbar}{2}, \tag{4.99}$$

which can be recognized as the standard displacement-momentum uncertainty principle (see Sect. 3.14). It turns out that the minimum uncertainty (*i.e.*, $\sigma_x\,\sigma_p = \hbar/2$) is only achieved by *Gaussian* wave packets (see Sect. 3.12): *i.e.*,

$$\psi(x) = \frac{e^{+i\,p_0\,x/\hbar}}{(2\pi\,\sigma_x^2)^{1/4}}\,e^{-(x-x_0)^2/4\,\sigma_x^2}, \tag{4.100}$$

$$\phi(p) = \frac{e^{-i\,p\,x_0/\hbar}}{(2\pi\,\sigma_p^2)^{1/4}}\,e^{-(p-p_0)^2/4\,\sigma_p^2}, \tag{4.101}$$

where $\phi(p)$ is the momentum-space equivalent of $\psi(x)$.

Energy and time are represented by the operators $H \equiv i\,\hbar\,\partial/\partial t$ and t, respectively. These operators do not commute, indicating that energy and time cannot be measured simultaneously. In fact,

$$[H, t] = i\,\hbar, \tag{4.102}$$

so

$$\sigma_E\,\sigma_t \geq \frac{\hbar}{2}. \tag{4.103}$$

This can be written, somewhat less exactly, as

$$\Delta E\,\Delta t \gtrsim \hbar, \tag{4.104}$$

where ΔE and Δt are the uncertainties in energy and time, respectively. The above expression is generally known as the *energy-time uncertainty principle*.

For instance, suppose that a particle passes some fixed point on the x-axis. Since the particle is, in reality, an extended wave packet, it takes a certain amount of time Δt for the particle to pass. Thus, there is an uncertainty, Δt, in the arrival time of the particle. Moreover, since $E = \hbar\,\omega$, the only wavefunctions which have unique energies are those with unique frequencies: *i.e.*, plane waves. Since a wave packet of finite extent is made up of a combination of plane waves of different wavenumbers, and, hence, different frequencies, there will be an uncertainty ΔE in the particle's energy which is proportional to the range of frequencies of the plane waves making up the wave packet. The more compact the wave packet (and, hence, the smaller Δt), the larger the range of frequencies of the constituent plane waves (and, hence, the large ΔE), and *vice versa*. To be more exact, if $\psi(t)$ is the wavefunction measured at the fixed point as a function of time, then we can write

$$\psi(t) = \frac{1}{\sqrt{2\pi\,\hbar}}\int_{-\infty}^{\infty}\chi(E)\,e^{-i\,E\,t/\hbar}\,dE. \tag{4.105}$$

In other words, we can express $\psi(t)$ as a linear combination of plane waves of definite energy E. Here, $\chi(E)$ is the complex amplitude of plane waves of energy E in this combination. By Fourier's theorem, we also have

$$\chi(E) = \frac{1}{\sqrt{2\pi\,\hbar}}\int_{-\infty}^{\infty}\psi(t)\,e^{+i\,E\,t/\hbar}\,dt. \tag{4.106}$$

For instance, if $\psi(t)$ is a Gaussian then it is easily shown that $\chi(E)$ is also a Gaussian: *i.e.*,

$$\psi(t) = \frac{e^{-i\,E_0\,t/\hbar}}{(2\pi\,\sigma_t^2)^{1/4}}\,e^{-(t-t_0)^2/4\,\sigma_t^2}, \tag{4.107}$$

$$\chi(E) = \frac{e^{+i\,E\,t_0/\hbar}}{(2\pi\,\sigma_E^2)^{1/4}}\,e^{-(E-E_0)^2/4\,\sigma_E^2}, \tag{4.108}$$

where $\sigma_E\,\sigma_t = \hbar/2$. As before, Gaussian wave packets satisfy the minimum uncertainty principle $\sigma_E\,\sigma_t = \hbar/2$. Conversely, non-Gaussian wave packets are characterized by $\sigma_E\,\sigma_t > \hbar/2$.

4.9 Eigenstates and Eigenvalues

Consider a general real-space operator $A(x)$. When this operator acts on a general wavefunction $\psi(x)$ the result is usually a wavefunction with a completely different shape. However, there are certain special wavefunctions which are such that when A acts on them the result is just a multiple of the original wavefunction. These special wavefunctions are called *eigenstates*, and the multiples are called *eigenvalues*. Thus, if

$$A\,\psi_a(x) = a\,\psi_a(x), \tag{4.109}$$

where a is a complex number, then ψ_a is called an eigenstate of A corresponding to the eigenvalue a.

Suppose that A is an Hermitian operator corresponding to some physical dynamical variable. Consider a particle whose wavefunction is ψ_a. The expectation of value A in this state is simply [see Eq. (4.56)]

$$\langle A\rangle = \int_{-\infty}^{\infty}\psi_a^*\,A\,\psi_a\,dx = a\int_{-\infty}^{\infty}\psi_a^*\,\psi_a\,dx = a, \tag{4.110}$$

where use has been made of Eq. (4.109) and the normalization condition (4.4). Moreover,

$$\langle A^2\rangle = \int_{-\infty}^{\infty}\psi_a^*\,A^2\,\psi_a\,dx = a\int_{-\infty}^{\infty}\psi_a^*\,A\,\psi_a\,dx = a^2\int_{-\infty}^{\infty}\psi_a^*\,\psi_a\,dx = a^2, \tag{4.111}$$

so the variance of A is [*cf.*, Eq. (4.24)]

$$\sigma_A^2 = \langle A^2\rangle - \langle A\rangle^2 = a^2 - a^2 = 0. \tag{4.112}$$

The fact that the variance is *zero* implies that every measurement of A is bound to yield the same result: namely, a. Thus, the eigenstate ψ_a is a state which is associated with a *unique* value of the dynamical variable corresponding to A. This unique value is simply the associated eigenvalue.

It is easily demonstrated that the eigenvalues of an Hermitian operator are all *real*. Recall [from Eq. (4.86)] that an Hermitian operator satisfies

$$\int_{-\infty}^{\infty} \psi_1^*\,(A\,\psi_2)\,dx = \int_{-\infty}^{\infty} (A\,\psi_1)^*\,\psi_2\,dx. \tag{4.113}$$

Hence, if $\psi_1 = \psi_2 = \psi_a$ then

$$\int_{-\infty}^{\infty} \psi_a^*\,(A\,\psi_a)\,dx = \int_{-\infty}^{\infty} (A\,\psi_a)^*\,\psi_a\,dx, \tag{4.114}$$

which reduces to [see Eq. (4.109)]

$$a = a^*, \tag{4.115}$$

assuming that ψ_a is properly normalized.

Two wavefunctions, $\psi_1(x)$ and $\psi_2(x)$, are said to be *orthogonal* if

$$\int_{-\infty}^{\infty} \psi_1^*\,\psi_2\,dx = 0. \tag{4.116}$$

Consider two eigenstates of A, ψ_a and $\psi_{a'}$, which correspond to the two *different* eigenvalues a and a', respectively. Thus,

$$A\,\psi_a = a\,\psi_a, \tag{4.117}$$

$$A\,\psi_{a'} = a'\,\psi_{a'}. \tag{4.118}$$

Multiplying the complex conjugate of the first equation by $\psi_{a'}$, and the second equation by ψ_a^*, and then integrating over all x, we obtain

$$\int_{-\infty}^{\infty} (A\,\psi_a)^*\,\psi_{a'}\,dx = a\int_{-\infty}^{\infty} \psi_a^*\,\psi_{a'}\,dx, \tag{4.119}$$

$$\int_{-\infty}^{\infty} \psi_a^*\,(A\,\psi_{a'})\,dx = a'\int_{-\infty}^{\infty} \psi_a^*\,\psi_{a'}\,dx. \tag{4.120}$$

However, from Eq. (4.113), the left-hand sides of the above two equations are equal. Hence, we can write

$$(a - a')\int_{-\infty}^{\infty} \psi_a^*\,\psi_{a'}\,dx = 0. \tag{4.121}$$

By assumption, $a \neq a'$, yielding

$$\int_{-\infty}^{\infty} \psi_a^*\,\psi_{a'}\,dx = 0. \tag{4.122}$$

In other words, eigenstates of an Hermitian operator corresponding to *different* eigenvalues are automatically *orthogonal*.

Consider two eigenstates of A, ψ_a and ψ'_a, which correspond to the *same* eigenvalue, a. Such eigenstates are termed *degenerate*. The above proof of the orthogonality of different eigenstates fails for degenerate eigenstates. Note, however, that *any* linear combination of ψ_a and ψ'_a is also an eigenstate of A corresponding to the eigenvalue a. Thus, even if ψ_a and ψ'_a are not orthogonal, we can always choose two linear combinations of these eigenstates which are orthogonal. For instance, if ψ_a and ψ'_a are properly normalized, and

$$\int_{-\infty}^{\infty} \psi_a^*\, \psi'_a\, dx = c, \tag{4.123}$$

then it is easily demonstrated that

$$\psi''_a = \frac{|c|}{\sqrt{1-|c|^2}} \left(\psi_a - c^{-1}\, \psi'_a\right) \tag{4.124}$$

is a properly normalized eigenstate of A, corresponding to the eigenvalue a, which is orthogonal to ψ_a. It is straightforward to generalize the above argument to three or more degenerate eigenstates. Hence, we conclude that the eigenstates of an Hermitian operator are, or can be chosen to be, *mutually orthogonal*.

It is also possible to demonstrate that the eigenstates of an Hermitian operator form a *complete set*: *i.e.*, that any general wavefunction can be written as a linear combination of these eigenstates. However, the proof is quite difficult, and we shall not attempt it here.

In summary, given an Hermitian operator A, any general wavefunction, $\psi(x)$, can be written

$$\psi = \sum_i c_i\, \psi_i, \tag{4.125}$$

where the c_i are complex weights, and the ψ_i are the properly normalized (and mutually orthogonal) eigenstates of A: *i.e.*,

$$A\, \psi_i = a_i\, \psi_i, \tag{4.126}$$

where a_i is the eigenvalue corresponding to the eigenstate ψ_i, and

$$\int_{-\infty}^{\infty} \psi_i^*\, \psi_j\, dx = \delta_{ij}. \tag{4.127}$$

Here, δ_{ij} is called the *Kronecker delta-function*, and takes the value unity when its two indices are equal, and zero otherwise.

It follows from Eqs. (4.125) and (4.127) that

$$c_i = \int_{-\infty}^{\infty} \psi_i^*\, \psi\, dx. \tag{4.128}$$

Thus, the expansion coefficients in Eq. (4.125) are easily determined, given the wavefunction ψ and the eigenstates ψ_i. Moreover, if ψ is a properly normalized wavefunction then Eqs. (4.125) and (4.127) yield

$$\sum_i |c_i|^2 = 1. \tag{4.129}$$

4.10 Measurement

Suppose that A is an Hermitian operator corresponding to some dynamical variable. By analogy with the discussion in Sect. 3.16, we expect that if a measurement of A yields the result a then the act of measurement will cause the wavefunction to collapse to a state in which a measurement of A is *bound* to give the result a. What sort of wavefunction, ψ, is such that a measurement of A is bound to yield a certain result, a? Well, expressing ψ as a linear combination of the eigenstates of A, we have

$$\psi = \sum_i c_i\,\psi_i, \tag{4.130}$$

where ψ_i is an eigenstate of A corresponding to the eigenvalue a_i. If a measurement of A is bound to yield the result a then

$$\langle A\rangle = a, \tag{4.131}$$

and

$$\sigma_A^2 = \langle A^2\rangle - \langle A\rangle = 0. \tag{4.132}$$

Now it is easily seen that

$$\begin{aligned} \langle A\rangle &= \sum_i |c_i|^2\,a_i, && (4.133)\\ \langle A^2\rangle &= \sum_i |c_i|^2\,a_i^2. && (4.134) \end{aligned}$$

Thus, Eq. (4.132) gives

$$\sum_i a_i^2\,|c_i|^2 - \left(\sum_i a_i\,|c_i|^2\right)^2 = 0. \tag{4.135}$$

Furthermore, the normalization condition yields

$$\sum_i |c_i|^2 = 1. \tag{4.136}$$

For instance, suppose that there are only two eigenstates. The above two equations then reduce to $|c_1|^2 = x$, and $|c_2|^2 = 1 - x$, where $0 \leq x \leq 1$, and

$$(a_1 - a_2)^2\,x\,(1-x) = 0. \tag{4.137}$$

The only solutions are $x = 0$ and $x = 1$. This result can easily be generalized to the case where there are more than two eigenstates. It follows that a state associated with a definite value of A is one in which one of the $|c_i|^2$ is unity, and all of the others are zero. In other words, the only states associated with definite values of A are the *eigenstates* of A. It immediately follows that *the result of a measurement of* A *must be one of the eigenvalues of* A. Moreover, if a general wavefunction is expanded as a linear combination of the eigenstates of A, like in Eq. (4.130), then it is clear from Eq. (4.133), and the general definition of a mean, that the probability of a measurement of A yielding the eigenvalue a_i is simply $|c_i|^2$, where c_i is the coefficient in front of the ith eigenstate in the expansion. Note, from Eq. (4.136), that these probabilities are properly normalized: *i.e.*, the probability of a measurement of A resulting in any possible answer is unity. Finally, if a measurement of A results in the eigenvalue a_i then immediately after the measurement the system will be left in the eigenstate corresponding to a_i.

Consider two physical dynamical variables represented by the two Hermitian operators A and B. Under what circumstances is it possible to simultaneously measure these two variables (exactly)? Well, the possible results of measurements of A and B are the eigenvalues of A and B, respectively. Thus, to simultaneously measure A and B (exactly) there must exist states which are *simultaneous eigenstates* of A and B. In fact, in order for A and B to be simultaneously measurable under all circumstances, we need *all* of the eigenstates of A to also be eigenstates of B, and *vice versa*, so that all states associated with unique values of A are also associated with unique values of B, and *vice versa*.

Now, we have already seen, in Sect. 4.8, that if A and B do not commute (*i.e.*, if $A\,B \neq B\,A$) then they cannot be simultaneously measured. This suggests that the condition for simultaneous measurement is that A and B should commute. Suppose that this is the case, and that the ψ_i and a_i are the normalized eigenstates and eigenvalues of A, respectively. It follows that

$$(A\,B - B\,A)\,\psi_i = (A\,B - B\,a_i)\,\psi_i = (A - a_i)\,B\,\psi_i = 0, \tag{4.138}$$

or

$$A\,(B\,\psi_i) = a_i\,(B\,\psi_i). \tag{4.139}$$

Thus, $B\,\psi_i$ is an eigenstate of A corresponding to the eigenvalue a_i (though not necessarily a normalized one). In other words, $B\,\psi_i \propto \psi_i$, or

$$B\,\psi_i = b_i\,\psi_i, \tag{4.140}$$

where b_i is a constant of proportionality. Hence, ψ_i is an eigenstate of B, and, thus, a simultaneous eigenstate of A and B. We conclude that if A and B commute then they possess simultaneous eigenstates, and are thus simultaneously measurable (exactly).

4.11 Continuous Eigenvalues

In the previous two sections, it was tacitly assumed that we were dealing with operators possessing *discrete* eigenvalues and square-integrable eigenstates. Unfortunately, some operators—most notably, x and p—possess eigenvalues which lie in a *continuous range* and non-square-integrable eigenstates (in fact, these two properties go hand in hand). Let us, therefore, investigate the eigenstates and eigenvalues of the displacement and momentum operators.

Let $\psi_x(x, x')$ be the eigenstate of x corresponding to the eigenvalue x'. It follows that

$$x\,\psi_x(x, x') = x'\,\psi_x(x, x') \tag{4.141}$$

for all x. Consider the Dirac delta-function $\delta(x - x')$. We can write

$$x\,\delta(x - x') = x'\,\delta(x - x'), \tag{4.142}$$

since $\delta(x - x')$ is only non-zero infinitesimally close to $x = x'$. Evidently, $\psi_x(x, x')$ is proportional to $\delta(x - x')$. Let us make the constant of proportionality unity, so that

$$\psi_x(x, x') = \delta(x - x'). \tag{4.143}$$

Now, it is easily demonstrated that

$$\int_{-\infty}^{\infty} \delta(x - x')\,\delta(x - x'')\,dx = \delta(x' - x''). \tag{4.144}$$

Hence, $\psi_x(x, x')$ satisfies the orthonormality condition

$$\int_{-\infty}^{\infty} \psi_x^*(x, x')\,\psi_x(x, x'')\,dx = \delta(x' - x''). \tag{4.145}$$

This condition is analogous to the orthonormality condition (4.127) satisfied by square-integrable eigenstates. Now, by definition, $\delta(x - x')$ satisfies

$$\int_{-\infty}^{\infty} f(x)\,\delta(x - x')\,dx = f(x'), \tag{4.146}$$

where $f(x)$ is a general function. We can thus write

$$\psi(x) = \int_{-\infty}^{\infty} c(x')\,\psi_x(x, x')\,dx', \tag{4.147}$$

where $c(x') = \psi(x')$, or

$$c(x') = \int_{-\infty}^{\infty} \psi_x^*(x, x')\,\psi(x)\,dx. \tag{4.148}$$

In other words, we can expand a general wavefunction $\psi(x)$ as a linear combination of the eigenstates, $\psi_x(x, x')$, of the displacement operator. Equations (4.147) and (4.148) are analogous to Eqs. (4.125) and (4.128), respectively, for square-integrable eigenstates. Finally, by analogy with the results in Sect. 4.9, the probability density of a measurement of x yielding the value x' is $|c(x')|^2$, which is equivalent to the standard result $|\psi(x')|^2$. Moreover, these probabilities are properly normalized provided $\psi(x)$ is properly normalized [*cf.*, Eq. (4.129)]: *i.e.*,

$$\int_{-\infty}^{\infty} |c(x')|^2\,dx' = \int_{-\infty}^{\infty} |\psi(x')|^2\,dx' = 1. \tag{4.149}$$

Finally, if a measurement of x yields the value x' then the system is left in the corresponding displacement eigenstate, $\psi_x(x, x')$, immediately after the measurement: *i.e.*, the wavefunction collapses to a "spike-function", $\delta(x - x')$, as discussed in Sect. 3.16.

Now, an eigenstate of the momentum operator $p \equiv -i\,\hbar\,\partial/\partial x$ corresponding to the eigenvalue p' satisfies

$$-\,i\,\hbar\,\frac{\partial\psi_p(x, p')}{\partial x} = p'\,\psi_p(x, p'). \tag{4.150}$$

It is evident that

$$\psi_p(x, p') \propto e^{+i\,p'\,x/\hbar}. \tag{4.151}$$

Now, we require $\psi_p(x, p')$ to satisfy an analogous orthonormality condition to Eq. (4.145): *i.e.*,

$$\int_{-\infty}^{\infty} \psi_p^*(x, p')\,\psi_p(x, p'')\,dx = \delta(p' - p''). \tag{4.152}$$

Thus, it follows from Eq. (4.74) that the constant of proportionality in Eq. (4.151) should be $(2\pi\,\hbar)^{-1/2}$: *i.e.*,

$$\psi_p(x, p') = \frac{e^{+i\,p'\,x/\hbar}}{(2\pi\,\hbar)^{1/2}}. \tag{4.153}$$

Furthermore, according to Eqs. (4.66) and (4.67),

$$\psi(x) = \int_{-\infty}^{\infty} c(p')\,\psi_p(x, p')\,dp', \tag{4.154}$$

where $c(p') = \phi(p')$ [see Eq. (4.67)], or

$$c(p') = \int_{-\infty}^{\infty} \psi_p^*(x, p')\,\psi(x)\,dx. \tag{4.155}$$

In other words, we can expand a general wavefunction $\psi(x)$ as a linear combination of the eigenstates, $\psi_p(x, p')$, of the momentum operator. Equations (4.154) and (4.155) are again analogous to Eqs. (4.125) and (4.128), respectively, for square-integrable eigenstates. Likewise, the probability density of a measurement of p yielding the result p' is $|c(p')|^2$, which is equivalent to the standard result $|\phi(p')|^2$. The probabilities are also properly normalized provided $\psi(x)$ is properly normalized [*cf.*, Eq. (4.85)]: *i.e.*,

$$\int_{-\infty}^{\infty} |c(p')|^2\,dp' = \int_{-\infty}^{\infty} |\phi(p')|^2\,dp' = \int_{-\infty}^{\infty} |\psi(x')|^2\,dx' = 1. \tag{4.156}$$

Finally, if a mesurement of p yields the value p' then the system is left in the corresponding momentum eigenstate, $\psi_p(x, p')$, immediately after the measurement.

4.12 Stationary States

An eigenstate of the energy operator $H \equiv i\,\hbar\,\partial/\partial t$ corresponding to the eigenvalue E_i satisfies

$$i\,\hbar\,\frac{\partial\psi_E(x,t,E_i)}{\partial t} = E_i\,\psi_E(x,t,E_i). \tag{4.157}$$

It is evident that this equation can be solved by writing

$$\psi_E(x,t,E_i) = \psi_i(x)\,e^{-i\,E_i\,t/\hbar}, \tag{4.158}$$

where $\psi_i(x)$ is a properly normalized stationary (*i.e.*, non-time-varying) wavefunction. The wavefunction $\psi_E(x,t,E_i)$ corresponds to a so-called *stationary state*, since the probability density $|\psi_E|^2$ is non-time-varying. Note that a stationary state is associated with a *unique* value for the energy. Substitution of the above expression into Schrödinger's equation (4.1) yields the equation satisfied by the stationary wavefunction:

$$\frac{\hbar^2}{2\,m}\,\frac{d^2\psi_i}{dx^2} = [V(x) - E_i]\,\psi_i. \tag{4.159}$$

This is known as the *time-independent Schrödinger equation*. More generally, this equation takes the form

$$H\,\psi_i = E_i\,\psi_i, \tag{4.160}$$

where H is assumed not to be an explicit function of t. Of course, the ψ_i satisfy the usual orthonormality condition:

$$\int_{-\infty}^{\infty}\psi_i^*\,\psi_j\,dx = \delta_{ij}. \tag{4.161}$$

Moreover, we can express a general wavefunction as a linear combination of energy eigenstates:

$$\psi(x,t) = \sum_i c_i\,\psi_i(x)\,e^{-i\,E_i\,t/\hbar}, \tag{4.162}$$

where

$$c_i = \int_{-\infty}^{\infty}\psi_i^*(x)\,\psi(x,0)\,dx. \tag{4.163}$$

Here, $|c_i|^2$ is the probability that a measurement of the energy will yield the eigenvalue E_i. Furthermore, immediately after such a measurement, the system is left in the corresponding energy eigenstate. The generalization of the above results to the case where H has continuous eigenvalues is straightforward.

If a dynamical variable is represented by some Hermitian operator A which commutes with H (so that it has simultaneous eigenstates with H), and contains no specific time dependence, then it is evident from Eqs. (4.161) and (4.162) that the expectation value and variance of A are *time independent*. In this sense, the dynamical variable in question is a constant of the motion.

Exercises

1. Monochromatic light with a wavelength of 6000 Å passes through a fast shutter that opens for 10^{-9} sec. What is the subsequent spread in wavelengths of the no longer monochromatic light?

2. Calculate $\langle x\rangle$, $\langle x^2\rangle$, and σ_x, as well as $\langle p\rangle$, $\langle p^2\rangle$, and σ_p, for the normalized wavefunction
$$\psi(x) = \sqrt{\frac{2\,a^3}{\pi}}\,\frac{1}{x^2+a^2}.$$
Use these to find $\sigma_x\,\sigma_p$. Note that $\int_{-\infty}^{\infty} dx/(x^2+a^2) = \pi/a$.

3. Classically, if a particle is not observed then the probability of finding it in a one-dimensional box of length L, which extends from $x = 0$ to $x = L$, is a constant $1/L$ per unit length. Show that the classical expectation value of x is $L/2$, the expectation value of x^2 is $L^2/3$, and the standard deviation of x is $L/\sqrt{12}$.

4. Demonstrate that if a particle in a one-dimensional stationary state is bound then the expectation value of its momentum must be zero.

5. Suppose that $V(x)$ is complex. Obtain an expression for $\partial P(x,t)/\partial t$ and $d/dt\int P(x,t)\,dx$ from Schrödinger's equation. What does this tell us about a complex $V(x)$?

6. $\psi_1(x)$ and $\psi_2(x)$ are normalized eigenfunctions corresponding to the same eigenvalue. If
$$\int_{-\infty}^{\infty}\psi_1^*\,\psi_2\,dx = c,$$
where c is real, find normalized linear combinations of ψ_1 and ψ_2 which are orthogonal to (a) ψ_1, (b) $\psi_1+\psi_2$.

7. Demonstrate that $p = -i\,\hbar\,\partial/\partial x$ is an Hermitian operator. Find the Hermitian conjugate of $a = x + i\,p$.

8. An operator A, corresponding to a physical quantity α, has two normalized eigenfunctions $\psi_1(x)$ and $\psi_2(x)$, with eigenvalues a_1 and a_2. An operator B, corresponding to another physical quantity β, has normalized eigenfunctions $\phi_1(x)$ and $\phi_2(x)$, with eigenvalues b_1 and b_2. The eigenfunctions are related via
$$\begin{aligned}\psi_1 &= (2\,\phi_1 + 3\,\phi_2)\big/\sqrt{13},\\ \psi_2 &= (3\,\phi_1 - 2\,\phi_2)\big/\sqrt{13}.\end{aligned}$$
α is measured and the value a_1 is obtained. If β is then measured and then α again, show that the probability of obtaining a_1 a second time is $97/169$.

9. Demonstrate that an operator which commutes with the Hamiltonian, and contains no explicit time dependence, has an expectation value which is constant in time.

10. For a certain system, the operator corresponding to the physical quantity A does not commute with the Hamiltonian. It has eigenvalues a_1 and a_2, corresponding to properly normalized eigenfunctions

$$\begin{aligned} \phi_1 &= (u_1 + u_2) \Big/ \sqrt{2}, \\ \phi_2 &= (u_1 - u_2) \Big/ \sqrt{2}, \end{aligned}$$

where u_1 and u_2 are properly normalized eigenfunctions of the Hamiltonian with eigenvalues E_1 and E_2. If the system is in the state $\psi = \phi_1$ at time $t = 0$, show that the expectation value of A at time t is

$$\langle A \rangle = \left(\frac{a_1 + a_2}{2} \right) + \left(\frac{a_1 - a_2}{2} \right) \cos\left(\frac{[E_1 - E_2]\, t}{\hbar} \right).$$

5 One-Dimensional Potentials

5.1 Introduction

In this chapter, we shall investigate the interaction of a non-relativistic particle of mass m and energy E with various one-dimensional potentials, $V(x)$. Since we are searching for stationary solutions with unique energies, we can write the wavefunction in the form (see Sect. 4.12)

$$\psi(x,t) = \psi(x)\, e^{-i\,E\,t/\hbar}, \tag{5.1}$$

where $\psi(x)$ satisfies the time-independent Schrödinger equation:

$$\frac{d^2\psi}{dx^2} = \frac{2\,m}{\hbar^2}\,[V(x) - E]\,\psi. \tag{5.2}$$

In general, the solution, $\psi(x)$, to the above equation must be *finite*, otherwise the probability density $|\psi|^2$ would become infinite (which is unphysical). Likewise, the solution must be *continuous*, otherwise the probability current (4.19) would become infinite (which is also unphysical).

5.2 Infinite Potential Well

Consider a particle of mass m and energy E moving in the following simple potential:

$$V(x) = \left\{\begin{array}{lll} 0 & \text{for } 0 \leq x \leq a \\ \infty & \text{otherwise} \end{array}\right. . \tag{5.3}$$

It follows from Eq. (5.2) that if $d^2\psi/dx^2$ (and, hence, ψ) is to remain finite then ψ must go to zero in regions where the potential is infinite. Hence, $\psi = 0$ in the regions $x \leq 0$ and $x \geq a$. Evidently, the problem is equivalent to that of a particle trapped in a one-dimensional box of length a. The boundary conditions on ψ in the region $0 < x < a$ are

$$\psi(0) = \psi(a) = 0. \tag{5.4}$$

Furthermore, it follows from Eq. (5.2) that ψ satisfies

$$\frac{d^2\psi}{dx^2} = -k^2\,\psi \tag{5.5}$$

in this region, where

$$k^2 = \frac{2\,m\,E}{\hbar^2}. \tag{5.6}$$

Here, we are assuming that $E > 0$. It is easily demonstrated that there are no solutions with $E < 0$ which are capable of satisfying the boundary conditions (5.4).

The solution to Eq. (5.5), subject to the boundary conditions (5.4), is

$$\psi_n(x) = A_n \sin(k_n x), \tag{5.7}$$

where the A_n are arbitrary (real) constants, and

$$k_n = \frac{n\pi}{a}, \tag{5.8}$$

for $n = 1, 2, 3, \cdots$. Now, it can be seen from Eqs. (5.6) and (5.8) that the energy E is only allowed to take certain discrete values: *i.e.*,

$$E_n = \frac{n^2\pi^2\hbar^2}{2\,m\,a^2}. \tag{5.9}$$

In other words, the eigenvalues of the energy operator are *discrete*. This is a general feature of *bounded* solutions: *i.e.*, solutions in which $|\psi| \rightarrow 0$ as $|x| \rightarrow \infty$. According to the discussion in Sect. 4.12, we expect the stationary eigenfunctions $\psi_n(x)$ to satisfy the orthonormality constraint

$$\int_0^a \psi_n(x)\,\psi_m(x)\,dx = \delta_{nm}. \tag{5.10}$$

It is easily demonstrated that this is the case, provided $A_n = \sqrt{2/a}$. Hence,

$$\psi_n(x) = \sqrt{\frac{2}{a}}\,\sin\left(n\pi\frac{x}{a}\right) \tag{5.11}$$

for $n = 1, 2, 3, \cdots$.

Finally, again from Sect. 4.12, the general *time-dependent* solution can be written as a linear superposition of stationary solutions:

$$\psi(x,t) = \sum_{n=0,\infty} c_n\,\psi_n(x)\,e^{-i\,E_n\,t/\hbar}, \tag{5.12}$$

where

$$c_n = \int_0^a \psi_n(x)\,\psi(x,0)\,dx. \tag{5.13}$$

5.3 Square Potential Barrier

Consider a particle of mass m and energy $E > 0$ interacting with the simple square potential barrier

$$V(x) = \left\{ \begin{array}{ll} V_0 & \text{for } 0 \leq x \leq a \\ 0 & \text{otherwise} \end{array} \right., \tag{5.14}$$

where $V_0 > 0$. In the regions to the left and to the right of the barrier, $\psi(x)$ satisfies

$$\frac{d^2\psi}{dx^2} = -k^2\,\psi, \tag{5.15}$$

where k is given by Eq. (5.6).

Let us adopt the following solution of the above equation to the left of the barrier (*i.e.*, $x < 0$):

$$\psi(x) = e^{i\,k\,x} + R\,e^{-i\,k\,x}. \tag{5.16}$$

This solution consists of a plane wave of unit amplitude traveling to the right [since the time-dependent wavefunction is multiplied by $\exp(-i\,\omega\,t)$, where $\omega = E/\hbar > 0$], and a plane wave of complex amplitude R traveling to the left. We interpret the first plane wave as an *incoming particle* (or, rather, a stream of incoming particles), and the second as a particle (or stream of particles) *reflected* by the potential barrier. Hence, $|R|^2$ is the probability of reflection. This can be seen by calculating the probability current (4.19) in the region $x < 0$, which takes the form

$$j_l = v\,(1 - |R|^2), \tag{5.17}$$

where $v = p/m = \hbar\,k/m$ is the classical particle velocity.

Let us adopt the following solution to Eq. (5.15) to the right of the barrier (*i.e.* $x > a$):

$$\psi(x) = T\,e^{i\,k\,x}. \tag{5.18}$$

This solution consists of a plane wave of complex amplitude T traveling to the right. We interpret this as a particle (or stream of particles) *transmitted* through the barrier. Hence, $|T|^2$ is the probability of transmission. The probability current in the region $x > a$ takes the form

$$j_r = v\,|T|^2. \tag{5.19}$$

Now, according to Eq. (4.35), in a stationary state (*i.e.*, $\partial|\psi|^2/\partial t = 0$), the probability current is a *spatial constant* (*i.e.*, $\partial j/\partial x = 0$). Hence, we must have $j_l = j_r$, or

$$|R|^2 + |T|^2 = 1. \tag{5.20}$$

In other words, the probabilities of reflection and transmission sum to unity, as must be the case, since reflection and transmission are the only possible outcomes for a particle incident on the barrier.

Inside the barrier (*i.e.*, $0 \leq x \leq a$), $\psi(x)$ satisfies

$$\frac{d^2\psi}{dx^2} = -q^2\,\psi, \tag{5.21}$$

where

$$q^2 = \frac{2\,m\,(E - V_0)}{\hbar^2}. \tag{5.22}$$

Let us, first of all, consider the case where $E > V_0$. In this case, the general solution to Eq. (5.21) inside the barrier takes the form

$$\psi(x) = A\,e^{i\,q\,x} + B\,e^{-i\,q\,x}, \tag{5.23}$$

where $q = \sqrt{2\,m\,(E-V_0)/\hbar^2}$.

Now, the boundary conditions at the edges of the barrier (*i.e.*, at $x = 0$ and $x = a$) are that ψ and $d\psi/dx$ are both continuous. These boundary conditions ensure that the probability current (4.19) remains finite and continuous across the edges of the boundary, as must be the case if it is to be a spatial constant.

Continuity of ψ and $d\psi/dx$ at the left edge of the barrier (*i.e.*, $x = 0$) yields

$$1 + R = A + B, \tag{5.24}$$

$$k\,(1-R) = q\,(A-B). \tag{5.25}$$

Likewise, continuity of ψ and $d\psi/dx$ at the right edge of the barrier (*i.e.*, $x = a$) gives

$$A\,e^{i\,q\,a} + B\,e^{-i\,q\,a} = T\,e^{i\,k\,a}, \tag{5.26}$$

$$q\left(A\,e^{i\,q\,a} - B\,e^{-i\,q\,a}\right) = k\,T\,e^{i\,k\,a}. \tag{5.27}$$

After considerable algebra, the above four equations yield

$$|R|^2 = \frac{(k^2-q^2)^2\,\sin^2(q\,a)}{4\,k^2\,q^2 + (k^2-q^2)^2\,\sin^2(q\,a)}, \tag{5.28}$$

and

$$|T|^2 = \frac{4\,k^2\,q^2}{4\,k^2\,q^2 + (k^2-q^2)^2\,\sin^2(q\,a)}. \tag{5.29}$$

Note that the above two expression satisfy the constraint (5.20).

It is instructive to compare the quantum mechanical probabilities of reflection and transmission—(5.28) and (5.29), respectively—with those derived from classical physics. Now, according to classical physics, if a particle of energy E is incident on a potential barrier of height $V_0 < E$ then the particle slows down as it passes through the barrier, but is otherwise unaffected. In other words, the classical probability of reflection is *zero*, and the classical probability of transmission is *unity*.

The reflection and transmission probabilities obtained from Eqs. (5.28) and (5.29), respectively, are plotted in Figs. 5.1 and 5.2. It can be seen, from Fig. 5.1, that the classical result, $|R|^2 = 0$ and $|T|^2 = 1$, is obtained in the limit where the height of the barrier is relatively small (*i.e.*, $V_0 \ll E$). However, when V_0 is of order E, there is a substantial probability that the incident particle will be *reflected* by the barrier. According to classical physics, reflection is impossible when $V_0 < E$.

It can also be seen, from Fig. 5.2, that at certain barrier widths the probability of reflection goes to *zero*. It turns out that this is true irrespective of the energy of the incident particle. It is evident, from Eq. (5.28), that these special barrier widths correspond to

$$q\,a = n\,\pi, \tag{5.30}$$

where $n = 1, 2, 3, \cdots$. In other words, the special barriers widths are integer multiples of half the de Broglie wavelength of the particle *inside* the barrier. There is no reflection at

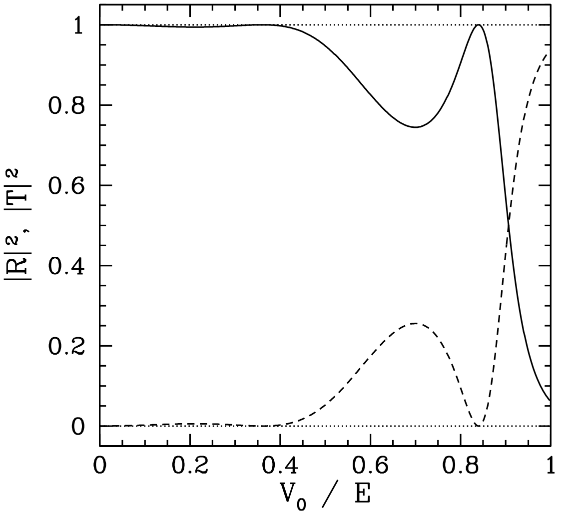

Figure 5.1: *Transmission (solid-curve) and reflection (dashed-curve) probabilities for a square potential barrier of width* $a = 1.25\,\lambda$, *where* λ *is the free-space de Broglie wavelength, as a function of the ratio of the height of the barrier,* V_0, *to the energy,* E, *of the incident particle.*

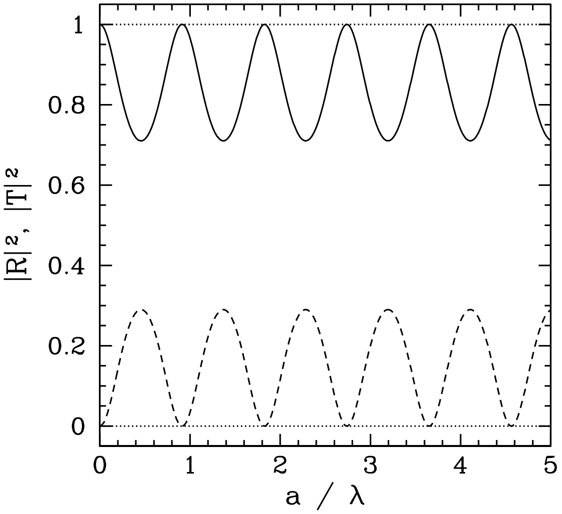

Figure 5.2: *Transmission (solid-curve) and reflection (dashed-curve) probabilities for a particle of energy* E *incident on a square potential barrier of height* $V_0 = 0.75\,E$, *as a function of the ratio of the width of the barrier,* a, *to the free-space de Broglie wavelength,* λ.

the special barrier widths because, at these widths, the backward traveling wave reflected from the left edge of the barrier interferes destructively with the similar wave reflected from the right edge of the barrier to give zero net reflected wave.

Let us, now, consider the case $E < V_0$. In this case, the general solution to Eq. (5.21) inside the barrier takes the form

$$\psi(x) = A\,e^{q\,x} + B\,e^{-q\,x}, \tag{5.31}$$

where $q = \sqrt{2\,m\,(V_0 - E)/\hbar^2}$. Continuity of ψ and $d\psi/dx$ at the left edge of the barrier (*i.e.*, $x = 0$) yields

$$1 + R = A + B, \tag{5.32}$$

$$i\,k\,(1 - R) = q\,(A - B). \tag{5.33}$$

Likewise, continuity of ψ and $d\psi/dx$ at the right edge of the barrier (*i.e.*, $x = a$) gives

$$A\,e^{q\,a} + B\,e^{-q\,a} = T\,e^{i\,k\,a}, \tag{5.34}$$

$$q\,(A\,e^{q\,a} - B\,e^{-q\,a}) = i\,k\,T\,e^{i\,k\,a}. \tag{5.35}$$

After considerable algebra, the above four equations yield

$$|R|^2 = \frac{(k^2 + q^2)^2\,\sinh^2(q\,a)}{4\,k^2\,q^2 + (k^2 + q^2)^2\,\sinh^2(q\,a)}, \tag{5.36}$$

and

$$|T|^2 = \frac{4\,k^2\,q^2}{4\,k^2\,q^2 + (k^2 + q^2)^2\,\sinh^2(q\,a)}. \tag{5.37}$$

These expressions can also be obtained from Eqs. (5.28) and (5.29) by making the substitution $q \rightarrow -i\,q$. Note that Eqs. (5.36) and (5.37) satisfy the constraint (5.20).

It is again instructive to compare the quantum mechanical probabilities of reflection and transmission—(5.36) and (5.37), respectively—with those derived from classical physics. Now, according to classical physics, if a particle of energy E is incident on a potential barrier of height $V_0 > E$ then the particle is reflected. In other words, the classical probability of reflection is *unity*, and the classical probability of transmission is *zero*.

The reflection and transmission probabilities obtained from Eqs. (5.36) and (5.37), respectively, are plotted in Figs. 5.3 and 5.4. It can be seen, from Fig. 5.3, that the classical result, $|R|^2 = 1$ and $|T|^2 = 0$, is obtained for relatively thin barriers (*i.e.*, $q\,a \sim 1$) in the limit where the height of the barrier is relatively large (*i.e.*, $V_0 \gg E$). However, when V_0 is of order E, there is a substantial probability that the incident particle will be *transmitted* by the barrier. According to classical physics, transmission is impossible when $V_0 > E$.

It can also be seen, from Fig. 5.4, that the transmission probability decays *exponentially* as the width of the barrier increases. Nevertheless, even for very wide barriers (*i.e.*, $q\,a \gg 1$), there is a small but *finite* probability that a particle incident on the barrier will be *transmitted*. This phenomenon, which is inexplicable within the context of classical physics, is called *tunneling*.

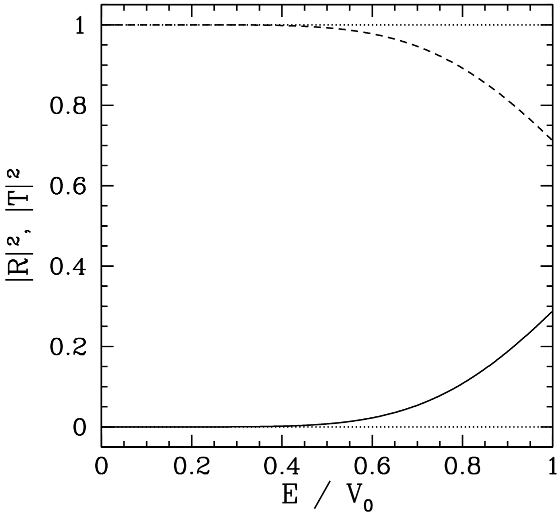

Figure 5.3: *Transmission (solid-curve) and reflection (dashed-curve) probabilities for a square potential barrier of width* $a = 0.5\,\lambda$, *where* λ *is the free-space de Broglie wavelength, as a function of the ratio of the energy,* E, *of the incoming particle to the height,* V_0, *of the barrier.*

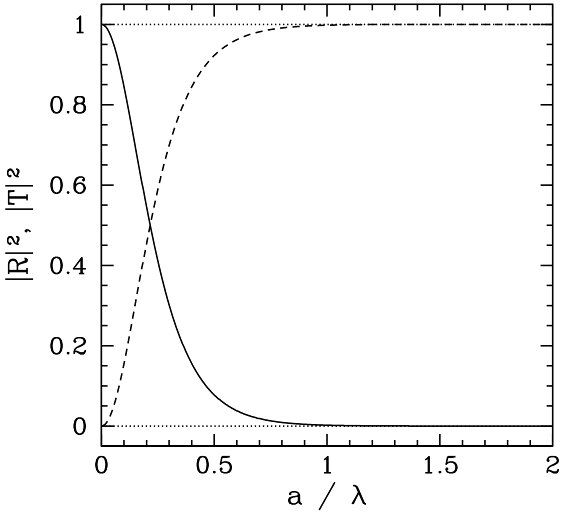

Figure 5.4: *Transmission (solid-curve) and reflection (dashed-curve) probabilities for a particle of energy* E *incident on a square potential barrier of height* $V_0 = (4/3)\,E$, *as a function of the ratio of the width of the barrier,* a, *to the free-space de Broglie wavelength,* λ.

5.4 WKB Approximation

Consider a particle of mass m and energy $E > 0$ moving through some *slowly varying* potential $V(x)$. The particle's wavefunction satisfies

$$\frac{d^2\psi(x)}{dx^2} = -k^2(x)\,\psi(x), \tag{5.38}$$

where

$$k^2(x) = \frac{2\,m\,[E - V(x)]}{\hbar^2}. \tag{5.39}$$

Let us try a solution to Eq. (5.38) of the form

$$\psi(x) = \psi_0\,\exp\left(\int_0^x i\,k(x')\,dx'\right), \tag{5.40}$$

where ψ_0 is a complex constant. Note that this solution represents a particle propagating in the positive x-direction [since the full wavefunction is multiplied by $\exp(-i\,\omega\,t)$, where $\omega = E/\hbar > 0$] with the continuously varying wavenumber $k(x)$. It follows that

$$\frac{d\psi(x)}{dx} = i\,k(x)\,\psi(x), \tag{5.41}$$

and

$$\frac{d^2\psi(x)}{dx^2} = i\,k'(x)\,\psi(x) - k^2(x)\,\psi(x), \tag{5.42}$$

where $k' \equiv dk/dx$. A comparison of Eqs. (5.38) and (5.42) reveals that Eq. (5.40) represents an approximate solution to Eq. (5.38) provided that the first term on its right-hand side is negligible compared to the second. This yields the validity criterion $|k'| \ll k^2$, or

$$\frac{k}{|k'|} \gg k^{-1}. \tag{5.43}$$

In other words, the variation length-scale of $k(x)$, which is approximately the same as the variation length-scale of $V(x)$, must be *much greater* than the particle's de Broglie wavelength (which is of order k^{-1}). Let us suppose that this is the case. Incidentally, the approximation involved in dropping the first term on the right-hand side of Eq. (5.42) is generally known as the *WKB approximation*.[1] Similarly, Eq. (5.40) is termed a WKB solution.

According to the WKB solution (5.40), the probability density remains constant: *i.e.*,

$$|\psi(x)|^2 = |\psi_0|^2, \tag{5.44}$$

as long as the particle moves through a region in which $E > V(x)$, and $k(x)$ is consequently real (*i.e.*, an allowed region according to classical physics). Suppose, however, that the

---

[1] After G. Wentzel, H.A. Kramers, and L. Brillouin.

particle encounters a potential barrier (*i.e.*, a region from which the particle is excluded according to classical physics). By definition, $E < V(x)$ inside such a barrier, and $k(x)$ is consequently imaginary. Let the barrier extend from $x = x_1$ to x_2, where $0 < x_1 < x_2$. The WKB solution inside the barrier is written

$$\psi(x) = \psi_1 \exp\left(-\int_{x_1}^{x} |k(x')|\, dx'\right), \tag{5.45}$$

where

$$\psi_1 = \psi_0 \exp\left(\int_0^{x_1} i\, k(x')\, dx'\right). \tag{5.46}$$

Here, we have neglected the unphysical exponentially growing solution.

According to the WKB solution (5.45), the probability density *decays exponentially* inside the barrier: *i.e.*,

$$|\psi(x)|^2 = |\psi_1|^2 \exp\left(-2\int_{x_1}^{x} |k(x')|\, dx'\right), \tag{5.47}$$

where $|\psi_1|^2$ is the probability density at the left-hand side of the barrier (*i.e.*, $x = x_1$). It follows that the probability density at the right-hand side of the barrier (*i.e.*, $x = x_2$) is

$$|\psi_2|^2 = |\psi_1|^2 \exp\left(-2\int_{x_1}^{x_2} |k(x')|\, dx'\right). \tag{5.48}$$

Note that $|\psi_2|^2 < |\psi_1|^2$. Of course, in the region to the right of the barrier (*i.e.*, $x > x_2$), the probability density takes the constant value $|\psi_2|^2$.

We can interpret the ratio of the probability densities to the right and to the left of the potential barrier as the probability, $|T|^2$, that a particle incident from the left will tunnel through the barrier and emerge on the other side: *i.e.*,

$$|T|^2 = \frac{|\psi_2|^2}{|\psi_1|^2} = \exp\left(-2\int_{x_1}^{x_2} |k(x')|\, dx'\right) \tag{5.49}$$

(see Sect. 5.3). It is easily demonstrated that the probability of a particle incident from the right tunneling through the barrier is the same.

Note that the criterion (5.43) for the validity of the WKB approximation implies that the above transmission probability is *very small*. Hence, the WKB approximation only applies to situations in which there is very little chance of a particle tunneling through the potential barrier in question. Unfortunately, the validity criterion (5.43) breaks down completely at the edges of the barrier (*i.e.*, at $x = x_1$ and x_2), since $k(x) = 0$ at these points. However, it can be demonstrated that the contribution of those regions, around $x = x_1$ and x_2, in which the WKB approximation breaks down to the integral in Eq. (5.49) is fairly negligible. Hence, the above expression for the tunneling probability is a reasonable approximation provided that the incident particle's de Broglie wavelength is much smaller than the spatial extent of the potential barrier.

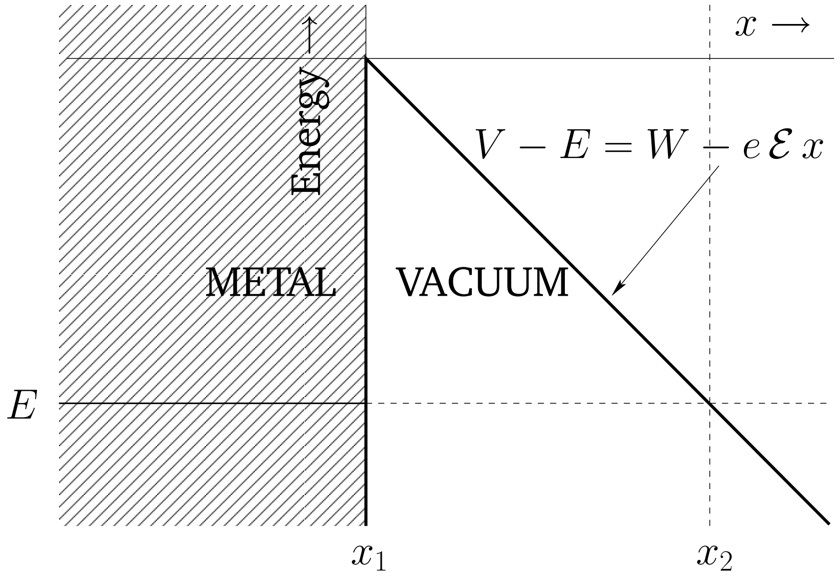

Figure 5.5: *The potential barrier for an electron in a metal surface subject to an external electric field.*

5.5 Cold Emission

Suppose that an unheated metal surface is subject to a large uniform external electric field of strength $\mathcal{E}$, which is directed such that it accelerates electrons *away* from the surface. We have already seen (in Sect. 3.6) that electrons just below the surface of a metal can be regarded as being in a potential well of depth W, where W is called the *work function* of the surface. Adopting a simple one-dimensional treatment of the problem, let the metal lie at $x < 0$, and the surface at $x = 0$. Now, the applied electric field is shielded from the interior of the metal. Hence, the energy, E, say, of an electron just below the surface is unaffected by the field. In the absence of the electric field, the potential barrier just above is the surface is simply $V(x) - E = W$. The electric field modifies this to $V(x) - E = W - e\,\mathcal{E}\,x$. The potential barrier is sketched in Fig. 5.5.

It can be seen, from Fig. 5.5, that an electron just below the surface of the metal is confined by a triangular potential barrier which extends from $x = x_1$ to x_2, where $x_1 = 0$ and $x_2 = W/e\,\mathcal{E}$. Making use of the WKB approximation (see the previous subsection), the probability of such an electron tunneling through the barrier, and consequently being emitted from the surface, is

$$|T|^2 = \exp\left(-\frac{2\sqrt{2\,m}}{\hbar}\int_{x_1}^{x_2}\sqrt{V(x) - E}\,dx\right), \tag{5.50}$$

or

$$|T|^2 = \exp\left(-\frac{2\sqrt{2\,m}}{\hbar}\int_{0}^{W/e\,\mathcal{E}}\sqrt{W - e\,\mathcal{E}\,x}\,dx\right). \tag{5.51}$$

This reduces to

$$|T|^2 = \exp\left(-2\sqrt{2}\,\frac{m^{1/2}W^{3/2}}{\hbar\, e\,\mathcal{E}}\int_0^1\sqrt{1-y}\,dy\right), \tag{5.52}$$

or

$$|T|^2 = \exp\left(-\frac{4\sqrt{2}}{3}\,\frac{m^{1/2}W^{3/2}}{\hbar\, e\,\mathcal{E}}\right). \tag{5.53}$$

The above result is known as the *Fowler-Nordheim* formula. Note that the probability of emission increases *exponentially* as the electric field-strength above the surface of the metal increases.

The cold emission of electrons from a metal surface is the basis of an important device known as a *scanning tunneling microscope*, or an STM. An STM consists of a very sharp conducting probe which is scanned over the surface of a metal (or any other solid conducting medium). A large voltage difference is applied between the probe and the surface. Now, the surface electric field-strength immediately below the probe tip is proportional to the applied potential difference, and inversely proportional to the spacing between the tip and the surface. Electrons tunneling between the surface and the probe tip give rise to a weak electric current. The magnitude of this current is proportional to the tunneling probability (5.53). It follows that the current is an *extremely sensitive* function of the surface electric field-strength, and, hence, of the spacing between the tip and the surface (assuming that the potential difference is held constant). An STM can thus be used to construct a very accurate contour map of the surface under investigation. In fact, STMs are capable of achieving sufficient resolution to image individual atoms

5.6 Alpha Decay

Many types of heavy atomic nucleus spontaneously decay to produce daughter nucleii via the emission of α-particles (*i.e.*, helium nucleii) of some characteristic energy. This process is know as α-decay. Let us investigate the α-decay of a particular type of atomic nucleus of radius R, charge-number Z, and mass-number A. Such a nucleus thus decays to produce a daughter nucleus of charge-number $Z_1 = Z-2$ and mass-number $A_1 = A-4$, and an α-particle of charge-number $Z_2 = 2$ and mass-number $A_2 = 4$. Let the characteristic energy of the α-particle be E. Incidentally, nuclear radii are found to satisfy the empirical formula

$$R = 1.5\times 10^{-15}\,A^{1/3}\,\mathrm{m} = 2.0\times 10^{-15}\,Z_1^{1/3}\,\mathrm{m} \tag{5.54}$$

for $Z \gg 1$.

In 1928, George Gamov proposed a very successful theory of α-decay, according to which the α-particle moves freely inside the nucleus, and is emitted after *tunneling* through the potential barrier between itself and the daughter nucleus. In other words, the α-particle, whose energy is E, is trapped in a potential well of radius R by the potential barrier

$$V(r) = \frac{Z_1\,Z_2\,e^2}{4\pi\,\epsilon_0\,r} \tag{5.55}$$

for $r > R$.

Making use of the WKB approximation (and neglecting the fact that r is a radial, rather than a Cartesian, coordinate), the probability of the α-particle tunneling through the barrier is

$$|T|^2 = \exp\left(-\frac{2\sqrt{2\,m}}{\hbar}\int_{r_1}^{r_2}\sqrt{V(r)-E\,dr}\right), \tag{5.56}$$

where $r_1 = R$ and $r_2 = Z_1\,Z_2\,e^2/(4\pi\,\epsilon_0\,E)$. Here, $m = 4\,m_p$ is the α-particle mass. The above expression reduces to

$$|T|^2 = \exp\left(-2\,\sqrt{2}\,\beta\int_1^{E_c/E}\left[\frac{1}{y}-\frac{E}{E_c}\right]^{1/2} dy\right), \tag{5.57}$$

where

$$\beta = \left(\frac{Z_1\,Z_2\,e^2\,m\,R}{4\pi\,\epsilon_0\,\hbar^2}\right)^{1/2} = 0.74\,Z_1^{2/3} \tag{5.58}$$

is a dimensionless constant, and

$$E_c = \frac{Z_1\,Z_2\,e^2}{4\pi\,\epsilon_0\,R} = 1.44\,Z_1^{2/3}\text{ MeV} \tag{5.59}$$

is the characteristic energy the α-particle would need in order to escape from the nucleus without tunneling. Of course, $E \ll E_c$. It is easily demonstrated that

$$\int_1^{1/\epsilon}\left[\frac{1}{y}-\epsilon\right]^{1/2} dy \simeq \frac{\pi}{2\,\sqrt{\epsilon}} - 2 \tag{5.60}$$

when $\epsilon \ll 1$. Hence.

$$|T|^2 \simeq \exp\left(-2\,\sqrt{2}\,\beta\left[\frac{\pi}{2}\sqrt{\frac{E_c}{E}} - 2\right]\right). \tag{5.61}$$

Now, the α-particle moves inside the nucleus with the characteristic velocity $v = \sqrt{2\,E/m}$. It follows that the particle bounces backward and forward within the nucleus at the frequency $\nu \simeq v/R$, giving

$$\nu \simeq 2\times 10^{28}\text{ yr}^{-1} \tag{5.62}$$

for a 1 MeV α-particle trapped inside a typical heavy nucleus of radius 10^{-14} m. Thus, the α-particle effectively attempts to tunnel through the potential barrier ν times a second. If each of these attempts has a probability $|T|^2$ of succeeding, then the probability of decay per unit time is $\nu\,|T|^2$. Hence, if there are $N(t) \gg 1$ undecayed nuclii at time t then there are only $N + dN$ at time $t + dt$, where

$$dN = -N\,\nu\,|T|^2\,dt. \tag{5.63}$$

This expression can be integrated to give

$$N(t) = N(0)\,\exp(-\nu\,|T|^2\,t). \tag{5.64}$$

Now, the *half-life*, τ, is defined as the time which must elapse in order for half of the nuclii originally present to decay. It follows from the above formula that

$$\tau = \frac{\ln 2}{\nu\,|T|^2}. \tag{5.65}$$

Note that the half-life is *independent* of $N(0)$.

Finally, making use of the above results, we obtain

$$\log_{10}[\tau(\mathrm{yr})] = -C_1 - C_2\,Z_1^{2/3} + C_3\,\frac{Z_1}{\sqrt{E(\mathrm{MeV})}}, \tag{5.66}$$

where

$$C_1 = 28.5, \tag{5.67}$$

$$C_2 = 1.83, \tag{5.68}$$

$$C_3 = 1.73. \tag{5.69}$$

The half-life, τ, the daughter charge-number, $Z_1 = Z - 2$, and the α-particle energy, E, for atomic nucleii which undergo α-decay are indeed found to satisfy a relationship of the form (5.66). The best fit to the data (see Fig. 5.6) is obtained using

$$C_1 = 28.9, \tag{5.70}$$

$$C_2 = 1.60, \tag{5.71}$$

$$C_3 = 1.61. \tag{5.72}$$

Note that these values are remarkably similar to those calculated above.

5.7 Square Potential Well

Consider a particle of mass m and energy E interacting with the simple square potential well

$$V(x) = \begin{cases} -V_0 & \text{for } -a/2 \leq x \leq a/2 \\ 0 & \text{otherwise} \end{cases}, \tag{5.73}$$

where $V_0 > 0$.

Now, if $E > 0$ then the particle is unbounded. Thus, when the particle encounters the well it is either reflected or transmitted. As is easily demonstrated, the reflection and

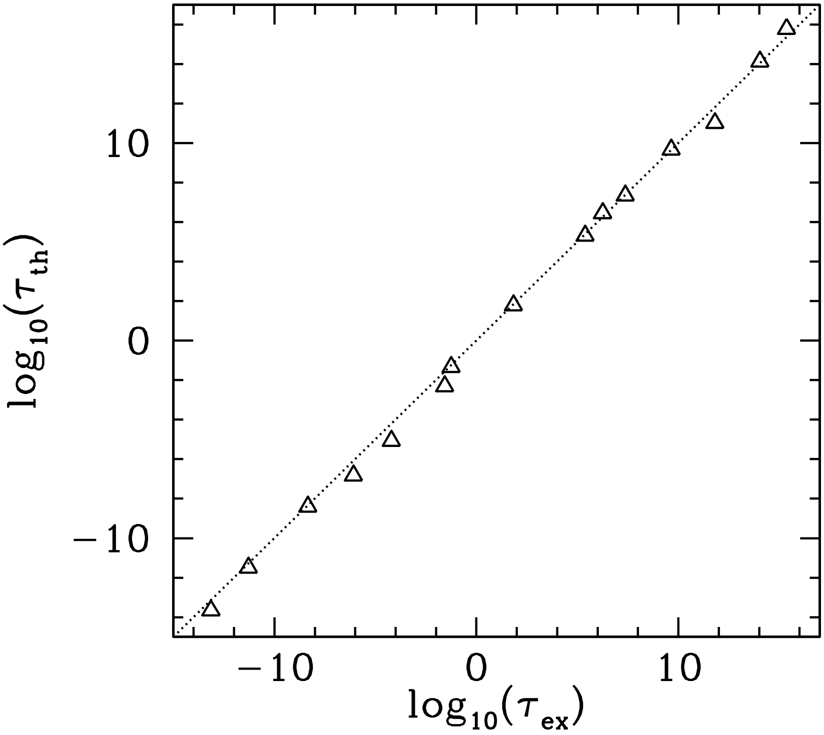

Figure 5.6: *The experimentally determined half-life,* τ_{ex}*, of various atomic nucleii which decay via* α *emission versus the best-fit theoretical half-life* $\log_{10}(\tau_{th}) = -28.9 - 1.60\, Z_1^{2/3} + 1.61\, Z_1/\sqrt{E}$. *Both half-lives are measured in years. Here,* $Z_1 = Z - 2$*, where* Z *is the charge number of the nucleus, and* E *the characteristic energy of the emitted* α*-particle in MeV. In order of increasing half-life, the points correspond to the following nucleii: Rn 215, Po 214, Po 216, Po 197, Fm 250, Ac 225, U 230, U 232, U 234, Gd 150, U 236, U 238, Pt 190, Gd 152, Nd 144. Data obtained from IAEA Nuclear Data Centre.*

transmission probabilities are given by Eqs. (5.28) and (5.29), respectively, where

$$k^2 = \frac{2\,m\,E}{\hbar^2}, \tag{5.74}$$

$$q^2 = \frac{2\,m\,(E+V_0)}{\hbar^2}. \tag{5.75}$$

Suppose, however, that $E < 0$. In this case, the particle is bounded (*i.e.*, $|\psi|^2 \rightarrow 0$ as $|x| \rightarrow \infty$). Is is possible to find bounded solutions of Schrödinger's equation in the finite square potential well (5.73)?

Now, it is easily seen that independent solutions of Schrödinger's equation (5.2) in the symmetric [*i.e.*, $V(-x) = V(x)$] potential (5.73) must be either totally symmetric [*i.e.*, $\psi(-x) = \psi(x)$], or totally anti-symmetric [*i.e.*, $\psi(-x) = -\psi(x)$]. Moreover, the solutions must satisfy the boundary condition

$$\psi \rightarrow 0 \qquad\qquad \text{as } |x| \rightarrow \infty. \tag{5.76}$$

Let us, first of all, search for a totally symmetric solution. In the region to the left of the well (*i.e.* $x < -a/2$), the solution of Schrödinger's equation which satisfies the boundary condition $\psi \rightarrow 0$ and $x \rightarrow -\infty$ is

$$\psi(x) = A\,e^{k\,x}, \tag{5.77}$$

where

$$k^2 = \frac{2\,m\,|E|}{\hbar^2}. \tag{5.78}$$

By symmetry, the solution in the region to the right of the well (*i.e.*, $x > a/2$) is

$$\psi(x) = A\,e^{-k\,x}. \tag{5.79}$$

The solution inside the well (*i.e.*, $|x| \leq a/2$) which satisfies the symmetry constraint $\psi(-x) = \psi(x)$ is

$$\psi(x) = B\,\cos(q\,x), \tag{5.80}$$

where

$$q^2 = \frac{2\,m\,(V_0+E)}{\hbar^2}. \tag{5.81}$$

Here, we have assumed that $E > -V_0$. The constraint that $\psi(x)$ and its first derivative be continuous at the edges of the well (*i.e.*, at $x = \pm a/2$) yields

$$k = q\,\tan(q\,a/2). \tag{5.82}$$

Let $y = q\,a/2$. It follows that

$$E = E_0\,y^2 - V_0, \tag{5.83}$$

where

$$E_0 = \frac{2\,\hbar^2}{m\,a^2}. \tag{5.84}$$

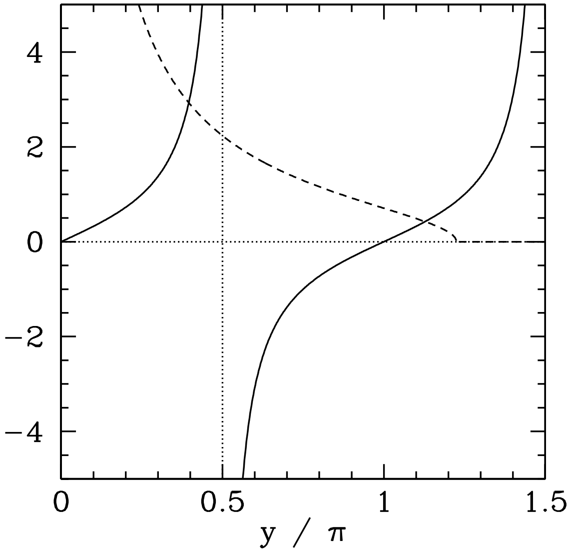

Figure 5.7: *The curves* $\tan y$ *(solid) and* $\sqrt{\lambda - y^2}/y$ *(dashed), calculated for* $\lambda = 1.5\,\pi^2$. *The latter curve takes the value* 0 *when* $y > \sqrt{\lambda}$.

Moreover, Eq. (5.82) becomes

$$\frac{\sqrt{\lambda - y^2}}{y} = \tan y, \tag{5.85}$$

with

$$\lambda = \frac{V_0}{E_0}. \tag{5.86}$$

Here, y must lie in the range $0 < y < \sqrt{\lambda}$: *i.e.*, E must lie in the range $-V_0 < E < 0$.

Now, the solutions to Eq. (5.85) correspond to the intersection of the curve $\sqrt{\lambda - y^2}/y$ with the curve $\tan y$. Figure 5.7 shows these two curves plotted for a particular value of λ. In this case, the curves intersect twice, indicating the existence of two totally symmetric bound states in the well. Moreover, it is evident, from the figure, that as λ increases (*i.e.*, as the well becomes deeper) there are more and more bound states. However, it is also evident that there is always at least one totally symmetric bound state, no matter how small λ becomes (*i.e.*, no matter how shallow the well becomes). In the limit $\lambda \gg 1$ (*i.e.*, the limit in which the well becomes very deep), the solutions to Eq. (5.85) asymptote to the roots of $\tan y = \infty$. This gives $y = (2\,j - 1)\,\pi/2$, where j is a positive integer, or

$$q = \frac{(2\,j - 1)\,\pi}{a}. \tag{5.87}$$

These solutions are equivalent to the odd-n infinite square well solutions specified by Eq. (5.8).

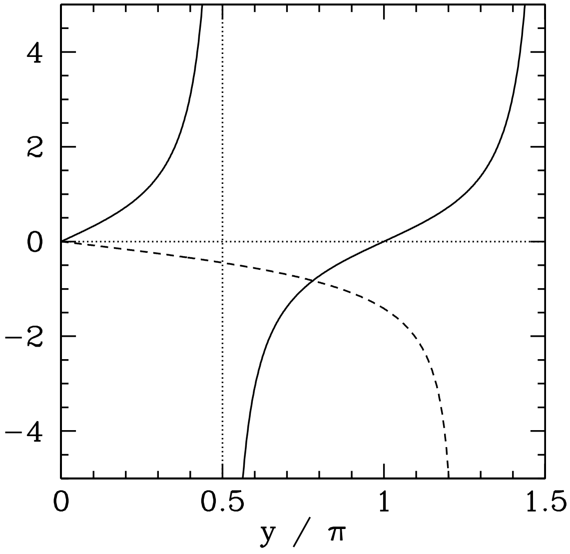

Figure 5.8: *The curves* $\tan y$ *(solid) and* $-y/\sqrt{\lambda - y^2}$ *(dashed), calculated for* $\lambda = 1.5\,\pi^2$.

For the case of a totally anti-symmetric bound state, similar analysis to the above yields

$$-\frac{y}{\sqrt{\lambda - y^2}} = \tan y. \tag{5.88}$$

The solutions of this equation correspond to the intersection of the curve $\tan y$ with the curve $-y/\sqrt{\lambda - y^2}$. Figure 5.8 shows these two curves plotted for the same value of λ as that used in Fig. 5.7. In this case, the curves intersect once, indicating the existence of a single totally anti-symmetric bound state in the well. It is, again, evident, from the figure, that as λ increases (*i.e.*, as the well becomes deeper) there are more and more bound states. However, it is also evident that when λ becomes sufficiently small [*i.e.*, $\lambda < (\pi/2)^2$] then there is no totally anti-symmetric bound state. In other words, a very shallow potential well always possesses a totally symmetric bound state, but does not generally possess a totally anti-symmetric bound state. In the limit $\lambda \gg 1$ (*i.e.*, the limit in which the well becomes very deep), the solutions to Eq. (5.88) asymptote to the roots of $\tan y = 0$. This gives $y = j\,\pi$, where j is a positive integer, or

$$q = \frac{2\,j\,\pi}{a}. \tag{5.89}$$

These solutions are equivalent to the even-n infinite square well solutions specified by Eq. (5.8).

5.8 Simple Harmonic Oscillator

The classical Hamiltonian of a simple harmonic oscillator is

$$H = \frac{p^2}{2\,m} + \frac{1}{2}\,K\,x^2, \tag{5.90}$$

where $K > 0$ is the so-called force constant of the oscillator. Assuming that the quantum mechanical Hamiltonian has the same form as the classical Hamiltonian, the time-independent Schrödinger equation for a particle of mass m and energy E moving in a simple harmonic potential becomes

$$\frac{d^2\psi}{dx^2} = \frac{2\,m}{\hbar^2}\left(\frac{1}{2}\,K\,x^2 - E\right)\psi. \tag{5.91}$$

Let $\omega = \sqrt{K/m}$, where ω is the oscillator's classical angular frequency of oscillation. Furthermore, let

$$y = \sqrt{\frac{m\,\omega}{\hbar}}\,x, \tag{5.92}$$

and

$$\epsilon = \frac{2\,E}{\hbar\,\omega}. \tag{5.93}$$

Equation (5.91) reduces to

$$\frac{d^2\psi}{dy^2} - (y^2 - \epsilon)\,\psi = 0. \tag{5.94}$$

We need to find solutions to the above equation which are bounded at infinity: *i.e.*, solutions which satisfy the boundary condition $\psi \rightarrow 0$ as $|y| \rightarrow \infty$.

Consider the behavior of the solution to Eq. (5.94) in the limit $|y| \gg 1$. As is easily seen, in this limit the equation simplifies somewhat to give

$$\frac{d^2\psi}{dy^2} - y^2\,\psi \simeq 0. \tag{5.95}$$

The approximate solutions to the above equation are

$$\psi(y) \simeq A(y)\,e^{\pm y^2/2}, \tag{5.96}$$

where $A(y)$ is a relatively slowly varying function of y. Clearly, if $\psi(y)$ is to remain bounded as $|y| \rightarrow \infty$ then we must chose the exponentially decaying solution. This suggests that we should write

$$\psi(y) = h(y)\,e^{-y^2/2}, \tag{5.97}$$

where we would expect $h(y)$ to be an algebraic, rather than an exponential, function of y.

Substituting Eq. (5.97) into Eq. (5.94), we obtain

$$\frac{d^2h}{dy^2} - 2\,y\,\frac{dh}{dy} + (\epsilon - 1)\,h = 0. \tag{5.98}$$

Let us attempt a power-law solution of the form

$$h(y) = \sum_{i=0}^{\infty} c_i\, y^i. \tag{5.99}$$

Inserting this test solution into Eq. (5.98), and equating the coefficients of y^i, we obtain the recursion relation

$$c_{i+2} = \frac{(2\,i - \epsilon + 1)}{(i+1)\,(i+2)}\, c_i. \tag{5.100}$$

Consider the behavior of $h(y)$ in the limit $|y| \rightarrow \infty$. The above recursion relation simplifies to

$$c_{i+2} \simeq \frac{2}{i}\, c_i. \tag{5.101}$$

Hence, at large $|y|$, when the higher powers of y dominate, we have

$$h(y) \sim C \sum_j \frac{y^{2\,j}}{j!} \sim C\, e^{\,y^2}. \tag{5.102}$$

It follows that $\psi(y) = h(y)\, \exp(-y^2/2)$ varies as $\exp(\,y^2/2)$ as $|y| \rightarrow \infty$. This behavior is unacceptable, since it does not satisfy the boundary condition $\psi \rightarrow 0$ as $|y| \rightarrow \infty$. The only way in which we can prevent ψ from blowing up as $|y| \rightarrow \infty$ is to demand that the power series (5.99) *terminate* at some finite value of i. This implies, from the recursion relation (5.100), that

$$\epsilon = 2\,n + 1, \tag{5.103}$$

where n is a non-negative integer. Note that the number of terms in the power series (5.99) is $n + 1$. Finally, using Eq. (5.93), we obtain

$$E = (n + 1/2)\, \hbar\, \omega, \tag{5.104}$$

for $n = 0, 1, 2, \cdots$.

Hence, we conclude that a particle moving in a harmonic potential has quantized energy levels which are *equally spaced*. The spacing between successive energy levels is $\hbar\,\omega$, where ω is the classical oscillation frequency. Furthermore, the lowest energy state ($n = 0$) possesses the *finite* energy $(1/2)\,\hbar\,\omega$. This is sometimes called *zero-point energy*. It is easily demonstrated that the (normalized) wavefunction of the lowest energy state takes the form

$$\psi_0(x) = \frac{e^{-x^2/2\,d^2}}{\pi^{1/4}\sqrt{d}}, \tag{5.105}$$

where $d = \sqrt{\hbar/m\,\omega}$.

Let $\psi_n(x)$ be an energy eigenstate of the harmonic oscillator corresponding to the eigenvalue

$$E_n = (n + 1/2)\,\hbar\,\omega. \tag{5.106}$$

Assuming that the ψ_n are properly normalized (and real), we have

$$\int_{-\infty}^{\infty}\psi_n\,\psi_m\,dx = \delta_{nm}. \tag{5.107}$$

Now, Eq. (5.94) can be written

$$\left(-\frac{d^2}{dy^2}+y^2\right)\psi_n = (2n+1)\,\psi_n, \tag{5.108}$$

where $x = d\,y$, and $d = \sqrt{\hbar/m\,\omega}$. It is helpful to define the operators

$$a_\pm = \frac{1}{\sqrt{2}}\left(\mp\frac{d}{dy}+y\right). \tag{5.109}$$

As is easily demonstrated, these operators satisfy the commutation relation

$$[a_+, a_-] = -1. \tag{5.110}$$

Using these operators, Eq. (5.108) can also be written in the forms

$$a_+\,a_-\,\psi_n = n\,\psi_n, \tag{5.111}$$

or

$$a_-\,a_+\,\psi_n = (n+1)\,\psi_n. \tag{5.112}$$

The above two equations imply that

$$a_+\,\psi_n = \sqrt{n+1}\,\psi_{n+1}, \tag{5.113}$$

$$a_-\,\psi_n = \sqrt{n}\,\psi_{n-1}. \tag{5.114}$$

We conclude that a_+ and a_- are *raising and lowering operators*, respectively, for the harmonic oscillator: *i.e.*, operating on the wavefunction with a_+ causes the quantum number n to increase by unity, and *vice versa*. The Hamiltonian for the harmonic oscillator can be written in the form

$$H = \hbar\,\omega\,\left(a_+\,a_-+\frac{1}{2}\right), \tag{5.115}$$

from which the result

$$H\,\psi_n = (n+1/2)\,\hbar\,\omega\,\psi_n = E_n\,\psi_n \tag{5.116}$$

is readily deduced. Finally, Eqs. (5.107), (5.113), and (5.114) yield the useful expression

$$\begin{aligned}\int_{-\infty}^{\infty}\psi_m\,x\,\psi_n\,dx &= \frac{d}{\sqrt{2}}\int_{-\infty}^{\infty}\psi_m\,(a_+ + a_-)\,\psi_n\,dx \qquad (5.117)\\ &= \sqrt{\frac{\hbar}{2\,m\,\omega}}\left(\sqrt{m}\,\delta_{m,n+1}+\sqrt{n}\,\delta_{m,n-1}\right).\end{aligned}$$

Exercises

1. Show that the wavefunction of a particle of mass m in an infinite one-dimensional square-well of width a returns to its original form after a quantum revival time $T = 4\,m\,a^2/\pi\,\hbar$.

2. A particle of mass m moves freely in one dimension between impenetrable walls located at $x = 0$ and a. Its initial wavefunction is

   $$\psi(x, 0) = \sqrt{2/a}\,\sin(3\pi\,x/a).$$

   What is the subsequent time evolution of the wavefunction? Suppose that the initial wavefunction is

   $$\psi(x, 0) = \sqrt{1/a}\,\sin(\pi\,x/a)\,[1 + 2\,\cos(\pi\,x/a)].$$

   What now is the subsequent time evolution? Calculate the probability of finding the particle between 0 and $a/2$ as a function of time in each case.

3. A particle of mass m is in the ground-state of an infinite one-dimensional square-well of width a. Suddenly the well expands to twice its original size, as the right wall moves from a to $2a$, leaving the wavefunction momentarily undisturbed. The energy of the particle is now measured. What is the most probable result? What is the probability of obtaining this result? What is the next most probable result, and what is its probability of occurrence? What is the expectation value of the energy?

4. A stream of particles of mass m and energy $E > 0$ encounter a potential step of height $W(< E)$: *i.e.*, $V(x) = 0$ for $x < 0$ and $V(x) = W$ for $x > 0$ with the particles incident from $-\infty$. Show that the fraction reflected is

   $$R = \left(\frac{k-q}{k+q}\right)^2,$$

   where $k^2 = (2m/\hbar^2)\,E$ and $q^{\,2} = (2m/\hbar^2)\,(E - W)$.

5. A stream of particles of mass m and energy $E > 0$ encounter the delta-function potential $V(x) = -\alpha\,\delta(x)$, where $\alpha > 0$. Show that the fraction reflected is

   $$R = \beta^2/(1 + \beta^2),$$

   where $\beta = m\,\alpha/\hbar^2\,k$, and $k^2 = (2m/\hbar^2)\,E$. Does such a potential have a bound state? If so, what is its energy?

6. Two potential wells of width a are separated by a distance $L \gg a$. A particle of mass m and energy E is in one of the wells. Estimate the time required for the particle to tunnel to the other well.

7. Consider the half-infinite potential well

   $$V(x) = \left\{\begin{array}{ll} \infty & x \leq 0 \\ -V_0 & 0 < x < L \\ 0 & x \geq L \end{array}\right.,$$

where $V_0 > 0$. Demonstrate that the bound-states of a particle of mass m and energy $-V_0 < E < 0$ satisfy

$$\tan\left(\sqrt{2\,m\,(V_0+E)}\;L/\hbar\right) = -\sqrt{(V_0+E)/(-E)}.$$

8. Find the properly normalized first two excited energy eigenstates of the harmonic oscillator, as well as the expectation value of the potential energy in the nth energy eigenstate. Hint: Consider the raising and lowering operators $a_\pm$ defined in Eq. (5.109).

6 Multi-Particle Systems

6.1 Introduction

In this chapter, we shall extend the single particle, one-dimensional formulation of non-relativistic quantum mechanics introduced in the previous sections in order to investigate one-dimensional chapters containing *multiple* particles.

6.2 Fundamental Concepts

We have already seen that the instantaneous state of a system consisting of a single non-relativistic particle, whose position coordinate is x, is fully specified by a complex wavefunction $\psi(x,t)$. This wavefunction is interpreted as follows. The probability of finding the particle between x and $x + dx$ at time t is given by $|\psi(x,t)|^2\,dx$. This interpretation only makes sense if the wavefunction is normalized such that

$$\int_{-\infty}^{\infty} |\psi(x,t)|^2\,dx = 1 \tag{6.1}$$

at all times. The physical significance of this normalization requirement is that the probability of the particle being found anywhere on the x-axis must always be unity (which corresponds to certainty).

Consider a system containing N non-relativistic particles, labeled $i = 1, N$, moving in one dimension. Let x_i and m_i be the position coordinate and mass, respectively, of the ith particle. By analogy with the single-particle case, the instantaneous state of a multi-particle system is specified by a complex wavefunction $\psi(x_1, x_2, \ldots, x_N, t)$. The probability of finding the first particle between x_1 and $x_1 + dx_1$, the second particle between x_2 and $x_2 + dx_2$, *etc.*, at time t is given by $|\psi(x_1, x_2, \ldots, x_N, t)|^2\,dx_1\,dx_2\ldots dx_N$. It follows that the wavefunction must satisfy the normalization condition

$$\int |\psi(x_1, x_2, \ldots, x_N, t)|^2\,dx_1\,dx_2\ldots dx_N = 1 \tag{6.2}$$

at all times, where the integration is taken over all $x_1\,x_2\ldots x_N$ space.

In a single-particle system, position is represented by the algebraic operator x, whereas momentum is represented by the differential operator $-i\,\hbar\,\partial/\partial x$ (see Sect. 4.6). By analogy, in a multi-particle system, the position of the ith particle is represented by the algebraic operator x_i, whereas the corresponding momentum is represented by the differential operator

$$p_i = -i\,\hbar\,\frac{\partial}{\partial x_i}. \tag{6.3}$$

Since the x_i are *independent* variables (*i.e.*, $\partial x_i/\partial x_j = \delta_{ij}$), we conclude that the various position and momentum operators satisfy the following commutation relations:

$$
\begin{aligned}
[x_i, x_j] &= 0, && (6.4)\\
[p_i, p_j] &= 0, && (6.5)\\
[x_i, p_j] &= i\,\hbar\,\delta_{ij}. && (6.6)
\end{aligned}
$$

Now, we know, from Sect. 4.10, that two dynamical variables can only be (exactly) measured *simultaneously* if the operators which represent them in quantum mechanics *commute* with one another. Thus, it is clear, from the above commutation relations, that the only restriction on measurement in a one-dimensional multi-particle system is that it is impossible to simultaneously measure the position and momentum of the *same* particle. Note, in particular, that a knowledge of the position or momentum of a given particle does not in any way preclude a similar knowledge for a different particle. The commutation relations (6.4)–(6.6) illustrate an important point in quantum mechanics: namely, that operators corresponding to *different degrees of freedom* of a dynamical system tend to *commute* with one another. In this case, the different degrees of freedom correspond to the different motions of the various particles making up the system.

Finally, if $H(x_1, x_2, \ldots, x_N, t)$ is the Hamiltonian of the system then the multi-particle wavefunction $\psi(x_1, x_2, \ldots, x_N, t)$ satisfies the usual time-dependent Schrödinger equation [see Eq. (4.63)]

$$i\,\hbar\,\frac{\partial\psi}{\partial t} = H\,\psi. \tag{6.7}$$

Likewise, a multi-particle state of definite energy E (*i.e.*, an eigenstate of the Hamiltonian with eigenvalue E) is written (see Sect. 4.12)

$$\psi(x_1, x_2, \ldots, x_N, t) = \psi_E(x_1, x_2, \ldots, x_N)\,e^{-i\,E\,t/\hbar}, \tag{6.8}$$

where the stationary wavefunction ψ_E satisfies the time-independent Schrödinger equation [see Eq. (4.160)]

$$H\,\psi_E = E\,\psi_E. \tag{6.9}$$

Here, H is assumed not to be an explicit function of t.

6.3 Non-Interacting Particles

In general, we expect the Hamiltonian of a multi-particle system to take the form

$$H(x_1, x_2, \ldots, x_N, t) = \sum_{i=1,N} \frac{p_i^{\,2}}{2\,m_i} + V(x_1, x_2, \ldots, x_N, t). \tag{6.10}$$

Here, the first term on the right-hand side represents the total kinetic energy of the system, whereas the potential V specifies the nature of the interaction between the various particles making up the system, as well as the interaction of the particles with any external forces.

Suppose that the particles do not interact with one another. This implies that each particle moves in a common potential: *i.e.*,

$$V(x_1, x_2, \ldots, x_N, t) = \sum_{i=1,N} V(x_i, t). \tag{6.11}$$

Hence, we can write

$$H(x_1, x_2, \ldots, x_N, t) = \sum_{i=1,N} H_i(x_i, t), \tag{6.12}$$

where

$$H_i = \frac{p_i^2}{2\,m_i} + V(x_i, t). \tag{6.13}$$

In other words, for the case of non-interacting particles, the multi-particle Hamiltonian of the system can be written as the sum of N independent single-particle Hamiltonians. Here, H_i represents the energy of the ith particle, and is completely unaffected by the energies of the other particles. Furthermore, given that the various particles which make up the system are non-interacting, we expect their instantaneous positions to be completely *uncorrelated* with one another. This immediately implies that the multi-particle wavefunction $\psi(x_1, x_2, \ldots x_N, t)$ can be written as the product of N independent single-particle wavefunctions: *i.e.*,

$$\psi(x_1, x_2, \ldots, x_N, t) = \psi_1(x_1, t)\,\psi_2(x_2, t)\ldots\psi_N(x_N, t). \tag{6.14}$$

Here, $|\psi_i(x_i, t)|^2\,dx_i$ is the probability of finding the ith particle between x_i and $x_i + dx_i$ at time t. This probability is completely unaffected by the positions of the other particles. It is evident that $\psi_i(x_i, t)$ must satisfy the normalization constraint

$$\int_{-\infty}^{\infty} |\psi_i(x_i, t)|^2\,dx_i = 1. \tag{6.15}$$

If this is the case then the normalization constraint (6.2) for the multi-particle wavefunction is automatically satisfied. Equation (6.14) illustrates an important point in quantum mechanics: namely, that we can generally write the total wavefunction of a many degree of freedom system as a product of different wavefunctions corresponding to each degree of freedom.

According to Eqs. (6.12) and (6.14), the time-dependent Schrödinger equation (6.7) for a system of N non-interacting particles factorizes into N independent equations of the form

$$i\,\hbar\,\frac{\partial\psi_i}{\partial t} = H_i\,\psi_i. \tag{6.16}$$

Assuming that $V(x, t) \equiv V(x)$, the time-independent Schrödinger equation (6.9) also factorizes to give

$$H_i\,\psi_{E_i} = E_i\,\psi_{E_i}, \tag{6.17}$$

where $\psi_i(x_i, t) = \psi_{E_i}(x_i)\,\exp(-i\,E_i\,t/\hbar)$, and E_i is the energy of the ith particle. Hence, a multi-particle state of definite energy E has a wavefunction of the form

$$\psi(x_1, x_2, \ldots, x_n, t) = \psi_E(x_1, x_2, \ldots, x_N)\, e^{-i\,E\,t/\hbar}, \tag{6.18}$$

where

$$\psi_E(x_1, x_2, \ldots, x_N) = \psi_{E_1}(x_1)\,\psi_{E_2}(x_2)\ldots\psi_{E_N}(x_N), \tag{6.19}$$

and

$$E = \sum_{i=1,N} E_i. \tag{6.20}$$

Clearly, for the case of non-interacting particles, the energy of the whole system is simply the sum of the energies of the component particles.

6.4 Two-Particle Systems

Consider a system consisting of two particles, mass m_1 and m_2, interacting via the potential $V(x_1 - x_2)$ which only depends on the *relative positions* of the particles. According to Eqs. (6.3) and (6.10), the Hamiltonian of the system is written

$$H(x_1, x_2) = -\frac{\hbar^2}{2\,m_1}\frac{\partial^2}{\partial x_1^2} - \frac{\hbar^2}{2\,m_2}\frac{\partial^2}{\partial x_2^2} + V(x_1 - x_2). \tag{6.21}$$

Let

$$x' = x_1 - x_2 \tag{6.22}$$

be the particles' relative position, and

$$X = \frac{m_1\,x_1 + m_2\,x_2}{m_1 + m_2} \tag{6.23}$$

the position of the center of mass. It is easily demonstrated that

$$\frac{\partial}{\partial x_1} = \frac{m_1}{m_1 + m_2}\frac{\partial}{\partial X} + \frac{\partial}{\partial x'}, \tag{6.24}$$

$$\frac{\partial}{\partial x_2} = \frac{m_2}{m_1 + m_2}\frac{\partial}{\partial X} - \frac{\partial}{\partial x'}. \tag{6.25}$$

Hence, when expressed in terms of the new variables, x' and X, the Hamiltonian becomes

$$H(x', X) = -\frac{\hbar^2}{2\,M}\frac{\partial^2}{\partial X^2} - \frac{\hbar^2}{2\,\mu}\frac{\partial^2}{\partial x'^2} + V(x'), \tag{6.26}$$

where

$$M = m_1 + m_2 \tag{6.27}$$

is the total mass of the system, and

$$\mu = \frac{m_1\,m_2}{m_1 + m_2} \tag{6.28}$$

the so-called *reduced mass*. Note that the total momentum of the system can be written

$$P = -i\,\hbar\left(\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}\right) = -i\,\hbar\,\frac{\partial}{\partial X}. \tag{6.29}$$

The fact that the Hamiltonian (6.26) is separable when expressed in terms of the new coordinates [*i.e.*, $H(x', X) = H_{x'}(x') + H_X(X)$] suggests, by analogy with the analysis in the previous section, that the wavefunction can be factorized: *i.e.*,

$$\psi(x_1, x_2, t) = \psi_{x'}(x', t)\,\psi_X(X, t). \tag{6.30}$$

Hence, the time-dependent Schrödinger equation (6.7) also factorizes to give

$$i\,\hbar\,\frac{\partial\psi_{x'}}{\partial t} = -\frac{\hbar^2}{2\,\mu}\frac{\partial^2\psi_{x'}}{\partial x'^2} + V(x')\,\psi_{x'}, \tag{6.31}$$

and

$$i\,\hbar\,\frac{\partial\psi_X}{\partial t} = -\frac{\hbar^2}{2\,M}\frac{\partial^2\psi_X}{\partial X^2}. \tag{6.32}$$

The above equation can be solved to give

$$\psi_X(X, t) = \psi_0\,e^{\,i\,(P'\,X/\hbar - E'\,t/\hbar)}, \tag{6.33}$$

where ψ_0, P', and $E' = P'^2/2\,M$ are constants. It is clear, from Eqs. (6.29), (6.30), and (6.33), that the total momentum of the system takes the constant value P': *i.e.*, momentum is conserved.

Suppose that we work in the *centre of mass frame* of the system, which is characterized by $P' = 0$. It follows that $\psi_X = \psi_0$. In this case, we can write the wavefunction of the system in the form $\psi(x_1, x_2, t) = \psi_{x'}(x', t)\,\psi_0 \equiv \psi(x_1 - x_2, t)$, where

$$i\,\hbar\,\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2\,\mu}\frac{\partial^2\psi}{\partial x^2} + V(x)\,\psi. \tag{6.34}$$

In other words, in the center of mass frame, two particles of mass m_1 and m_2, moving in the potential $V(x_1 - x_2)$, are equivalent to a single particle of mass μ, moving in the potential $V(x)$, where $x = x_1 - x_2$. This is a familiar result from classical dynamics.

6.5 Identical Particles

Consider a system consisting of two *identical* particles of mass m. As before, the instantaneous state of the system is specified by the complex wavefunction $\psi(x_1, x_2, t)$. However, the only thing which this wavefunction tells us is that the probability of finding the first particle between x_1 and $x_1 + dx_1$, and the second between x_2 and $x_2 + dx_2$, at time t is $|\psi(x_1, x_2, t)|^2\, dx_1\, dx_2$. However, since the particles are identical, this must be the same as the probability of finding the first particle between x_2 and $x_2 + dx_2$, and the second between x_1 and $x_1 + dx_1$, at time t (since, in both cases, the physical outcome of the measurement is exactly the same). Hence, we conclude that

$$|\psi(x_1, x_2, t)|^2 = |\psi(x_2, x_1, t)|^2, \tag{6.35}$$

or

$$\psi(x_1, x_2, t) = e^{i\varphi}\, \psi(x_2, x_1, t), \tag{6.36}$$

where φ is a real constant. However, if we swap the labels on particles 1 and 2 (which are, after all, arbitrary for identical particles), and repeat the argument, we also conclude that

$$\psi(x_2, x_1, t) = e^{i\varphi}\, \psi(x_1, x_2, t). \tag{6.37}$$

Hence,

$$e^{2i\varphi} = 1. \tag{6.38}$$

The only solutions to the above equation are $\varphi = 0$ and $\varphi = \pi$. Thus, we infer that for a system consisting of two identical particles, the wavefunction must be either *symmetric* or *anti-symmetric* under interchange of particle labels: *i.e.*, either

$$\psi(x_2, x_1, t) = \psi(x_1, x_2, t), \tag{6.39}$$

or

$$\psi(x_2, x_1, t) = -\psi(x_1, x_2, t). \tag{6.40}$$

The above argument can easily be extended to systems containing more than two identical particles.

It turns out that whether the wavefunction of a system containing many identical particles is symmetric or anti-symmetric under interchange of the labels on any two particles is determined by the nature of the particles themselves. Particles with wavefunctions which are *symmetric* under label interchange are said to obey *Bose-Einstein* statistics, and are called *bosons*—for instance, photons are bosons. Particles with wavefunctions which are *anti-symmetric* under label interchange are said to obey *Fermi-Dirac* statistics, and are called *fermions*—for instance, electrons, protons, and neutrons are fermions.

Consider a system containing two identical and non-interacting bosons. Let $\psi(x, E)$ be a properly normalized, single-particle, stationary wavefunction corresponding to a state of definite energy E. The stationary wavefunction of the whole system is written

$$\psi_{E\,\rm boson}(x_1, x_2) = \frac{1}{\sqrt{2}}\left[\psi(x_1, E_a)\, \psi(x_2, E_b) + \psi(x_2, E_a)\, \psi(x_1, E_b)\right], \tag{6.41}$$

when the energies of the two particles are E_a and E_b. This expression automatically satisfies the symmetry requirement on the wavefunction. Incidentally, since the particles are identical, we cannot be sure which particle has energy E_a, and which has energy E_b—only that one particle has energy E_a, and the other E_b.

For a system consisting of two identical and non-interacting fermions, the stationary wavefunction of the whole system takes the form

$$\psi_{E\,\rm fermion}(x_1,x_2) = \frac{1}{\sqrt{2}}\left[\psi(x_1,E_a)\,\psi(x_2,E_b) - \psi(x_2,E_a)\,\psi(x_1,E_b)\right], \tag{6.42}$$

Again, this expression automatically satisfies the symmetry requirement on the wavefunction. Note that if $E_a = E_b$ then the total wavefunction becomes zero everywhere. Now, in quantum mechanics, a null wavefunction corresponds to the absence of a state. We thus conclude that it is impossible for the two fermions in our system to occupy the same single-particle stationary state.

Finally, if the two particles are somehow distinguishable then the stationary wavefunction of the system is simply

$$\psi_{E\,\rm dist}(x_1,x_2) = \psi(x_1,E_a)\,\psi(x_2,E_b). \tag{6.43}$$

Let us evaluate the variance of the distance, $x_1 - x_2$, between the two particles, using the above three wavefunctions. It is easily demonstrated that if the particles are distinguishable then

$$\langle (x_1-x_2)^2\rangle_{\rm dist} = \langle x^2\rangle_a + \langle x^2\rangle_b - 2\,\langle x\rangle_a\,\langle x\rangle_b, \tag{6.44}$$

where

$$\langle x^n\rangle_{a,b} = \int_{-\infty}^{\infty}\psi^\ast(x,E_{a,b})\,x^n\,\psi(x,E_{a,b})\,dx. \tag{6.45}$$

For the case of two identical bosons, we find

$$\langle (x_1-x_2)^2\rangle_{\rm boson} = \langle (x_1-x_2)^2\rangle_{\rm dist} - 2\,|\langle x\rangle_{ab}|^2, \tag{6.46}$$

where

$$\langle x\rangle_{ab} = \int_{-\infty}^{\infty}\psi^\ast(x,E_a)\,x\,\psi(x,E_b)\,dx. \tag{6.47}$$

Here, we have assumed that $E_a \neq E_b$, so that

$$\int_{-\infty}^{\infty}\psi^\ast(x,E_a)\,\psi(x,E_b)\,dx = 0. \tag{6.48}$$

Finally, for the case of two identical fermions, we obtain

$$\langle (x_1-x_2)^2\rangle_{\rm fermion} = \langle (x_1-x_2)^2\rangle_{\rm dist} + 2\,|\langle x\rangle_{ab}|^2, \tag{6.49}$$

Equation (6.46) shows that the symmetry requirement on the total wavefunction of two identical *bosons* forces the particles to be, on average, *closer together* than two similar distinguishable particles. Conversely, Eq. (6.49) shows that the symmetry requirement on the total wavefunction of two identical *fermions* forces the particles to be, on average, *further apart* than two similar distinguishable particles. However, the strength of this effect depends on square of the magnitude of $\langle x\rangle_{ab}$, which measures the *overlap* between the wavefunctions $\psi(x, E_a)$ and $\psi(x, E_b)$. It is evident, then, that if these two wavefunctions do not overlap to any great extent then identical bosons or fermions will act very much like distinguishable particles.

For a system containing N identical and non-interacting fermions, the anti-symmetric stationary wavefunction of the system is written

$$\psi_E(x_1, x_2, \ldots x_N) = \frac{1}{\sqrt{N!}}\begin{vmatrix} \psi(x_1, E_1) & \psi(x_2, E_1) & \ldots & \psi(x_N, E_1) \\ \psi(x_1, E_2) & \psi(x_2, E_2) & \ldots & \psi(x_N, E_2) \\ \vdots & \vdots & \vdots & \vdots \\ \psi(x_1, E_N) & \psi(x_2, E_N) & \ldots & \psi(x_N, E_N) \end{vmatrix}. \tag{6.50}$$

This expression is known as the *Slater determinant*, and automatically satisfies the symmetry requirements on the wavefunction. Here, the energies of the particles are $E_1, E_2, \ldots, E_N$. Note, again, that if any two particles in the system have the same energy (*i.e.*, if $E_i = E_j$ for some $i \neq j$) then the total wavefunction is null. We conclude that it is impossible for any two identical fermions in a multi-particle system to occupy the same single-particle stationary state. This important result is known as the *Pauli exclusion principle*.

Exercises (N.B. Neglect spin in the following questions.)

1. Consider a system consisting of two non-interacting particles, and three one-particle states, $\psi_a(x)$, $\psi_b(x)$, and $\psi_c(x)$. How many different two-particle states can be constructed if the particles are (a) distinguishable, (b) indistinguishable bosons, or (c) indistinguishable fermions?

2. Consider two non-interacting particles, each of mass m, in a one-dimensional harmonic oscillator potential of classical oscillation frequency ω. If one particle is in the ground-state, and the other in the first excited state, calculate $\langle (x_1 - x_2)^2\rangle$ assuming that the particles are (a) distinguishable, (b) indistinguishable bosons, or (c) indistinguishable fermions.

3. Two non-interacting particles, with the same mass m, are in a one-dimensional box of length a. What are the four lowest energies of the system? What are the degeneracies of these energies if the two particles are (a) distinguishable, (b) indistinguishable bosons, or (c) indistingishable fermions?

4. Two particles in a one-dimensional box of length a occupy the $n = 4$ and $n' = 3$ states. Write the properly normalized wavefunctions if the particles are (a) distinguishable, (b) indistinguishable bosons, or (c) indistinguishable fermions.

7 Three-Dimensional Quantum Mechanics

7.1 Introduction

In this chapter, we shall extend our previous one-dimensional formulation of non-relativistic quantum mechanics to produce a fully three-dimensional theory.

7.2 Fundamental Concepts

We have seen that in one dimension the instantaneous state of a single non-relativistic particle is fully specified by a complex wavefunction, $\psi(x,t)$. The probability of finding the particle at time t between x and $x+dx$ is $P(x,t)\,dx$, where

$$P(x,t) = |\psi(x,t)|^2. \tag{7.1}$$

Moreover, the wavefunction is normalized such that

$$\int_{-\infty}^{\infty} |\psi(x,t)|^2\,dx = 1 \tag{7.2}$$

at all times.

In three dimensions, the instantaneous state of a single particle is also fully specified by a complex wavefunction, $\psi(x,y,z,t)$. By analogy with the one-dimensional case, the probability of finding the particle at time t between x and $x+dx$, between y and $y+dx$, and between z and $z+dz$, is $P(x,y,z,t)\,dx\,dy\,dz$, where

$$P(x,y,z,t) = |\psi(x,y,z,t)|^2. \tag{7.3}$$

As usual, this interpretation of the wavefunction only makes sense if the wavefunction is normalized such that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\psi(x,y,z,t)|^2\,dx\,dy\,dz = 1. \tag{7.4}$$

This normalization constraint ensures that the probability of finding the particle anywhere is space is always unity.

In one dimension, we can write the probability conservation equation (see Sect. 4.5)

$$\frac{\partial|\psi|^2}{\partial t} + \frac{\partial j}{\partial x} = 0, \tag{7.5}$$

where

$$j = \frac{i\,\hbar}{2\,m}\left(\psi\,\frac{\partial\psi^*}{\partial x} - \psi^*\,\frac{\partial\psi}{\partial x}\right) \tag{7.6}$$

is the flux of probability along the x-axis. Integrating Eq. (7.5) over all space, and making use of the fact that $\psi \rightarrow 0$ as $|x| \rightarrow \infty$ if ψ is to be square-integrable, we obtain

$$\frac{d}{dt}\int_{-\infty}^{\infty} |\psi(x,t)|^2\,dx = 0. \tag{7.7}$$

In other words, if the wavefunction is initially normalized then it stays normalized as time progresses. This is a necessary criterion for the viability of our basic interpretation of $|\psi|^2$ as a probability density.

In three dimensions, by analogy with the one dimensional case, the probability conservation equation becomes

$$\frac{\partial|\psi|^2}{\partial t} + \frac{\partial j_x}{\partial x} + \frac{\partial j_y}{\partial y} + \frac{\partial j_z}{\partial z} = 0. \tag{7.8}$$

Here,

$$j_x = \frac{i\,\hbar}{2\,m}\left(\psi\,\frac{\partial\psi^*}{\partial x} - \psi^*\,\frac{\partial\psi}{\partial x}\right) \tag{7.9}$$

is the flux of probability along the x-axis, and

$$j_y = \frac{i\,\hbar}{2\,m}\left(\psi\,\frac{\partial\psi^*}{\partial y} - \psi^*\,\frac{\partial\psi}{\partial y}\right) \tag{7.10}$$

the flux of probability along the y-axis, *etc*. Integrating Eq. (7.8) over all space, and making use of the fact that $\psi \rightarrow 0$ as $|\mathbf{r}| \rightarrow \infty$ if ψ is to be square-integrable, we obtain

$$\frac{d}{dt}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\psi(x,y,z,t)|^2\,dx\,dy\,dz = 0. \tag{7.11}$$

Thus, the normalization of the wavefunction is again preserved as time progresses, as must be the case if $|\psi|^2$ is to be interpreted as a probability density.

In one dimension, position is represented by the algebraic operator x, whereas momentum is represented by the differential operator $-i\,\hbar\,\partial/\partial x$ (see Sect. 4.6). By analogy, in three dimensions, the Cartesian coordinates x, y, and z are represented by the algebraic operators x, y, and z, respectively, whereas the three Cartesian components of momentum, p_x, p_y, and p_z, have the following representations:

$$p_x \equiv -i\,\hbar\,\frac{\partial}{\partial x}, \tag{7.12}$$

$$p_y \equiv -i\,\hbar\,\frac{\partial}{\partial y}, \tag{7.13}$$

$$p_z \equiv -i\,\hbar\,\frac{\partial}{\partial z}. \tag{7.14}$$

Let $x_1 = x$, $x_2 = y$, $x_3 = z$, and $p_1 = p_x$, *etc.* Since the x_i are *independent* variables (*i.e.*, $\partial x_i/\partial x_j = \delta_{ij}$), we conclude that the various position and momentum operators satisfy the following commutation relations:

$$[x_i, x_j] = 0, \tag{7.15}$$

$$[p_i, p_j] = 0, \tag{7.16}$$

$$[x_i, p_j] = i\,\hbar\,\delta_{ij}. \tag{7.17}$$

Now, we know, from Sect. 4.10, that two dynamical variables can only be (exactly) measured *simultaneously* if the operators which represent them in quantum mechanics *commute* with one another. Thus, it is clear, from the above commutation relations, that the only restriction on measurement in a system consisting of a single particle moving in three dimensions is that it is impossible to simultaneously measure a given position coordinate and the corresponding component of momentum. Note, however, that it is perfectly possible to simultaneously measure two different positions coordinates, or two different components of the momentum. The commutation relations (7.15)–(7.17) again illustrate the point that quantum mechanical operators corresponding to different degrees of freedom of a dynamical system (in this case, motion in different directions) tend to commute with one another (see Sect. 6.2).

In one dimension, the time evolution of the wavefunction is given by [see Eq. (4.63)]

$$i\,\hbar\,\frac{\partial\psi}{\partial t} = H\,\psi, \tag{7.18}$$

where H is the Hamiltonian. The same equation governs the time evolution of the wavefunction in three dimensions.

Now, in one dimension, the Hamiltonian of a non-relativistic particle of mass m takes the form

$$H = \frac{p_x^{\,2}}{2\,m} + V(x,t), \tag{7.19}$$

where $V(x)$ is the potential energy. In three dimensions, this expression generalizes to

$$H = \frac{p_x^{\,2} + p_y^{\,2} + p_z^{\,2}}{2\,m} + V(x,y,z,t). \tag{7.20}$$

Hence, making use of Eqs. (7.12)–(7.14) and (7.18), the three-dimensional version of the time-dependent Schröndiger equation becomes [see Eq. (4.1)]

$$i\,\hbar\,\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2\,m}\,\nabla^2\psi + V\,\psi. \tag{7.21}$$

Here, the differential operator

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{7.22}$$

is known as the *Laplacian*. Incidentally, the probability conservation equation (7.8) is easily derivable from Eq. (7.21). An eigenstate of the Hamiltonian corresponding to the eigenvalue E satisfies

$$H\,\psi = E\,\psi. \tag{7.23}$$

It follows from Eq. (7.18) that (see Sect. 4.12)

$$\psi(x,y,z,t) = \psi(x,y,z)\, e^{-i\,E\,t/\hbar}, \tag{7.24}$$

where the stationary wavefunction $\psi(x,y,z)$ satisfies the three-dimensional version of the time-independent Schröndiger equation [see Eq. (4.159)]:

$$\nabla^2\psi = \frac{2\,m}{\hbar^2}\,(V-E)\,\psi, \tag{7.25}$$

where V is assumed not to depend explicitly on t.

7.3 Particle in a Box

Consider a particle of mass m trapped inside a cubic box of dimension a (see Sect. 5.2). The particle's stationary wavefunction, $\psi(x,y,z)$, satisfies

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)\psi = -\frac{2\,m}{\hbar^2}\,E\,\psi, \tag{7.26}$$

where E is the particle energy. The wavefunction satisfies the boundary condition that it must be zero at the edges of the box.

Let us search for a separable solution to the above equation of the form

$$\psi(x,y,z) = X(x)\,Y(y)\,Z(z). \tag{7.27}$$

The factors of the wavefunction satisfy the boundary conditions $X(0)=X(a)=0$, $Y(0)=Y(a)=0$, and $Z(0)=Z(a)=0$. Substituting (7.27) into Eq. (7.26), and rearranging, we obtain

$$\frac{X''}{X}+\frac{Y''}{Y}+\frac{Z''}{Z} = -\frac{2\,m}{\hbar^2}\,E, \tag{7.28}$$

where $'$ denotes a derivative with respect to argument. It is evident that the only way in which the above equation can be satisfied at all points within the box is if

$$\frac{X''}{X} = -k_x^2, \tag{7.29}$$

$$\frac{Y''}{Y} = -k_y^2, \tag{7.30}$$

$$\frac{Z''}{Z} = -k_z^2, \tag{7.31}$$

where k_x^2, k_y^2, and k_z^2 are *spatial constants*. Note that the right-hand sides of the above equations must contain negative, rather than positive, spatial constants, because it would not otherwise be possible to satisfy the boundary conditions. The solutions to the above equations which are properly normalized, and satisfy the boundary conditions, are [see Eq. (5.11)]

$$X(x) = \sqrt{\frac{2}{a}}\sin(k_x\,x), \tag{7.32}$$

$$Y(y) = \sqrt{\frac{2}{a}}\sin(k_y\,y), \tag{7.33}$$

$$Z(z) = \sqrt{\frac{2}{a}}\sin(k_z\,z), \tag{7.34}$$

where

$$k_x = \frac{l_x\,\pi}{a}, \tag{7.35}$$

$$k_y = \frac{l_y\,\pi}{a}, \tag{7.36}$$

$$k_z = \frac{l_z\,\pi}{a}. \tag{7.37}$$

Here, l_x, l_y, and l_z are *positive integers*. Thus, from Eqs. (7.28)–(7.31), the energy is written [see Eq. (5.9)]

$$E = \frac{l^2\,\pi^2\,\hbar^2}{2\,m\,a^2}. \tag{7.38}$$

where

$$l^2 = l_x^{\,2} + l_y^{\,2} + l_z^{\,2}. \tag{7.39}$$

7.4 Degenerate Electron Gases

Consider N electrons trapped in a cubic box of dimension a. Let us treat the electrons as essentially *non-interacting* particles. According to Sect. 6.3, the total energy of a system consisting of many non-interacting particles is simply the sum of the single-particle energies of the individual particles. Furthermore, electrons are subject to the *Pauli exclusion principle* (see Sect. 6.5), since they are indistinguishable fermions. The exclusion principle states that no two electrons in our system can occupy the same single-particle energy level. Now, from the previous section, the single-particle energy levels for a particle in a box are characterized by the three quantum numbers l_x, l_y, and l_z. Thus, we conclude that no two electrons in our system can have the same set of values of l_x, l_y, and l_z. It turns out that this is not quite true, because electrons possess an intrinsic angular momentum called *spin* (see Cha. 10). The spin states of an electron are governed by an additional quantum number, which can take one of two different values. Hence, when spin is taken into account, we conclude that a maximum of *two* electrons (with different spin quantum numbers) can occupy a single-particle energy level corresponding to a particular set of values of l_x, l_y, and l_z. Note, from Eqs. (7.38) and (7.39), that the associated particle energy is proportional to $l^2 = l_x^2 + l_y^2 + l_z^2$.

Suppose that our electrons are *cold*: *i.e.*, they have comparatively little thermal energy. In this case, we would expect them to fill the lowest single-particle energy levels available to them. We can imagine the single-particle energy levels as existing in a sort of three-dimensional quantum number space whose Cartesian coordinates are l_x, l_y, and l_z. Thus, the energy levels are uniformly distributed in this space on a cubic lattice. Moreover, the distance between nearest neighbour energy levels is unity. This implies that the number of energy levels per unit volume is also unity. Finally, the energy of a given energy level is proportional to its distance, $l^2 = l_x^2 + l_y^2 + l_z^2$, from the origin.

Since we expect cold electrons to occupy the lowest energy levels available to them, but only two electrons can occupy a given energy level, it follows that if the number of electrons, N, is very large then the filled energy levels will be approximately distributed in a *sphere* centered on the origin of quantum number space. The number of energy levels contained in a sphere of radius l is approximately equal to the volume of the sphere—since the number of energy levels per unit volume is unity. It turns out that this is not quite correct, because we have forgotten that the quantum numbers l_x, l_y, and l_z can only take *positive* values. Hence, the filled energy levels actually only occupy one *octant* of a sphere. The radius l_F of the octant of filled energy levels in quantum number space can be calculated by equating the number of energy levels it contains to the number of electrons, N. Thus, we can write

$$N = 2 \times \frac{1}{8} \times \frac{4\pi}{3}\, l_F^3. \tag{7.40}$$

Here, the factor 2 is to take into account the two spin states of an electron, and the factor 1/8 is to take account of the fact that l_x, l_y, and l_z can only take positive values. Thus,

$$l_F = \left(\frac{3\,N}{\pi}\right)^{1/3}. \tag{7.41}$$

According to Eq. (7.38), the energy of the most energetic electrons—which is known as the *Fermi energy*—is given by

$$E_F = \frac{l_F^2\,\pi^2\,\hbar^2}{2\,m_e\,a^2} = \frac{\pi^2\,\hbar^2}{2\,m\,a^2}\left(\frac{3\,N}{\pi}\right)^{2/3}, \tag{7.42}$$

where m_e is the electron mass. This can also be written as

$$E_F = \frac{\pi^2\,\hbar^2}{2\,m_e}\left(\frac{3\,n}{\pi}\right)^{2/3}, \tag{7.43}$$

where $n = N/a^3$ is the number of electrons per unit volume (in real space). Note that the Fermi energy only depends on the *number density* of the confined electrons.

The mean energy of the electrons is given by

$$\bar{E} = E_F \int_0^{l_F} l^2\, 4\pi\, l^2\, dl \bigg/ \frac{4}{3}\,\pi\, l_F^5 = \frac{3}{5}\, E_F, \tag{7.44}$$

since $E \propto l^2$, and the energy levels are uniformly distributed in quantum number space inside an octant of radius l_F. Now, according to classical physics, the mean thermal energy of the electrons is $(3/2)\, k_B\, T$, where T is the electron temperature, and k_B the Boltzmann constant. Thus, if $k_B\, T \ll E_F$ then our original assumption that the electrons are cold is valid. Note that, in this case, the electron energy is *much larger* than that predicted by classical physics—electrons in this state are termed *degenerate*. On the other hand, if $k_B\, T \gg E_F$ then the electrons are hot, and are essentially governed by classical physics—electrons in this state are termed *non-degenerate*.

The total energy of a degenerate electron gas is

$$E_{total} = N\, \bar{E} = \frac{3}{5}\, N\, E_F. \tag{7.45}$$

Hence, the gas pressure takes the form

$$P = -\frac{\partial E_{total}}{\partial V} = \frac{2}{5}\, n\, E_F, \tag{7.46}$$

since $E_F \propto a^{-2} = V^{-2/3}$ [see Eq. (7.42)]. Now, the pressure predicted by classical physics is $P = n\, k_B\, T$. Thus, a degenerate electron gas has a *much higher* pressure than that which would be predicted by classical physics. This is an entirely quantum mechanical effect, and is due to the fact that identical fermions cannot get significantly closer together than a de Broglie wavelength without violating the Pauli exclusion principle. Note that, according to Eq. (7.43), the mean spacing between degenerate electrons is

$$d \sim n^{-1/3} \sim \frac{\hbar}{\sqrt{m_e\, E}} \sim \frac{\hbar}{p} \sim \lambda, \tag{7.47}$$

where λ is the de Broglie wavelength. Thus, an electron gas is non-degenerate when the mean spacing between the electrons is much greater than the de Broglie wavelength, and becomes degenerate as the mean spacing approaches the de Broglie wavelength.

In turns out that the conduction (*i.e.*, free) electrons inside metals are highly degenerate (since the number of electrons per unit volume is very large, and $E_F \propto n^{2/3}$). Indeed, most metals are hard to compress as a direct consequence of the high degeneracy pressure of their conduction electrons. To be more exact, resistance to compression is usually measured in terms of a quantity known as the *bulk modulus*, which is defined

$$B = -V\, \frac{\partial P}{\partial V} \tag{7.48}$$

Now, for a fixed number of electrons, $P \propto V^{-5/3}$ [see Eqs. (7.42) and (7.46)]. Hence,

$$B = \frac{5}{3}\, P = \frac{\pi^3\, \hbar^2}{9\, m} \left(\frac{3\, n}{\pi}\right)^{5/3}. \tag{7.49}$$

For example, the number density of free electrons in magnesium is $n \sim 8.6\times 10^{28}\,\mathrm{m}^{-3}$. This leads to the following estimate for the bulk modulus: $B \sim 6.4\times 10^{10}\,\mathrm{N\,m^{-2}}$. The actual bulk modulus is $B = 4.5\times 10^{10}\,\mathrm{N\,m^{-2}}$.

7.5 White-Dwarf Stars

A main-sequence hydrogen-burning star, such as the Sun, is maintained in equilibrium via the balance of the gravitational attraction tending to make it collapse, and the thermal pressure tending to make it expand. Of course, the thermal energy of the star is generated by nuclear reactions occurring deep inside its core. Eventually, however, the star will run out of burnable fuel, and, therefore, start to collapse, as it radiates away its remaining thermal energy. What is the ultimate fate of such a star?

A burnt-out star is basically a gas of electrons and ions. As the star collapses, its density increases, and so the mean separation between its constituent particles decreases. Eventually, the mean separation becomes of order the de Broglie wavelength of the electrons, and the electron gas becomes *degenerate*. Note, that the de Broglie wavelength of the ions is much smaller than that of the electrons, so the ion gas remains non-degenerate. Now, even at zero temperature, a degenerate electron gas exerts a substantial pressure, because the Pauli exclusion principle prevents the mean electron separation from becoming significantly smaller than the typical de Broglie wavelength (see previous section). Thus, it is possible for a burnt-out star to maintain itself against complete collapse under gravity via the *degeneracy pressure* of its constituent electrons. Such stars are termed *white-dwarfs*. Let us investigate the physics of white-dwarfs in more detail.

The total energy of a white-dwarf star can be written

$$\mathcal{E} = K + U, \tag{7.50}$$

where K is the kinetic energy of the degenerate electrons (the kinetic energy of the ion is negligible), and U is the gravitational potential energy. Let us assume, for the sake of simplicity, that the density of the star is *uniform*. In this case, the gravitational potential energy takes the form

$$U = -\frac{3}{5}\frac{G\,M^2}{R}, \tag{7.51}$$

where G is the gravitational constant, M is the stellar mass, and R is the stellar radius.

From the previous subsection, the kinetic energy of a degenerate electron gas is simply

$$K = N\,\bar{E} = \frac{3}{5}\,N\,E_F = \frac{3}{5}\,N\,\frac{\pi^2\,\hbar^2}{2\,m_e}\left(\frac{3\,N}{\pi\,V}\right)^{2/3}, \tag{7.52}$$

where N is the number of electrons, V the volume of the star, and m_e the electron mass.

The interior of a white-dwarf star is composed of atoms like C^{12} and O^{16} which contain equal numbers of protons, neutrons, and electrons. Thus,

$$M = 2\,N\,m_p, \tag{7.53}$$

where m_p is the proton mass.

Equations (7.50)–(7.53) can be combined to give

$$\mathcal{E} = \frac{A}{R^2} - \frac{B}{R}, \tag{7.54}$$

where

$$A = \frac{3}{20}\left(\frac{9\pi}{8}\right)^{2/3} \frac{\hbar^2}{m_e}\left(\frac{M}{m_p}\right)^{5/3}, \tag{7.55}$$

$$B = \frac{3}{5}\,G\,M^2. \tag{7.56}$$

The equilibrium radius of the star, R_*, is that which *minimizes* the total energy $\mathcal{E}$. In fact, it is easily demonstrated that

$$R_* = \frac{2\,A}{B}, \tag{7.57}$$

which yields

$$R_* = \frac{(9\pi)^{2/3}}{8}\,\frac{\hbar^2}{G\,m_e\,m_p^{5/3}\,M^{1/3}}. \tag{7.58}$$

The above formula can also be written

$$\frac{R_*}{R_\odot} = 0.010\left(\frac{M_\odot}{M}\right)^{1/3}, \tag{7.59}$$

where $R_\odot = 7 \times 10^5$ km is the solar radius, and $M_\odot = 2 \times 10^{30}$ kg the solar mass. It follows that the radius of a typical solar mass white-dwarf is about 7000 km: *i.e.*, about the same as the radius of the Earth. The first white-dwarf to be discovered (in 1862) was the companion of Sirius. Nowadays, thousands of white-dwarfs have been observed, all with properties similar to those described above.

Note from Eqs. (7.52), (7.53), and (7.59) that $\bar{E} \propto M^{4/3}$. In other words, the mean energy of the electrons inside a white dwarf *increases* as the stellar mass increases. Hence, for a sufficiently massive white dwarf, the electrons can become *relativistic*. It turns out that the degeneracy pressure for relativistic electrons only scales as R^{-1}, rather that R^{-2}, and thus is unable to balance the gravitational pressure [which also scales as R^{-1}—see Eq. (7.54)]. It follows that electron degeneracy pressure is only able to halt the collapse of a burnt-out star provided that the stellar mass does not exceed some critical value, known as the *Chandrasekhar limit*, which turns out to be about 1.4 times the mass of the Sun. Stars whose mass exceeds the Chandrasekhar limit inevitably collapse to produce extremely compact objects, such as neutron stars (which are held up by the degeneracy pressure of their constituent neutrons), or black holes.

Exercises

1. Consider a particle of mass m moving in a three-dimensional isotropic harmonic oscillator potential of force constant k. Solve the problem via the separation of variables, and obtain an expression for the allowed values of the total energy of the system (in a stationary state).

2. Repeat the calculation of the Fermi energy of a gas of fermions by assuming that the fermions are massless, so that the energy-momentum relation is $E = p\,c$.

3. Calculate the density of states of an electron gas in a cubic box of volume L^3, bearing in mind that there are two electrons per energy state. In other words, calculate the number of electron states in the interval E to $E + dE$. This number can be written $dN = \rho(E)\,dE$, where ρ is the density of states.

4. Repeat the above calculation for a two-dimensional electron gas in a square box of area L^2.

5. Given that the number density of free electrons in copper is $8.5 \times 10^{28}\,\mathrm{m}^{-3}$, calculate the Fermi energy in electron volts, and the velocity of an electron whose kinetic energy is equal to the Fermi energy.

6. Obtain an expression for the Fermi energy (in eV) of an electron in a white dwarf star as a function of the stellar mass (in solar masses). At what mass does the Fermi energy equal the rest mass energy?

8 Orbital Angular Momentum

8.1 Introduction

As is well-known, angular momentum plays a vitally important role in the classical description of three-dimensional motion. Let us now investigate the role of angular momentum in the quantum mechanical description of such motion.

8.2 Angular Momentum Operators

In classical mechanics, the vector angular momentum, **L**, of a particle of position vector **r** and linear momentum **p** is defined as

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}. \tag{8.1}$$

It follows that

$$L_x = y\,p_z - z\,p_y, \tag{8.2}$$

$$L_y = z\,p_x - x\,p_z, \tag{8.3}$$

$$L_z = x\,p_y - y\,p_x. \tag{8.4}$$

Let us, first of all, consider whether it is possible to use the above expressions as the definitions of the operators corresponding to the components of angular momentum in quantum mechanics, assuming that the x_i and p_i (where $x_1 \equiv x$, $p_1 \equiv p_x$, $x_2 \equiv y$, *etc.*) correspond to the appropriate quantum mechanical position and momentum operators. The first point to note is that expressions (8.2)–(8.4) are *unambiguous* with respect to the order of the terms in multiplicative factors, since the various position and momentum operators appearing in them all *commute* with one another [see Eqs. (7.17)]. Moreover, given that the x_i and the p_i are Hermitian operators, it is easily seen that the L_i are also *Hermitian*. This is important, since only Hermitian operators can represent physical variables in quantum mechanics (see Sect. 4.6). We, thus, conclude that Eqs. (8.2)–(8.4) are plausible definitions for the quantum mechanical operators which represent the components of angular momentum.

Let us now derive the commutation relations for the L_i. For instance,

$$\begin{aligned}[L_x, L_y] &= [y\,p_z - z\,p_y, z\,p_x - x\,p_z] = y\,p_x\,[p_z, z] + x\,p_y\,[z, p_z] \\ &= i\,\hbar\,(x\,p_y - y\,p_x) = i\,\hbar\,L_z,\end{aligned} \tag{8.5}$$

where use has been made of the definitions of the L_i [see Eqs. (8.2)–(8.4)], and commutation relations (7.15)–(7.17) for the x_i and p_i. There are two similar commutation relations: one for L_y and L_z, and one for L_z and L_x. Collecting all of these commutation relations together, we obtain

$$\begin{aligned}
[L_x, L_y] &= i\,\hbar\,L_z, &(8.6)\\
[L_y, L_z] &= i\,\hbar\,L_x, &(8.7)\\
[L_z, L_x] &= i\,\hbar\,L_y. &(8.8)
\end{aligned}$$

By analogy with classical mechanics, the operator L^2, which represents the *magnitude squared* of the angular momentum vector, is defined

$$L^2 = L_x^2 + L_y^2 + L_z^2. \tag{8.9}$$

Now, it is easily demonstrated that if A and B are two general operators then

$$[A^2, B] = A\,[A, B] + [A, B]\,A. \tag{8.10}$$

Hence,

$$\begin{aligned}
[L^2, L_x] &= [L_y^2, L_x] + [L_z^2, L_x]\\
&= L_y\,[L_y, L_x] + [L_y, L_x]\,L_y + L_z\,[L_z, L_x] + [L_z, L_x]\,L_z\\
&= i\,\hbar\,(-L_y\,L_z - L_z\,L_y + L_z\,L_y + L_y\,L_z) = 0, &(8.11)
\end{aligned}$$

where use has been made of Eqs. (8.6)–(8.8). In other words, L^2 commutes with L_x. Likewise, it is easily demonstrated that L^2 also commutes with L_y, and with L_z. Thus,

$$[L^2, L_x] = [L^2, L_y] = [L^2, L_z] = 0. \tag{8.12}$$

Recall, from Sect. 4.10, that in order for two physical quantities to be (exactly) measured *simultaneously*, the operators which represent them in quantum mechanics must *commute* with one another. Hence, the commutation relations (8.6)–(8.8) and (8.12) imply that we can only simultaneously measure the magnitude squared of the angular momentum vector, L^2, together with, at most, *one* of its Cartesian components. By convention, we shall always choose to measure the z-component, L_z.

Finally, it is helpful to define the operators

$$L_\pm = L_x \pm i\,L_y. \tag{8.13}$$

Note that L_+ and L_- are not Hermitian operators, but are the Hermitian conjugates of one another (see Sect. 4.6): *i.e.*,

$$(L_\pm)^\dagger = L_\mp, \tag{8.14}$$

Moreover, it is easily seen that

$$\begin{aligned}
L_+\,L_- &= (L_x + i\,L_y)\,(L_x - i\,L_y) = L_x^2 + L_y^2 - i\,[L_x, L_y] = L_x^2 + L_y^2 + \hbar\,L_z\\
&= L^2 - L_z^2 + \hbar\,L_z. &(8.15)
\end{aligned}$$

Likewise,

$$L_- \, L_+ = L^2 - L_z^2 - \hbar \, L_z, \tag{8.16}$$

giving

$$[L_+, L_-] = 2 \, \hbar \, L_z. \tag{8.17}$$

We also have

$$[L_+, L_z] \;=\; [L_x, L_z] + i \, [L_y, L_z] = -i \, \hbar \, L_y - \hbar \, L_x = -\hbar \, L_+, \tag{8.18}$$

and, similarly,

$$[L_-, L_z] = \hbar \, L_-. \tag{8.19}$$

8.3 Representation of Angular Momentum

Now, we saw earlier, in Sect. 7.2, that the operators, p_i, which represent the Cartesian components of linear momentum in quantum mechanics, can be represented as the spatial differential operators $-i \, \hbar \, \partial/\partial x_i$. Let us now investigate whether angular momentum operators can similarly be represented as spatial differential operators.

It is most convenient to perform our investigation using conventional *spherical polar coordinates*: *i.e.*, r, θ, and ϕ. These are defined with respect to our usual Cartesian coordinates as follows:

$$x \;=\; r \, \sin\theta \, \cos\phi, \tag{8.20}$$

$$y \;=\; r \, \sin\theta \, \sin\phi, \tag{8.21}$$

$$z \;=\; r \, \cos\theta. \tag{8.22}$$

It follows, after some tedious analysis, that

$$\frac{\partial}{\partial x} \;=\; \sin\theta \, \cos\phi \, \frac{\partial}{\partial r} + \frac{\cos\theta \, \cos\phi}{r} \, \frac{\partial}{\partial \theta} - \frac{\sin\phi}{r \, \sin\theta} \, \frac{\partial}{\partial \phi}, \tag{8.23}$$

$$\frac{\partial}{\partial y} \;=\; \sin\theta \, \sin\phi \, \frac{\partial}{\partial r} + \frac{\cos\theta \, \sin\phi}{r} \, \frac{\partial}{\partial \theta} + \frac{\cos\phi}{r \, \sin\theta} \, \frac{\partial}{\partial \phi}, \tag{8.24}$$

$$\frac{\partial}{\partial z} \;=\; \cos\theta \, \frac{\partial}{\partial r} - \frac{\sin\theta}{r} \, \frac{\partial}{\partial \theta}. \tag{8.25}$$

Making use of the definitions (8.2)–(8.4), (8.9), and (8.13), the fundamental representation (7.12)–(7.14) of the p_i operators as spatial differential operators, the Eqs. (8.20)–(8.25), and a great deal of tedious algebra, we finally obtain

$$L_x \;=\; -i \, \hbar \left(-\sin\phi \, \frac{\partial}{\partial \theta} - \cos\phi \, \cot\theta \, \frac{\partial}{\partial \phi} \right), \tag{8.26}$$

$$L_y \;=\; -i \, \hbar \left(\cos\phi \, \frac{\partial}{\partial \theta} - \sin\phi \, \cot\theta \, \frac{\partial}{\partial \phi} \right), \tag{8.27}$$

$$L_z \;=\; -i \, \hbar \, \frac{\partial}{\partial \phi}, \tag{8.28}$$

as well as

$$L^2 = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right], \tag{8.29}$$

and

$$L_\pm = \hbar\,e^{\pm i\,\phi}\left(\pm\frac{\partial}{\partial\theta} + i\,\cot\theta\,\frac{\partial}{\partial\phi}\right). \tag{8.30}$$

We, thus, conclude that all of our angular momentum operators can be represented as differential operators involving the *angular* spherical coordinates, θ and ϕ, but not involving the *radial* coordinate, r.

8.4 Eigenstates of Angular Momentum

Let us find the simultaneous eigenstates of the angular momentum operators L_z and L^2. Since both of these operators can be represented as purely angular differential operators, it stands to reason that their eigenstates only depend on the angular coordinates θ and ϕ. Thus, we can write

$$L_z\,Y_{l,m}(\theta,\phi) = m\,\hbar\,Y_{l,m}(\theta,\phi), \tag{8.31}$$

$$L^2\,Y_{l,m}(\theta,\phi) = l\,(l+1)\,\hbar^2\,Y_{l,m}(\theta,\phi). \tag{8.32}$$

Here, the $Y_{l,m}(\theta,\phi)$ are the eigenstates in question, whereas the dimensionless quantities m and l parameterize the eigenvalues of L_z and L^2, which are $m\,\hbar$ and $l\,(l+1)\,\hbar^2$, respectively. Of course, we expect the $Y_{l,m}$ to be both mutually orthogonal and properly normalized (see Sect. 4.9), so that

$$\oint Y^*_{l',m'}(\theta,\phi)\,Y_{l,m}(\theta,\phi)\,d\Omega = \delta_{ll'}\,\delta_{mm'}, \tag{8.33}$$

where $d\Omega = \sin\theta\,d\theta\,d\phi$ is an element of solid angle, and the integral is over all solid angle.

Now,

$$\begin{aligned} L_z\,(L_+\,Y_{l,m}) &= (L_+\,L_z + [L_z, L_+])\,Y_{l,m} = (L_+\,L_z + \hbar\,L_+)\,Y_{l,m} \\ &= (m+1)\,\hbar\,(L_+\,Y_{l,m}), \end{aligned} \tag{8.34}$$

where use has been made of Eq. (8.18). We, thus, conclude that when the operator L_+ operates on an eigenstate of L_z corresponding to the eigenvalue $m\,\hbar$ it converts it to an eigenstate corresponding to the eigenvalue $(m+1)\,\hbar$. Hence, L_+ is known as the *raising operator* (for L_z). It is also easily demonstrated that

$$L_z\,(L_-\,Y_{l,m}) = (m-1)\,\hbar\,(L_-\,Y_{l,m}). \tag{8.35}$$

In other words, when L_- operates on an eigenstate of L_z corresponding to the eigenvalue $m\,\hbar$ it converts it to an eigenstate corresponding to the eigenvalue $(m-1)\,\hbar$. Hence, L_- is known as the *lowering operator* (for L_z).

Writing

$$\begin{aligned} L_+\,Y_{l,m} &= c^+_{l,m}\,Y_{l,m+1}, && (8.36)\\ L_-\,Y_{l,m} &= c^-_{l,m}\,Y_{l,m-1}, && (8.37)\end{aligned}$$

we obtain

$$L_-\,L_+\,Y_{l,m} = c^+_{l,m}\,c^-_{l,m+1}\,Y_{l,m} = [l\,(l+1) - m\,(m+1)]\,\hbar^2\,Y_{l,m}, \tag{8.38}$$

where use has been made of Eq. (8.16). Likewise,

$$L_+\,L_-\,Y_{l,m} = c^+_{l,m-1}\,c^-_{l,m}\,Y_{l,m} = [l\,(l+1) - m\,(m-1)]\,\hbar^2\,Y_{l,m}, \tag{8.39}$$

where use has been made of Eq. (8.15). It follows that

$$\begin{aligned} c^+_{l,m}\,c^-_{l,m+1} &= [l\,(l+1) - m\,(m+1)]\,\hbar^2, && (8.40)\\ c^+_{l,m-1}\,c^-_{l,m} &= [l\,(l+1) - m\,(m-1)]\,\hbar^2. && (8.41)\end{aligned}$$

These equations are satisfied when

$$c^\pm_{l,m} = [l\,(l+l) - m\,(m\pm 1)]^{1/2}\,\hbar. \tag{8.42}$$

Hence, we can write

$$\begin{aligned} L_+\,Y_{l,m} &= [l\,(l+1) - m\,(m+1)]^{1/2}\,\hbar\,Y_{l,m+1}, && (8.43)\\ L_-\,Y_{l,m} &= [l\,(l+1) - m\,(m-1)]^{1/2}\,\hbar\,Y_{l,m-1}. && (8.44)\end{aligned}$$

8.5 Eigenvalues of L_z

It seems reasonable to attempt to write the eigenstate $Y_{l,m}(\theta,\phi)$ in the separable form

$$Y_{l,m}(\theta,\phi) = \Theta_{l,m}(\theta)\,\Phi_m(\phi). \tag{8.45}$$

We can satisfy the orthonormality constraint (8.33) provided that

$$\begin{aligned} \int_{-\pi}^{\pi}\Theta^*_{l',m'}(\theta)\,\Theta_{l,m}(\theta)\,\sin\theta\,d\theta &= \delta_{ll'}, && (8.46)\\ \int_0^{2\pi}\Phi^*_{m'}(\phi)\,\Phi_m(\phi)\,d\phi &= \delta_{mm'}. && (8.47)\end{aligned}$$

Note, from Eq. (8.28), that the differential operator which represents L_z only depends on the azimuthal angle ϕ, and is independent of the polar angle θ. It therefore follows from Eqs. (8.28), (8.31), and (8.45) that

$$-\,i\,\hbar\frac{d\Phi_m}{d\phi} = m\,\hbar\,\Phi_m. \tag{8.48}$$

The solution to this equation is

$$\Phi_m(\phi) \sim e^{\,i\,m\,\phi}. \tag{8.49}$$

Here, the symbol $\sim$ just means that we are neglecting multiplicative constants.

Now, our basic interpretation of a wavefunction as a quantity whose modulus squared represents the probability density of finding a particle at a particular point in space suggests that a physical wavefunction must be *single-valued* in space. Otherwise, the probability density at a given point would not, in general, have a unique value, which does not make physical sense. Hence, we demand that the wavefunction (8.49) be single-valued: *i.e.*, $\Phi_m(\phi + 2\,\pi) = \Phi_m(\phi)$ for all ϕ. This immediately implies that the quantity m is *quantized*. In fact, m can only take *integer values*. Thus, we conclude that the eigenvalues of L_z are also quantized, and take the values $m\,\hbar$, where m is an integer. [A more rigorous argument is that $\Phi_m(\phi)$ must be continuous in order to ensure that L_z is an Hermitian operator, since the proof of hermiticity involves an integration by parts in ϕ that has canceling contributions from $\phi = 0$ and $\phi = 2\pi$.]

Finally, we can easily normalize the eigenstate (8.49) by making use of the orthonormality constraint (8.47). We obtain

$$\Phi_m(\phi) = \frac{e^{\,i\,m\,\phi}}{\sqrt{2\pi}}. \tag{8.50}$$

This is the properly normalized eigenstate of L_z corresponding to the eigenvalue $m\,\hbar$.

8.6 Eigenvalues of L^2

Consider the angular wavefunction $\psi(\theta, \phi) = L_+\,Y_{l,m}(\theta, \phi)$. We know that

$$\oint \psi^*(\theta, \phi)\,\psi(\theta, \phi)\,d\Omega \geq 0, \tag{8.51}$$

since $\psi^*\,\psi \equiv |\psi|^2$ is a positive-definite real quantity. Hence, making use of Eqs. (4.58) and (8.14), we find that

$$\begin{aligned} \oint (L_+\,Y_{l,m})^*\,(L_+\,Y_{l,m})\,d\Omega &= \oint Y_{l,m}^*\,(L_+)^\dagger\,(L_+\,Y_{l,m})\,d\Omega \\ &= \oint Y_{l,m}^*\,L_-\,L_+\,Y_{l,m}\,d\Omega \geq 0. \end{aligned} \tag{8.52}$$

It follows from Eqs. (8.16), and (8.31)–(8.33) that

$$\begin{aligned} \oint Y_{l,m}^*\,(L^2 - L_z^2 - \hbar\,L_z)\,Y_{l,m}\,d\Omega &= \oint Y_{l,m}^*\,\hbar^2\,[l\,(l+1) - m\,(m+1)]\,Y_{l,m}\,d\Omega \\ &= \hbar^2\;[l\,(l+1) - m\,(m+1)]\oint Y_{l,m}^*\,Y_{l,m}\,d\Omega \\ &= \hbar^2\;[l\,(l+1) - m\,(m+1)] \geq 0. \end{aligned} \tag{8.53}$$

We, thus, obtain the constraint

$$l\,(l+1) \geq m\,(m+1). \tag{8.54}$$

Likewise, the inequality

$$\oint (L_-\,Y_{l,m})^*\,(L_-\,Y_{l,m})\,d\Omega = \oint Y_{l,m}^*\,L_+\,L_-\,Y_{l,m}\,d\Omega \geq 0 \tag{8.55}$$

leads to a second constraint:

$$l\,(l+1) \geq m\,(m-1). \tag{8.56}$$

Without loss of generality, we can assume that $l \geq 0$. This is reasonable, from a physical standpoint, since $l\,(l+1)\,\hbar^2$ is supposed to represent the magnitude squared of something, and should, therefore, only take non-negative values. If l is non-negative then the constraints (8.54) and (8.56) are equivalent to the following constraint:

$$-\,l \leq m \leq l. \tag{8.57}$$

We, thus, conclude that the quantum number m can only take a *restricted range* of integer values.

Well, if m can only take a restricted range of integer values then there must exist a lowest possible value it can take. Let us call this special value m_-, and let Y_{l,m_-} be the corresponding eigenstate. Suppose we act on this eigenstate with the lowering operator L_-. According to Eq. (8.35), this will have the effect of converting the eigenstate into that of a state with a lower value of m. However, no such state exists. A non-existent state is represented in quantum mechanics by the null wavefunction, $\psi = 0$. Thus, we must have

$$L_-\,Y_{l,m_-} = 0. \tag{8.58}$$

Now, from Eq. (8.15),

$$L^2 = L_+\,L_- + L_z^2 - \hbar\,L_z \tag{8.59}$$

Hence,

$$L^2\,Y_{l,m_-} = (L_+\,L_- + L_z^2 - \hbar\,L_z)\,Y_{l,m_-}, \tag{8.60}$$

or

$$l\,(l+1)\,Y_{l,m_-} = m_-\,(m_- - 1)\,Y_{l,m_-}, \tag{8.61}$$

where use has been made of (8.31), (8.32), and (8.58). It follows that

$$l\,(l+1) = m_-\,(m_- - 1). \tag{8.62}$$

Assuming that m_- is negative, the solution to the above equation is

$$m_- = -l. \tag{8.63}$$

We can similarly show that the largest possible value of m is

$$m_+ = +l. \tag{8.64}$$

The above two results imply that l is an *integer*, since m_- and m_+ are both constrained to be integers.

We can now formulate the rules which determine the allowed values of the quantum numbers l and m. The quantum number l takes the non-negative integer values $0, 1, 2, 3, \cdots$. Once l is given, the quantum number m can take any integer value in the range

$$-\,l,\ -l+1,\ \cdots\ 0,\ \cdots,\ l-1,\ l. \tag{8.65}$$

Thus, if $l = 0$ then m can only take the value 0, if $l = 1$ then m can take the values $-1, 0, +1$, if $l = 2$ then m can take the values $-2, -1, 0, +1, +2$, and so on.

8.7 Spherical Harmonics

The simultaneous eigenstates, $Y_{l,m}(\theta, \phi)$, of L^2 and L_z are known as the *spherical harmonics*. Let us investigate their functional form.

Now, we know that

$$L_+\, Y_{l,l}(\theta, \phi) = 0, \tag{8.66}$$

since there is no state for which m has a larger value than $+l$. Writing

$$Y_{l,l}(\theta, \phi) = \Theta_{l,l}(\theta)\, e^{\,i\,l\,\phi} \tag{8.67}$$

[see Eqs. (8.45) and (8.49)], and making use of Eq. (8.30), we obtain

$$\hbar\, e^{\,i\,\phi} \left(\frac{\partial}{\partial\theta} + i\,\cot\theta\,\frac{\partial}{\partial\phi}\right) \Theta_{l,l}(\theta)\, e^{\,i\,l\,\phi} = 0. \tag{8.68}$$

This equation yields

$$\frac{d\Theta_{l,l}}{d\theta} - l\,\cot\theta\,\Theta_{l,l} = 0. \tag{8.69}$$

which can easily be solved to give

$$\Theta_{l,l} \sim (\sin\theta)^l. \tag{8.70}$$

Hence, we conclude that

$$Y_{l,l}(\theta, \phi) \sim (\sin\theta)^l\, e^{\,i\,l\,\phi}. \tag{8.71}$$

Likewise, it is easy to demonstrate that

$$Y_{l,-l}(\theta, \phi) \sim (\sin\theta)^l\, e^{-i\,l\,\phi}. \tag{8.72}$$

Once we know $Y_{l,l}$, we can obtain $Y_{l,l-1}$ by operating on $Y_{l,l}$ with the lowering operator L_-. Thus,

$$Y_{l,l-1} \sim L_- \, Y_{l,l} \sim e^{-i\,\phi}\left(-\frac{\partial}{\partial\theta} + i\,\cot\theta\,\frac{\partial}{\partial\phi}\right)(\sin\theta)^l\,e^{\,i\,l\,\phi}, \tag{8.73}$$

where use has been made of Eq. (8.30). The above equation yields

$$Y_{l,l-1} \sim e^{\,i\,(l-1)\,\phi}\left(\frac{d}{d\theta} + l\,\cot\theta\right)(\sin\theta)^l. \tag{8.74}$$

Now,

$$\left(\frac{d}{d\theta} + l\,\cot\theta\right)f(\theta) \equiv \frac{1}{(\sin\theta)^l}\frac{d}{d\theta}\left[(\sin\theta)^l\,f(\theta)\right], \tag{8.75}$$

where $f(\theta)$ is a general function. Hence, we can write

$$Y_{l,l-1}(\theta,\phi) \sim \frac{e^{\,i\,(l-1)\,\phi}}{(\sin\theta)^{l-1}}\left(\frac{1}{\sin\theta}\frac{d}{d\theta}\right)(\sin\theta)^{2\,l}. \tag{8.76}$$

Likewise, we can show that

$$Y_{l,-l+1}(\theta,\phi) \sim L_+\,Y_{l,-l} \sim \frac{e^{-i\,(l-1)\,\phi}}{(\sin\theta)^{l-1}}\left(\frac{1}{\sin\theta}\frac{d}{d\theta}\right)(\sin\theta)^{2\,l}. \tag{8.77}$$

We can now obtain $Y_{l,l-2}$ by operating on $Y_{l,l-1}$ with the lowering operator. We get

$$Y_{l,l-2} \sim L_-\,Y_{l,l-1} \sim e^{-i\,\phi}\left(-\frac{\partial}{\partial\theta} + i\,\cot\theta\,\frac{\partial}{\partial\phi}\right)\frac{e^{\,i\,(l-1)\,\phi}}{(\sin\theta)^{l-1}}\left(\frac{1}{\sin\theta}\frac{d}{d\theta}\right)(\sin\theta)^{2\,l}, \tag{8.78}$$

which reduces to

$$Y_{l,l-2} \sim e^{-i\,(l-2)\,\phi}\left[\frac{d}{d\theta} + (l-1)\,\cot\theta\right]\frac{1}{(\sin\theta)^{l-1}}\left(\frac{1}{\sin\theta}\frac{d}{d\theta}\right)(\sin\theta)^{2\,l}. \tag{8.79}$$

Finally, making use of Eq. (8.75), we obtain

$$Y_{l,l-2}(\theta,\phi) \sim \frac{e^{\,i\,(l-2)\,\phi}}{(\sin\theta)^{l-2}}\left(\frac{1}{\sin\theta}\frac{d}{d\theta}\right)^2(\sin\theta)^{2\,l}. \tag{8.80}$$

Likewise, we can show that

$$Y_{l,-l+2}(\theta,\phi) \sim L_+\,Y_{l,-l+1} \sim \frac{e^{-i\,(l-2)\,\phi}}{(\sin\theta)^{l-2}}\left(\frac{1}{\sin\theta}\frac{d}{d\theta}\right)^2(\sin\theta)^{2\,l}. \tag{8.81}$$

A comparison of Eqs. (8.71), (8.76), and (8.80) reveals the general functional form of the spherical harmonics:

$$Y_{l,m}(\theta,\phi) \sim \frac{e^{\,i\,m\,\phi}}{(\sin\theta)^m}\left(\frac{1}{\sin\theta}\frac{d}{d\theta}\right)^{l-m}(\sin\theta)^{2\,l}. \tag{8.82}$$

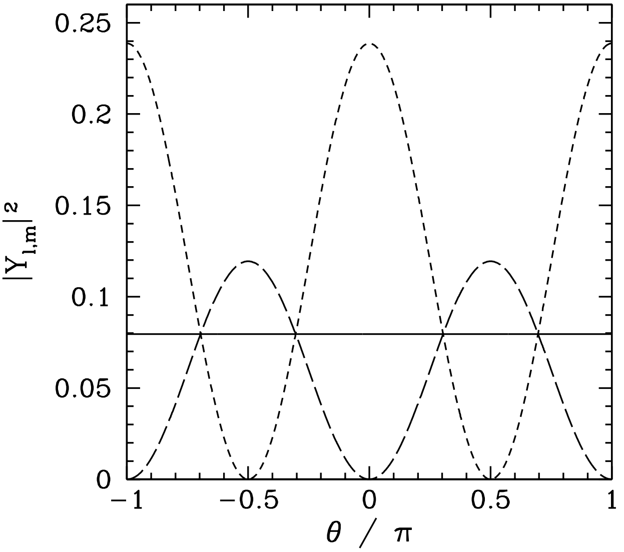

Figure 8.1: *The* $|Y_{l,m}(\theta,\phi)|^2$ *plotted as a functions of* θ*. The solid, short-dashed, and long-dashed curves correspond to* $l, m = 0, 0$*, and* $1, 0$*, and* $1, \pm 1$*, respectively.*

Here, m is assumed to be non-negative. Making the substitution $u = \cos\theta$, we can also write

$$Y_{l,m}(u,\phi) \sim e^{\,i\,m\,\phi}\,(1-u^2)^{-m/2}\left(\frac{d}{du}\right)^{l-m}(1-u^2)^{l}. \tag{8.83}$$

Finally, it is clear from Eqs. (8.72), (8.77), and (8.81) that

$$Y_{l,-m} \sim Y_{l,m}^{*}. \tag{8.84}$$

We now need to normalize our spherical harmonic functions so as to ensure that

$$\oint |Y_{l,m}(\theta,\phi)|^2\,d\Omega = 1. \tag{8.85}$$

After a great deal of tedious analysis, the normalized spherical harmonic functions are found to take the form

$$Y_{l,m}(\theta,\phi) = (-1)^m\left[\frac{2\,l+1}{4\pi}\,\frac{(l-m)!}{(l+m)!}\right]^{1/2} P_{l,m}(\cos\theta)\,e^{\,i\,m\,\phi} \tag{8.86}$$

for $m \geq 0$, where the $P_{l,m}$ are known as *associated Legendre polynomials*, and are written

$$P_{l,m}(u) = (-1)^{l+m}\,\frac{(l+m)!}{(l-m)!}\,\frac{(1-u^2)^{-m/2}}{2^l\,l!}\left(\frac{d}{du}\right)^{l-m}(1-u^2)^{l} \tag{8.87}$$

for $m \geq 0$. Alternatively,

$$P_{l,m}(u) = (-1)^{l}\,\frac{(1-u^2)^{m/2}}{2^l\,l!}\left(\frac{d}{du}\right)^{l+m}(1-u^2)^{l}, \tag{8.88}$$

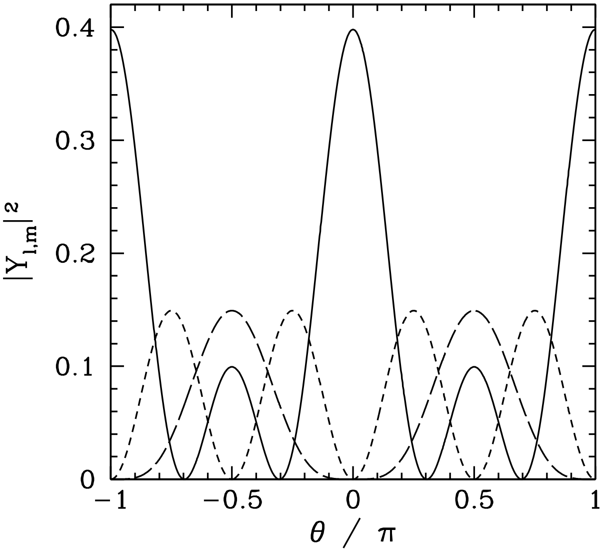

Figure 8.2: *The* $|Y_{l,m}(\theta,\phi)|^2$ *plotted as a functions of* θ. *The solid, short-dashed, and long-dashed curves correspond to* $l, m = 2, 0$, *and* $2, \pm 1$, *and* $2, \pm 2$, *respectively.*

for $m \geq 0$. The spherical harmonics characterized by $m < 0$ can be calculated from those characterized by $m > 0$ via the identity

$$Y_{l,-m} = (-1)^m\, Y_{l,m}^{*}. \tag{8.89}$$

The spherical harmonics are *orthonormal*: *i.e.*,

$$\oint Y_{l',m'}^{*}\, Y_{l,m}\, d\Omega = \delta_{ll'}\, \delta_{mm'}, \tag{8.90}$$

and also form a *complete set*. In other words, *any* function of θ and ϕ can be represented as a superposition of spherical harmonics. Finally, and most importantly, the spherical harmonics are the simultaneous eigenstates of L_z and L^2 corresponding to the eigenvalues $m\,\hbar$ and $l\,(l+1)\,\hbar^2$, respectively.

All of the $l = 0$, $l = 1$, and $l = 2$ spherical harmonics are listed below:

$$Y_{0,0} = \frac{1}{\sqrt{4\pi}}, \tag{8.91}$$

$$Y_{1,0} = \sqrt{\frac{3}{4\pi}}\,\cos\theta, \tag{8.92}$$

$$Y_{1,\pm 1} = \mp\sqrt{\frac{3}{8\pi}}\,\sin\theta\, e^{\pm i\,\phi}, \tag{8.93}$$

$$Y_{2,0} = \sqrt{\frac{5}{16\pi}}\,(3\,\cos^2\theta - 1), \tag{8.94}$$

$$Y_{2,\pm1} = \mp\sqrt{\frac{15}{8\pi}}\,\sin\theta\,\cos\theta\,e^{\pm i\,\phi}, \tag{8.95}$$

$$Y_{2,\pm2} = \sqrt{\frac{15}{32\pi}}\,\sin^2\theta\,e^{\pm 2i\,\phi}. \tag{8.96}$$

The θ variation of these functions is illustrated in Figs. 8.1 and 8.2.

Exercises

1. A system is in the state $\psi = Y_{l,m}(\theta,\phi)$. Calculate $\langle L_x\rangle$ and $\langle L_x^2\rangle$.

2. Find the eigenvalues and eigenfunctions (in terms of the angles θ and ϕ) of L_x.

3. Consider a beam of particles with $l = 1$. A measurement of L_x yields the result $\hbar$. What values will be obtained by a subsequent measurement of L_z, and with what probabilities? Repeat the calculation for the cases in which the measurement of L_x yields the results 0 and $-\hbar$.

4. The Hamiltonian for an axially symmetric rotator is given by

$$H = \frac{L_x^{\,2}+L_y^{\,2}}{2\,I_1} + \frac{L_z^{\,2}}{2\,I_2}.$$

   What are the eigenvalues of H?

9 Central Potentials

9.1 Introduction

In this chapter, we shall investigate the interaction of a non-relativistic particle of mass m and energy E with various so-called *central potentials*, $V(r)$, where $r = \sqrt{x^2+y^2+z^2}$ is the radial distance from the origin. It is, of course, most convenient to work in spherical polar coordinates—r, θ, ϕ—during such an investigation (see Sect. 8.3). Thus, we shall be searching for stationary wavefunctions, $\psi(r,\theta,\phi)$, which satisfy the time-independent Schrödinger equation (see Sect. 4.12)

$$H\,\psi = E\,\psi, \tag{9.1}$$

where the Hamiltonian takes the standard non-relativistic form

$$H = \frac{p^2}{2\,m} + V(r). \tag{9.2}$$

9.2 Derivation of Radial Equation

Now, we have seen that the Cartesian components of the momentum, $\mathbf{p}$, can be represented as (see Sect. 7.2)

$$p_i = -i\,\hbar\,\frac{\partial}{\partial x_i} \tag{9.3}$$

for $i = 1,2,3$, where $x_1 \equiv x$, $x_2 \equiv y$, $x_3 \equiv z$, and $\mathbf{r} \equiv (x_1, x_2, x_3)$. Likewise, it is easily demonstrated, from the above expressions, and the basic definitions of the spherical polar coordinates [see Eqs. (8.20)–(8.25)], that the radial component of the momentum can be represented as

$$p_r \equiv \frac{\mathbf{p}\cdot\mathbf{r}}{r} = -i\,\hbar\,\frac{\partial}{\partial r}. \tag{9.4}$$

Recall that the angular momentum vector, $\mathbf{L}$, is defined [see Eq. (8.1)]

$$\mathbf{L} = \mathbf{r}\times\mathbf{p}. \tag{9.5}$$

This expression can also be written in the following form:

$$L_i = \epsilon_{ijk}\,x_j\,p_k. \tag{9.6}$$

Here, the ϵ_{ijk} (where i, j, k all run from 1 to 3) are elements of the so-called *totally anti-symmetric tensor*. The values of the various elements of this tensor are determined via a simple rule:

$$\epsilon_{ijk} = \left\{ \begin{array}{cl} 0 & \text{if } i,j,k \text{ not all different} \\ 1 & \text{if } i,j,k \text{ are cyclic permutation of } 1,2,3 \\ -1 & \text{if } i,j,k \text{ are anti-cyclic permutation of } 1,2,3 \end{array} \right. . \tag{9.7}$$

Thus, $\epsilon_{123} = \epsilon_{231} = 1$, $\epsilon_{321} = \epsilon_{132} = -1$, and $\epsilon_{112} = \epsilon_{131} = 0$, *etc.* Equation (9.6) also makes use of the *Einstein summation convention*, according to which repeated indices are summed (from 1 to 3). For instance, $a_i\,b_i \equiv a_1\,b_1 + a_2\,b_2 + a_3\,b_3$. Making use of this convention, as well as Eq. (9.7), it is easily seen that Eqs. (9.5) and (9.6) are indeed equivalent.

Let us calculate the value of L^2 using Eq. (9.6). According to our new notation, L^2 is the same as $L_i\,L_i$. Thus, we obtain

$$L^2 = \epsilon_{ijk}\,x_j\,p_k\,\epsilon_{ilm}\,x_l\,p_m = \epsilon_{ijk}\,\epsilon_{ilm}\,x_j\,p_k\,x_l\,p_m. \tag{9.8}$$

Note that we are able to shift the position of ϵ_{ilm} because its elements are just numbers, and, therefore, commute with all of the x_i and the p_i. Now, it is easily demonstrated that

$$\epsilon_{ijk}\,\epsilon_{ilm} \equiv \delta_{jl}\,\delta_{km} - \delta_{jm}\,\delta_{kl}. \tag{9.9}$$

Here δ_{ij} is the usual *Kronecker delta*, whose elements are determined according to the rule

$$\delta_{ij} = \left\{ \begin{array}{lll} 1 & & \text{if } i \text{ and } j \text{ the same} \\ 0 & & \text{if } i \text{ and } j \text{ different} \end{array} \right. . \tag{9.10}$$

It follows from Eqs. (9.8) and (9.9) that

$$L^2 = x_i\,p_j\,x_i\,p_j - x_i\,p_j\,x_j\,p_i. \tag{9.11}$$

Here, we have made use of the fairly self-evident result that $\delta_{ij}\,a_i\,b_j \equiv a_i\,b_i$. We have also been careful to preserve the order of the various terms on the right-hand side of the above expression, since the x_i and the p_i do not necessarily commute with one another.

We now need to rearrange the order of the terms on the right-hand side of Eq. (9.11). We can achieve this by making use of the fundamental commutation relation for the x_i and the p_i [see Eq. (7.17)]:

$$[x_i, p_j] = i\,\hbar\,\delta_{ij}. \tag{9.12}$$

Thus,

$$\begin{aligned} L^2 &= x_i\,(x_i\,p_j - [x_i, p_j])\,p_j - x_i\,p_j\,\left(p_i\,x_j + [x_j, p_i]\right) \\ &= x_i\,x_i\,p_j\,p_j - i\,\hbar\,\delta_{ij}\,x_i\,p_j - x_i\,p_j\,p_i\,x_j - i\,\hbar\,\delta_{ij}\,x_i\,p_j \\ &= x_i\,x_i\,p_j\,p_j - x_i\,p_i\,p_j\,x_j - 2\,i\,\hbar\,x_i\,p_i. \end{aligned} \tag{9.13}$$

Here, we have made use of the fact that $p_j\,p_i = p_i\,p_j$, since the p_i commute with one another [see Eq. (7.16)]. Next,

$$L^2 = x_i\,x_i\,p_j\,p_j - x_i\,p_i\,(x_j\,p_j - [x_j, p_j]) - 2\,i\,\hbar\,x_i\,p_i. \tag{9.14}$$

Now, according to (9.12),

$$[x_j, p_j] \equiv [x_1, p_1] + [x_2, p_2] + [x_3, p_3] = 3\,i\,\hbar. \tag{9.15}$$

Hence, we obtain

$$L^2 = x_i\,x_i\,p_j\,p_j - x_i\,p_i\,x_j\,p_j + i\,\hbar\,x_i\,p_i. \tag{9.16}$$

When expressed in more conventional vector notation, the above expression becomes

$$L^2 = r^2\,p^2 - (\mathbf{r}\cdot\mathbf{p})^2 + i\,\hbar\,\mathbf{r}\cdot\mathbf{p}. \tag{9.17}$$

Note that if we had attempted to derive the above expression directly from Eq. (9.5), using standard vector identities, then we would have missed the final term on the right-hand side. This term originates from the lack of commutation between the x_i and p_i operators in quantum mechanics. Of course, standard vector analysis assumes that all terms commute with one another.

Equation (9.17) can be rearranged to give

$$p^2 = r^{-2}\left[(\mathbf{r}\cdot\mathbf{p})^2 - i\,\hbar\,\mathbf{r}\cdot\mathbf{p} + L^2\right]. \tag{9.18}$$

Now,

$$\mathbf{r}\cdot\mathbf{p} = r\,p_r = -i\,\hbar\,r\,\frac{\partial}{\partial r}, \tag{9.19}$$

where use has been made of Eq. (9.4). Hence, we obtain

$$p^2 = -\hbar^2\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\,\frac{\partial}{\partial r}\right) + \frac{1}{r}\frac{\partial}{\partial r} - \frac{L^2}{\hbar^2\,r^2}\right]. \tag{9.20}$$

Finally, the above equation can be combined with Eq. (9.2) to give the following expression for the Hamiltonian:

$$H = -\frac{\hbar^2}{2\,m}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} - \frac{L^2}{\hbar^2\,r^2}\right) + V(r). \tag{9.21}$$

Let us now consider whether the above Hamiltonian commutes with the angular momentum operators L_z and L^2. Recall, from Sect. 8.3, that L_z and L^2 are represented as differential operators which depend solely on the angular spherical polar coordinates, θ and ϕ, and do not contain the radial polar coordinate, r. Thus, any function of r, or any differential operator involving r (but not θ and ϕ), will automatically commute with L^2 and L_z. Moreover, L^2 commutes both with itself, and with L_z (see Sect. 8.2). It is, therefore, clear that the above Hamiltonian commutes with both L_z and L^2.

Now, according to Sect. 4.10, if two operators commute with one another then they possess simultaneous eigenstates. We thus conclude that *for a particle moving in a central potential the eigenstates of the Hamiltonian are simultaneous eigenstates of* L_z *and* L^2. Now, we have already found the simultaneous eigenstates of L_z and L^2—they are the spherical harmonics, $Y_{l,m}(\theta,\phi)$, discussed in Sect. 8.7. It follows that the spherical harmonics are also eigenstates of the Hamiltonian. This observation leads us to try the following separable form for the stationary wavefunction:

$$\psi(r,\theta,\phi) = R(r)\,Y_{l,m}(\theta,\phi). \tag{9.22}$$

It immediately follows, from (8.31) and (8.32), and the fact that L_z and L^2 both obviously commute with $R(r)$, that

$$L_z\,\psi = m\,\hbar\,\psi, \tag{9.23}$$

$$L^2\,\psi = l\,(l+1)\,\hbar^2\,\psi. \tag{9.24}$$

Recall that the quantum numbers m and l are restricted to take certain integer values, as explained in Sect. 8.6.

Finally, making use of Eqs. (9.1), (9.21), and (9.24), we obtain the following differential equation which determines the radial variation of the stationary wavefunction:

$$-\frac{\hbar^2}{2\,m}\left(\frac{d^2}{dr^2}+\frac{2}{r}\frac{d}{dr}-\frac{l\,(l+1)}{r^2}\right)R_{n,l}+V\,R_{n,l}=E\,R_{n,l}. \tag{9.25}$$

Here, we have labeled the function $R(r)$ by two quantum numbers, n and l. The second quantum number, l, is, of course, related to the eigenvalue of L^2. [Note that the azimuthal quantum number, m, does not appear in the above equation, and, therefore, does not influence either the function $R(r)$ or the energy, E.] As we shall see, the first quantum number, n, is determined by the constraint that the radial wavefunction be square-integrable.

9.3 Infinite Spherical Potential Well

Consider a particle of mass m and energy $E > 0$ moving in the following simple central potential:

$$V(r)=\left\{\begin{array}{ll}0 & \text{for } 0\leq r\leq a\\ \infty & \text{otherwise}\end{array}\right. . \tag{9.26}$$

Clearly, the wavefunction ψ is only non-zero in the region $0 \leq r \leq a$. Within this region, it is subject to the physical boundary conditions that it be well behaved (*i.e.*, square-integrable) at $r = 0$, and that it be zero at $r = a$ (see Sect. 5.2). Writing the wavefunction in the standard form

$$\psi(r,\theta,\phi)=R_{n,l}(r)\,Y_{l,m}(\theta,\phi), \tag{9.27}$$

we deduce (see previous section) that the radial function $R_{n,l}(r)$ satisfies

$$\frac{d^2R_{n,l}}{dr^2}+\frac{2}{r}\frac{dR_{n,l}}{dr}+\left(k^2-\frac{l\,(l+1)}{r^2}\right)R_{n,l}=0 \tag{9.28}$$

in the region $0 \leq r \leq a$, where

$$k^2=\frac{2\,m\,E}{\hbar^2}. \tag{9.29}$$

Defining the scaled radial variable $z = k\,r$, the above differential equation can be transformed into the standard form

$$\frac{d^2R_{n,l}}{dz^2}+\frac{2}{z}\frac{dR_{n,l}}{dz}+\left[1-\frac{l\,(l+1)}{z^2}\right]R_{n,l}=0. \tag{9.30}$$

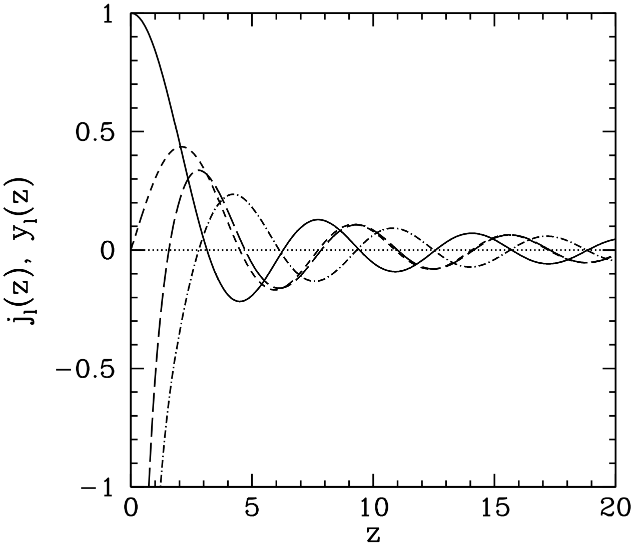

Figure 9.1: *The first few spherical Bessel functions. The solid, short-dashed, long-dashed, and dot-dashed curves show* $j_0(z)$, $j_1(z)$, $y_0(z)$, *and* $y_1(z)$, *respectively.*

The two independent solutions to this well-known second-order differential equation are called *spherical Bessel functions,*[1] and can be written

$$j_l(z) = z^l\left(-\frac{1}{z}\frac{d}{dz}\right)^l\left(\frac{\sin z}{z}\right), \tag{9.31}$$

$$y_l(z) = -z^l\left(-\frac{1}{z}\frac{d}{dz}\right)^l\left(\frac{\cos z}{z}\right). \tag{9.32}$$

Thus, the first few spherical Bessel functions take the form

$$j_0(z) = \frac{\sin z}{z}, \tag{9.33}$$

$$j_1(z) = \frac{\sin z}{z^2} - \frac{\cos z}{z}, \tag{9.34}$$

$$y_0(z) = -\frac{\cos z}{z}, \tag{9.35}$$

$$y_1(z) = -\frac{\cos z}{z^2} - \frac{\sin z}{z}. \tag{9.36}$$

These functions are also plotted in Fig. 9.1. It can be seen that the spherical Bessel functions are oscillatory in nature, passing through zero many times. However, the $y_l(z)$ functions are badly behaved (*i.e.*, they are not square-integrable) at $z = 0$, whereas the $j_l(z)$

---

[1] M. Abramowitz, and I.A. Stegun, *Handbook of Mathematical Functions* (Dover, New York NY, 1965), Sect. 10.1.

	$n=1$	$n=2$	$n=3$	$n=4$
$l=0$	3.142	6.283	9.425	12.566
$l=1$	4.493	7.725	10.904	14.066
$l=2$	5.763	9.095	12.323	15.515
$l=3$	6.988	10.417	13.698	16.924
$l=4$	8.183	11.705	15.040	18.301

Table 9.1: *The first few zeros of the spherical Bessel function* $j_l(z)$.

functions are well behaved everywhere. It follows from our boundary condition at $r = 0$ that the $y_l(z)$ are unphysical, and that the radial wavefunction $R_{n,l}(r)$ is thus proportional to $j_l(k\,r)$ only. In order to satisfy the boundary condition at $r = a$ [*i.e.*, $R_{n,l}(a) = 0$], the value of k must be chosen such that $z = k\,a$ corresponds to one of the zeros of $j_l(z)$. Let us denote the nth zero of $j_l(z)$ as $z_{n,l}$. It follows that

$$k\,a = z_{n,l}, \tag{9.37}$$

for $n = 1, 2, 3, \ldots$. Hence, from (9.29), the allowed energy levels are

$$E_{n,l} = z_{n,l}^2\,\frac{\hbar^2}{2\,m\,a^2}. \tag{9.38}$$

The first few values of $z_{n,l}$ are listed in Table 9.1. It can be seen that $z_{n,l}$ is an increasing function of both n and l.

We are now in a position to interpret the three quantum numbers—n, l, and m—which determine the form of the wavefunction specified in Eq. (9.27). As is clear from Sect. 8, the azimuthal quantum number m determines the number of nodes in the wavefunction as the azimuthal angle ϕ varies between 0 and 2π. Thus, $m = 0$ corresponds to no nodes, $m = 1$ to a single node, $m = 2$ to two nodes, *etc*. Likewise, the polar quantum number l determines the number of nodes in the wavefunction as the polar angle θ varies between 0 and π. Again, $l = 0$ corresponds to no nodes, $l = 1$ to a single node, *etc*. Finally, the radial quantum number n determines the number of nodes in the wavefunction as the radial variable r varies between 0 and a (not counting any nodes at $r = 0$ or $r = a$). Thus, $n = 1$ corresponds to no nodes, $n = 2$ to a single node, $n = 3$ to two nodes, *etc*. Note that, for the case of an infinite potential well, the only restrictions on the values that the various quantum numbers can take are that n must be a positive integer, l must be a non-negative integer, and m must be an integer lying between $-l$ and l. Note, further, that the allowed energy levels (9.38) only depend on the values of the quantum numbers n and l. Finally, it is easily demonstrated that the spherical Bessel functions are mutually orthogonal: *i.e.*,

$$\int_0^a j_l(z_{n,l}\,r/a)\,j_l(z_{n',l}\,r/a)\,r^2\,dr = 0 \tag{9.39}$$

when $n \neq n'$. Given that the $Y_{l,m}(\theta, \phi)$ are mutually orthogonal (see Sect. 8), this ensures that wavefunctions (9.27) corresponding to distinct sets of values of the quantum numbers n, l, and m are mutually orthogonal.

9.4 Hydrogen Atom

A hydrogen atom consists of an electron, of charge $-e$ and mass m_e, and a proton, of charge $+e$ and mass m_p, moving in the Coulomb potential

$$V(\mathbf{r}) = -\frac{e^2}{4\pi\,\epsilon_0\,|\mathbf{r}|}, \tag{9.40}$$

where $\mathbf{r}$ is the position vector of the electron with respect to the proton. Now, according to the analysis in Sect. 6.4, this two-body problem can be converted into an equivalent one-body problem. In the latter problem, a particle of mass

$$\mu = \frac{m_e\,m_p}{m_e + m_p} \tag{9.41}$$

moves in the central potential

$$V(r) = -\frac{e^2}{4\pi\,\epsilon_0\,r}. \tag{9.42}$$

Note, however, that since $m_e/m_p \simeq 1/1836$ the difference between m_e and μ is very small. Hence, in the following, we shall write neglect this difference entirely.

Writing the wavefunction in the usual form,

$$\psi(r,\theta,\phi) = R_{n,l}(r)\,Y_{l,m}(\theta,\phi), \tag{9.43}$$

it follows from Sect. 9.2 that the radial function $R_{n,l}(r)$ satisfies

$$-\frac{\hbar^2}{2\,m_e}\left(\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{l\,(l+1)}{r^2}\right)R_{n,l} - \left(\frac{e^2}{4\pi\,\epsilon_0\,r} + E\right)R_{n,l} = 0. \tag{9.44}$$

Let $r = a\,z$, with

$$a = \sqrt{\frac{\hbar^2}{2\,m_e\,(-E)}} = \sqrt{\frac{E_0}{E}}\,a_0, \tag{9.45}$$

where E_0 and a_0 are defined in Eqs. (9.57) and (9.58), respectively. Here, it is assumed that $E < 0$, since we are only interested in bound-states of the hydrogen atom. The above differential equation transforms to

$$\left(\frac{d^2}{dz^2} + \frac{2}{z}\frac{d}{dz} - \frac{l\,(l+1)}{z^2} + \frac{\zeta}{z} - 1\right)R_{n,l} = 0, \tag{9.46}$$

where

$$\zeta = \frac{2\,m_e\,a\,e^2}{4\pi\,\epsilon_0\,\hbar^2} = 2\sqrt{\frac{E_0}{E}}. \tag{9.47}$$

Suppose that $R_{n,l}(r) = Z(r/a)\,\exp(-r/a)/(r/a)$. It follows that

$$\left(\frac{d^2}{dz^2} - 2\,\frac{d}{dz} - \frac{l\,(l+1)}{z^2} + \frac{\zeta}{z}\right)Z = 0. \tag{9.48}$$

We now need to solve the above differential equation in the domain $z = 0$ to $z = \infty$, subject to the constraint that $R_{n,l}(r)$ be square-integrable.

Let us look for a power-law solution of the form

$$Z(z) = \sum_k c_k\, z^k. \tag{9.49}$$

Substituting this solution into Eq. (9.48), we obtain

$$\sum_k c_k \left\{k\,(k-1)\,z^{k-2} - 2\,k\,z^{k-1} - l\,(l+1)\,z^{k-2} + \zeta\,z^{k-1}\right\} = 0. \tag{9.50}$$

Equating the coefficients of z^{k-2} gives the recursion relation

$$c_k\;[k\,(k-1) - l\,(l+1)] = c_{k-1}\;[2\,(k-1) - \zeta]\,. \tag{9.51}$$

Now, the power series (9.49) must terminate at small k, at some positive value of k, otherwise $Z(z)$ behaves unphysically as $z \to 0$ [*i.e.*, it yields an $R_{n,l}(r)$ that is not square-integrable as $r \to 0$]. From the above recursion relation, this is only possible if $[k_{min}\,(k_{min} - 1) - l\,(l+1)] = 0$, where the first term in the series is $c_{k_{min}}\, z^{k_{min}}$. There are two possibilities: $k_{min} = -l$ or $k_{min} = l + 1$. However, the former possibility predicts unphysical behaviour of $Z(z)$ at $z = 0$. Thus, we conclude that $k_{min} = l + 1$. Note that, since $R_{n,l}(r) \simeq Z(r/a)/(r/a) \simeq (r/a)^l$ at small r, there is a finite probability of finding the electron at the nucleus for an $l = 0$ state, whereas there is zero probability of finding the electron at the nucleus for an $l > 0$ state [*i.e.*, $|\psi|^2 = 0$ at $r = 0$, except when $l = 0$].

For large values of z, the ratio of successive coefficients in the power series (9.49) is

$$\frac{c_k}{c_{k-1}} = \frac{2}{k}, \tag{9.52}$$

according to Eq. (9.51). This is the same as the ratio of successive coefficients in the power series

$$\sum_k \frac{(2\,z)^k}{k!}, \tag{9.53}$$

which converges to $\exp(2\,z)$. We conclude that $Z(z) \to \exp(2\,z)$ as $z \to \infty$. It thus follows that $R_{n,l}(r) \sim Z(r/a)\,\exp(-r/a)/(r/a) \to \exp(r/a)/(r/a)$ as $r \to \infty$. This does not correspond to physically acceptable behaviour of the wavefunction, since $\int |\psi|^2\, dV$ must be finite. The only way in which we can avoid this unphysical behaviour is if the power series (9.49) *terminates* at some maximum value of k. According to the recursion relation (9.51), this is only possible if

$$\frac{\zeta}{2} = n, \tag{9.54}$$

where n is an integer, and the last term in the series is $c_n\, z^n$. Since the first term in the series is $c_{l+1}\, z^{l+1}$, it follows that n must be greater than l, otherwise there are no terms in the series at all. Finally, it is clear from Eqs. (9.45), (9.47), and (9.54) that

$$E = \frac{E_0}{n^2} \tag{9.55}$$

and

$$a = n\,a_0, \tag{9.56}$$

where

$$E_0 = -\frac{m_e\,e^4}{2\,(4\pi\,\epsilon_0)^2\,\hbar^2} = -\frac{e^2}{8\pi\,\epsilon_0\,a_0} = -13.6\,\text{eV}, \tag{9.57}$$

and

$$a_0 = \frac{4\pi\,\epsilon_0\,\hbar^2}{m_e\,e^2} = 5.3\times 10^{-11}\,\text{m}. \tag{9.58}$$

Here, E_0 is the energy of so-called *ground-state* (or lowest energy state) of the hydrogen atom, and the length a_0 is known as the *Bohr radius*. Note that $|E_0| \sim \alpha^2\,m_e\,c^2$, where $\alpha = e^2/(4\pi\,\epsilon_0\,\hbar\,c) \simeq 1/137$ is the dimensionless *fine-structure constant*. The fact that $|E_0| \ll m_e\,c^2$ is the ultimate justification for our non-relativistic treatment of the hydrogen atom.

We conclude that the wavefunction of a hydrogen atom takes the form

$$\psi_{n,l,m}(r,\theta,\phi) = R_{n,l}(r)\,Y_{l,m}(\theta,\phi). \tag{9.59}$$

Here, the $Y_{l,m}(\theta,\phi)$ are the spherical harmonics (see Sect 8.7), and $R_{n,l}(z = r/a)$ is the solution of

$$\left(\frac{1}{z^2}\,\frac{d}{dz}\,z^2\,\frac{d}{dz} - \frac{l\,(l+1)}{z^2} + \frac{2\,n}{z} - 1\right)R_{n,l} = 0 \tag{9.60}$$

which varies as z^l at small z. Furthermore, the quantum numbers n, l, and m can only take values which satisfy the inequality

$$|m| \leq l < n, \tag{9.61}$$

where n is a positive integer, l a non-negative integer, and m an integer.

Now, we expect the stationary states of the hydrogen atom to be orthonormal: *i.e.*,

$$\int \psi^\ast_{n',l',m'}\,\psi_{n,l,m}\,dV = \delta_{nn'}\,\delta_{ll'}\,\delta_{mm'}, \tag{9.62}$$

where dV is a volume element, and the integral is over all space. Of course, $dV = r^2\,dr\,d\Omega$, where $d\Omega$ is an element of solid angle. Moreover, we already know that the spherical harmonics are orthonormal [see Eq. (8.90)]: *i.e.*,

$$\oint Y^\ast_{l',m'}\,Y_{l,m}\,d\Omega = \delta_{ll'}\,\delta_{mm'}. \tag{9.63}$$

It, thus, follows that the radial wavefunction satisfies the orthonormality constraint

$$\int_0^\infty R^\ast_{n',l}\,R_{n,l}\,r^2\,dr = \delta_{nn'}. \tag{9.64}$$

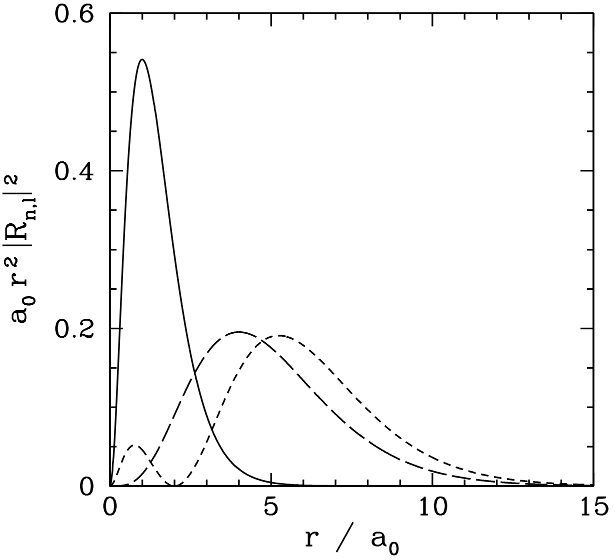

Figure 9.2: *The* $a_0\,r^2\,|R_{n,l}(r)|^2$ *plotted as a functions of* r/a_0*. The solid, short-dashed, and long-dashed curves correspond to* $n, l = 1, 0$*, and* $2, 0$*, and* $2, 1$*, respectively.*

The first few radial wavefunctions for the hydrogen atom are listed below:

$$R_{1,0}(r) = \frac{2}{a_0^{3/2}}\,\exp\left(-\frac{r}{a_0}\right), \tag{9.65}$$

$$R_{2,0}(r) = \frac{2}{(2\,a_0)^{3/2}}\left(1-\frac{r}{2\,a_0}\right)\exp\left(-\frac{r}{2\,a_0}\right), \tag{9.66}$$

$$R_{2,1}(r) = \frac{1}{\sqrt{3}\,(2\,a_0)^{3/2}}\,\frac{r}{a_0}\,\exp\left(-\frac{r}{2\,a_0}\right), \tag{9.67}$$

$$R_{3,0}(r) = \frac{2}{(3\,a_0)^{3/2}}\left(1-\frac{2\,r}{3\,a_0}+\frac{2\,r^2}{27\,a_0^2}\right)\exp\left(-\frac{r}{3\,a_0}\right), \tag{9.68}$$

$$R_{3,1}(r) = \frac{4\sqrt{2}}{9\,(3\,a_0)^{3/2}}\,\frac{r}{a_0}\left(1-\frac{r}{6\,a_0}\right)\,\exp\left(-\frac{r}{3\,a_0}\right), \tag{9.69}$$

$$R_{3,2}(r) = \frac{2\sqrt{2}}{27\sqrt{5}\,(3\,a_0)^{3/2}}\left(\frac{r}{a_0}\right)^2\exp\left(-\frac{r}{3\,a_0}\right). \tag{9.70}$$

These functions are illustrated in Figs. 9.2 and 9.3.

Given the (properly normalized) hydrogen wavefunction (9.59), plus our interpretation of $|\psi|^2$ as a probability density, we can calculate

$$\langle r^k\rangle = \int_0^\infty r^{2+k}\,|R_{n,l}(r)|^2\,dr, \tag{9.71}$$

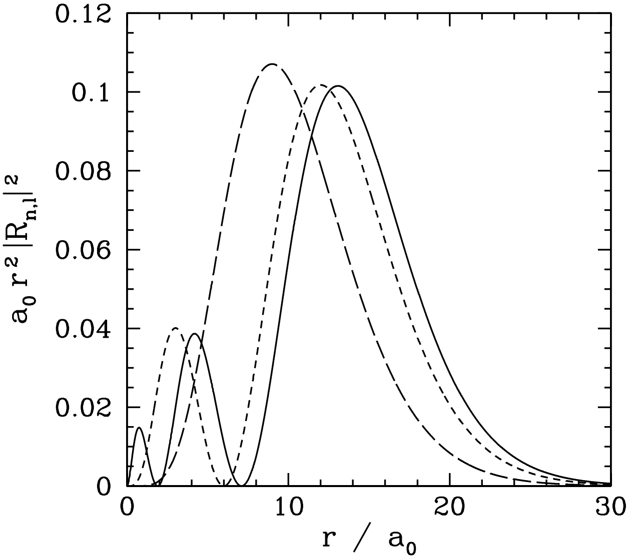

Figure 9.3: *The* $a_0\, r^2\, |R_{n,l}(r)|^2$ *plotted as a functions of* r/a_0*. The solid, short-dashed, and long-dashed curves correspond to* $n, l = 3, 0$*, and* $3, 1$*, and* $3, 2$*, respectively.*

where the angle-brackets denote an expectation value. For instance, it can be demonstrated (after much tedious algebra) that

$$\langle r^2\rangle = \frac{a_0^2\, n^2}{2}\, [5\, n^2 + 1 - 3\, l\, (l+1)], \tag{9.72}$$

$$\langle r\rangle = \frac{a_0}{2}\, [3\, n^2 - l\, (l+1)], \tag{9.73}$$

$$\left\langle \frac{1}{r}\right\rangle = \frac{1}{n^2\, a_0}, \tag{9.74}$$

$$\left\langle \frac{1}{r^2}\right\rangle = \frac{1}{(l+1/2)\, n^3\, a_0^2}, \tag{9.75}$$

$$\left\langle \frac{1}{r^3}\right\rangle = \frac{1}{l\, (l+1/2)\, (l+1)\, n^3\, a_0^3}. \tag{9.76}$$

According to Eq. (9.55), the energy levels of the bound-states of a hydrogen atom only depend on the radial quantum number n. It turns out that this is a special property of a $1/r$ potential. For a general central potential, $V(r)$, the quantized energy levels of a bound-state depend on both n and l (see Sect. 9.3).

The fact that the energy levels of a hydrogen atom only depend on n, and not on l and m, implies that the energy spectrum of a hydrogen atom is *highly degenerate*: *i.e.*, there are many different states which possess the same energy. According to the inequality (9.61) (and the fact that n, l, and m are integers), for a given value of l, there are $2\,l+1$ different allowed values of m (*i.e.*, $-l, -l+1, \cdots, l-1, l$). Likewise, for a given value of n, there are n different allowed values of l (*i.e.*, $0, 1, \cdots, n-1$). Now, all states possessing the same value of n have the same energy (*i.e.*, they are degenerate). Hence, the total number of degenerate states corresponding to a given value of n is

$$1 + 3 + 5 + \cdots + 2\,(n-1) + 1 = n^2. \tag{9.77}$$

Thus, the ground-state ($n = 1$) is not degenerate, the first excited state ($n = 2$) is four-fold degenerate, the second excited state ($n = 3$) is nine-fold degenerate, *etc.* [Actually, when we take into account the two spin states of an electron (see Sect. 10), the degeneracy of the nth energy level becomes $2\,n^2$.]

9.5 Rydberg Formula

An electron in a given stationary state of a hydrogen atom, characterized by the quantum numbers n, l, and m, should, in principle, remain in that state indefinitely. In practice, if the state is slightly perturbed—*e.g.*, by interacting with a photon—then the electron can make a transition to another stationary state with different quantum numbers.

Suppose that an electron in a hydrogen atom makes a transition from an initial state whose radial quantum number is n_i to a final state whose radial quantum number is n_f. According to Eq. (9.55), the energy of the electron will change by

$$\Delta E = E_0 \left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right). \tag{9.78}$$

If ΔE is negative then we would expect the electron to *emit* a photon of frequency $\nu = -\Delta E/h$ [see Eq. (3.32)]. Likewise, if ΔE is positive then the electron must *absorb* a photon of energy $\nu = \Delta E/h$. Given that $\lambda^{-1} = \nu/c$, the possible wavelengths of the photons emitted by a hydrogen atom as its electron makes transitions between different energy levels are

$$\frac{1}{\lambda} = R\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right), \tag{9.79}$$

where

$$R = \frac{-E_0}{h\,c} = \frac{m_e\,e^4}{(4\pi)^3\,\epsilon_0^2\,\hbar^3\,c} = 1.097 \times 10^7\,\mathrm{m}^{-1}. \tag{9.80}$$

Here, it is assumed that $n_f < n_i$. Note that the emission spectrum of hydrogen is *quantized*: *i.e.*, a hydrogen atom can only emit photons with certain fixed set of wavelengths. Likewise, a hydrogen atom can only absorb photons which have the same fixed set of wavelengths. This set of wavelengths constitutes the characteristic emission/absorption spectrum of the hydrogen atom, and can be observed as "spectral lines" using a spectroscope.

Equation (9.79) is known as the *Rydberg formula*. Likewise, R is called the *Rydberg constant*. The Rydberg formula was actually discovered empirically in the nineteenth century by spectroscopists, and was first explained theoretically by Bohr in 1913 using a primitive version of quantum mechanics. Transitions to the ground-state ($n_f = 1$) give rise to spectral lines in the ultraviolet band—this set of lines is called the *Lyman series*. Transitions to the first excited state ($n_f = 2$) give rise to spectral lines in the visible band—this set of lines is called the *Balmer series*. Transitions to the second excited state ($n_f = 3$) give rise to spectral lines in the infrared band—this set of lines is called the *Paschen series*, and so on.

Exercises

1. A particle of mass m is placed in a *finite* spherical well:
$$V(r) = \begin{cases} -V_0 & \text{for } r \leq a \\ 0 & \text{for } r > a \end{cases},$$
with $V_0 > 0$ and $a > 0$. Find the ground-state by solving the radial equation with $l = 0$. Show that there is no ground-state if $V_0\,a^2 < \pi^2\,\hbar^2/8\,m$.

2. Consider a particle of mass m in the three-dimensional harmonic oscillator potential $V(r) = (1/2)\,m\,\omega^2\,r^2$. Solve the problem by separation of variables in spherical polar coordinates, and, hence, determine the energy eigenvalues of the system.

3. The normalized wavefunction for the ground-state of a hydrogen-like atom (neutral hydrogen, He$^+$, Li^{++}, *etc.*) with nuclear charge $Z\,e$ has the form
$$\psi = A\,\exp(-\beta\,r),$$
where A and β are constants, and r is the distance between the nucleus and the electron. Show the following:
    (a) $A^2 = \beta^3/\pi$.
    (b) $\beta = Z/a_0$, where $a_0 = (\hbar^2/m_e)\,(4\pi\,\epsilon_0/e^2)$.
    (c) The energy is $E = -Z^2\,E_0$ where $E_0 = (m_e/2\,\hbar^2)\,(e^2/4\pi\,\epsilon_0)^2$.
    (d) The expectation values of the potential and kinetic energies are $2\,E$ and $-E$, respectively.
    (e) The expectation value of r is $(3/2)\,(a_0/Z)$.
    (f) The most probable value of r is a_0/Z.

4. An atom of tritium is in its ground-state. Suddenly the nucleus decays into a helium nucleus, via the emission of a fast electron which leaves the atom without perturbing the extranuclear electron, Find the probability that the resulting He$^+$ ion will be left in an $n = 1$, $l = 0$ state. Find the probability that it will be left in a $n = 2$, $l = 0$ state. What is the probability that the ion will be left in an $l > 0$ state?

5. Calculate the wavelengths of the photons emitted from the $n = 2$, $l = 1$ to $n = 1$, $l = 0$ transition in hydrogen, deuterium, and positronium.

6. To conserve linear momentum, an atom emitting a photon must recoil, which means that not all of the energy made available in the downward jump goes to the photon. Find a hydrogen atom's recoil energy when it emits a photon in an $n = 2$ to $n = 1$ transition. What fraction of the transition energy is the recoil energy?

10 Spin Angular Momentum

10.1 Introduction

Broadly speaking, a classical extended object (*e.g.*, the Earth) can possess two types of angular momentum. The first type is due to the rotation of the object's center of mass about some fixed external point (*e.g.*, the Sun)—this is generally known as *orbital angular momentum*. The second type is due to the object's internal motion—this is generally known as *spin angular momentum* (since, for a rigid object, the internal motion consists of spinning about an axis passing through the center of mass). By analogy, quantum particles can possess both orbital angular momentum due to their motion through space (see Cha. 8), and spin angular momentum due to their internal motion. Actually, the analogy with classical extended objects is not entirely accurate, since electrons, for instance, are structureless point particles. In fact, in quantum mechanics, it is best to think of spin angular momentum as a kind of *intrinsic* angular momentum possessed by particles. It turns out that each type of elementary particle has a characteristic spin angular momentum, just as each type has a characteristic charge and mass.

10.2 Spin Operators

Since spin is a type of angular momentum, it is reasonable to suppose that it possesses similar properties to orbital angular momentum. Thus, by analogy with Sect. 8.2, we would expect to be able to define three operators—S_x, S_y, and S_z—which represent the three Cartesian components of spin angular momentum. Moreover, it is plausible that these operators possess analogous commutation relations to the three corresponding orbital angular momentum operators, L_x, L_y, and L_z [see Eqs. (8.6)–(8.8)]. In other words,

$$[S_x, S_y] = i\,\hbar\,S_z, \tag{10.1}$$

$$[S_y, S_z] = i\,\hbar\,S_x, \tag{10.2}$$

$$[S_z, S_x] = i\,\hbar\,S_y. \tag{10.3}$$

We can represent the magnitude squared of the spin angular momentum vector by the operator

$$S^2 = S_x^{\,2} + S_y^{\,2} + S_z^{\,2}. \tag{10.4}$$

By analogy with the analysis in Sect. 8.2, it is easily demonstrated that

$$[S^2, S_x] = [S^2, S_y] = [S^2, S_z] = 0. \tag{10.5}$$

We thus conclude (see Sect. 4.10) that we can simultaneously measure the magnitude squared of the spin angular momentum vector, together with, at most, one Cartesian component. By convention, we shall always choose to measure the z-component, S_z.

By analogy with Eq. (8.13), we can define raising and lowering operators for spin angular momentum:

$$S_\pm = S_x \pm i\, S_y. \tag{10.6}$$

If S_x, S_y, and S_z are Hermitian operators, as must be the case if they are to represent physical quantities, then $S_\pm$ are the Hermitian conjugates of one another: *i.e.*,

$$(S_\pm)^\dagger = S_\mp. \tag{10.7}$$

Finally, by analogy with Sect. 8.2, it is easily demonstrated that

$$\begin{aligned} S_+ S_- &= S^2 - S_z^2 + \hbar\, S_z, &(10.8)\\ S_- S_+ &= S^2 - S_z^2 - \hbar\, S_z, &(10.9)\\ [S_+, S_z] &= -\hbar\, S_+, &(10.10)\\ [S_-, S_z] &= +\hbar\, S_-. &(10.11) \end{aligned}$$

10.3 Spin Space

We now have to discuss the wavefunctions upon which the previously introduced spin operators act. Unlike regular wavefunctions, spin wavefunctions do not exist in real space. Likewise, the spin angular momentum operators cannot be represented as differential operators in real space. Instead, we need to think of spin wavefunctions as existing in an abstract (complex) vector space. The different members of this space correspond to the different internal configurations of the particle under investigation. Note that only the *directions* of our vectors have any physical significance (just as only the shape of a regular wavefunction has any physical significance). Thus, if the vector χ corresponds to a particular internal state then $c\,\chi$ corresponds to the same state, where c is a complex number. Now, we expect the internal states of our particle to be *superposable*, since the superposability of states is one of the fundamental assumptions of quantum mechanics. It follows that the vectors making up our vector space must also be superposable. Thus, if χ_1 and χ_2 are two vectors corresponding to two different internal states then $c_1\,\chi_1 + c_2\,\chi_2$ is another vector corresponding to the state obtained by superposing c_1 times state 1 with c_2 times state 2 (where c_1 and c_2 are complex numbers). Finally, the dimensionality of our vector space is simply the number of linearly independent vectors required to span it (*i.e.*, the number of linearly independent internal states of the particle under investigation).

We now need to define the length of our vectors. We can do this by introducing a second, or *dual*, vector space whose elements are in one to one correspondence with the elements of our first space. Let the element of the second space which corresponds to the element χ of the first space be called $\chi^\dagger$. Moreover, the element of the second space which corresponds to $c\,\chi$ is $c^\ast\,\chi^\dagger$. We shall assume that it is possible to combine χ and $\chi^\dagger$ in a multiplicative fashion to generate a real positive-definite number which we interpret as the length, or *norm*, of χ. Let us denote this number $\chi^\dagger\,\chi$. Thus, we have

$$\chi^\dagger\,\chi \geq 0 \tag{10.12}$$

for all χ. We shall also assume that it is possible to combine unlike states in an analogous multiplicative fashion to produce complex numbers. The product of two unlike states χ and χ' is denoted $\chi^\dagger\,\chi'$. Two states χ and χ' are said to be mutually orthogonal, or independent, if $\chi^\dagger\,\chi' = 0$.

Now, when a general spin operator, A, operates on a general spin-state, χ, it coverts it into a different spin-state which we shall denote $A\,\chi$. The dual of this state is $(A\,\chi)^\dagger \equiv \chi^\dagger\,A^\dagger$, where $A^\dagger$ is the Hermitian conjugate of A (this is the definition of an Hermitian conjugate in spin space). An eigenstate of A corresponding to the eigenvalue a satisfies

$$A\,\chi_a = a\,\chi_a. \tag{10.13}$$

As before, if A corresponds to a physical variable then a measurement of A will result in one of its eigenvalues (see Sect. 4.10). In order to ensure that these eigenvalues are all real, A must be Hermitian: *i.e.*, $A^\dagger = A$ (see Sect. 4.9). We expect the χ_a to be mutually orthogonal. We can also normalize them such that they all have unit length. In other words,

$$\chi_a^\dagger\,\chi_{a'} = \delta_{aa'}. \tag{10.14}$$

Finally, a general spin state can be written as a superposition of the normalized eigenstates of A: *i.e.*,

$$\chi = \sum_a c_a\,\chi_a. \tag{10.15}$$

A measurement of χ will then yield the result a with probability $|c_a|^2$.

10.4 Eigenstates of S_z and S^2

Since the operators S_z and S^2 commute, they must possess simultaneous eigenstates (see Sect. 4.10). Let these eigenstates take the form [see Eqs. (8.31) and (8.32)]:

$$S_z\,\chi_{s,m_s} = m_s\,\hbar\,\chi_{s,m_s}, \tag{10.16}$$

$$S^2\,\chi_{s,m_s} = s\,(s+1)\,\hbar^2\,\chi_{s,m_s}. \tag{10.17}$$

Now, it is easily demonstrated, from the commutation relations (10.10) and (10.11), that

$$S_z\,(S_+\,\chi_{s,m_s}) = (m_s+1)\,\hbar\,(S_+\,\chi_{s,m_s}), \tag{10.18}$$

and

$$S_z\,(S_-\,\chi_{s,m_s}) = (m_s-1)\,\hbar\,(S_-\,\chi_{s,m_s}). \tag{10.19}$$

Thus, S_+ and S_- are indeed the raising and lowering operators, respectively, for spin angular momentum (see Sect. 8.4). The eigenstates of S_z and S^2 are assumed to be orthonormal: *i.e.*,

$$\chi_{s,m_s}^\dagger\,\chi_{s',m_s'} = \delta_{ss'}\,\delta_{m_s\,m_s'}. \tag{10.20}$$

Consider the wavefunction $\chi = S_+ \chi_{s,m_s}$. Since we know, from Eq. (10.12), that $\chi^\dagger \chi \geq 0$, it follows that

$$(S_+ \chi_{s,m_s})^\dagger (S_+ \chi_{s,m_s}) = \chi_{s,m_s}^\dagger S_+^\dagger S_+ \chi_{s,m_s} = \chi_{s,m_s}^\dagger S_- S_+ \chi_{s,m_s} \geq 0, \tag{10.21}$$

where use has been made of Eq. (10.7). Equations (10.9), (10.16), (10.17), and (10.20) yield

$$s(s+1) \geq m_s(m_s+1). \tag{10.22}$$

Likewise, if $\chi = S_- \chi_{s,m_s}$ then we obtain

$$s(s+1) \geq m_s(m_s-1). \tag{10.23}$$

Assuming that $s \geq 0$, the above two inequalities imply that

$$-s \leq m_s \leq s. \tag{10.24}$$

Hence, at fixed s, there is both a maximum and a minimum possible value that m_s can take.

Let $m_{s\,min}$ be the minimum possible value of m_s. It follows that (see Sect. 8.6)

$$S_- \chi_{s,m_{s\,min}} = 0. \tag{10.25}$$

Now, from Eq. (10.8),

$$S^2 = S_+ S_- + S_z^2 - \hbar S_z. \tag{10.26}$$

Hence,

$$S^2 \chi_{s,m_{s\,min}} = (S_+ S_- + S_z^2 - \hbar S_z)\, \chi_{s,m_{s\,min}}, \tag{10.27}$$

giving

$$s(s+1) = m_{s\,min}(m_{s\,min} - 1). \tag{10.28}$$

Assuming that $m_{s\,min} < 0$, this equation yields

$$m_{s\,min} = -s. \tag{10.29}$$

Likewise, it is easily demonstrated that

$$m_{s\,max} = +s. \tag{10.30}$$

Moreover,

$$S_- \chi_{s,-s} = S_+ \chi_{s,s} = 0. \tag{10.31}$$

Now, the raising operator S_+, acting upon $\chi_{s,-s}$, converts it into some multiple of $\chi_{s,-s+1}$. Employing the raising operator a second time, we obtain a multiple of $\chi_{s,-s+2}$. However, this process cannot continue indefinitely, since there is a maximum possible value of m_s. Indeed, after acting upon $\chi_{s,-s}$ a sufficient number of times with the raising operator S_+, we must obtain a multiple of $\chi_{s,s}$, so that employing the raising operator one more time leads to the null state [see Eq. (10.31)]. If this is not the case then we will inevitably obtain eigenstates of S_z corresponding to $m_s > s$, which we have already demonstrated is impossible.

It follows, from the above argument, that

$$m_{s\,max} - m_{s\,min} = 2\,s = k, \tag{10.32}$$

where k is a positive integer. Hence, the quantum number s can either take *positive integer* or *positive half-integer* values. Up to now, our analysis has been very similar to that which we used earlier to investigate orbital angular momentum (see Sect. 8). Recall, that for orbital angular momentum the quantum number m, which is analogous to m_s, is restricted to take *integer* values (see Cha. 8.5). This implies that the quantum number l, which is analogous to s, is also restricted to take integer values. However, the origin of these restrictions is the representation of the orbital angular momentum operators as differential operators in real space (see Sect. 8.3). There is no equivalent representation of the corresponding spin angular momentum operators. Hence, we conclude that there is no reason why the quantum number s cannot take half-integer, as well as integer, values.

In 1940, Wolfgang Pauli proved the so-called *spin-statistics theorem* using relativistic quantum mechanics. According to this theorem, all *fermions* possess *half-integer spin* (*i.e.*, a half-integer value of s), whereas all *bosons* possess *integer spin* (*i.e.*, an integer value of s). In fact, all presently known fermions, including electrons and protons, possess *spin one-half*. In other words, electrons and protons are characterized by $s = 1/2$ and $m_s = \pm 1/2$.

10.5 Pauli Representation

Let us denote the two independent spin eigenstates of an electron as

$$\chi_\pm \equiv \chi_{1/2,\pm 1/2}. \tag{10.33}$$

It thus follows, from Eqs. (10.16) and (10.17), that

$$\begin{aligned} S_z\,\chi_\pm &= \pm\frac{1}{2}\,\hbar\,\chi_\pm, & (10.34)\\ S^2\,\chi_\pm &= \frac{3}{4}\,\hbar^2\,\chi_\pm. & (10.35) \end{aligned}$$

Note that χ_+ corresponds to an electron whose spin angular momentum vector has a positive component along the z-axis. Loosely speaking, we could say that the spin vector points in the $+z$-direction (or its spin is "up"). Likewise, χ_- corresponds to an electron whose spin points in the $-z$-direction (or whose spin is "down"). These two eigenstates satisfy the orthonormality requirements

$$\chi_+^\dagger\,\chi_+ = \chi_-^\dagger\,\chi_- = 1, \tag{10.36}$$

and

$$\chi_+^\dagger\,\chi_- = 0. \tag{10.37}$$

A general spin state can be represented as a linear combination of χ_+ and χ_-: *i.e.*,

$$\chi = c_+\,\chi_+ + c_-\,\chi_-. \tag{10.38}$$

It is thus evident that electron spin space is *two-dimensional*.

Up to now, we have discussed spin space in rather abstract terms. In the following, we shall describe a particular representation of electron spin space due to Pauli. This so-called *Pauli representation* allows us to visualize spin space, and also facilitates calculations involving spin.

Let us attempt to represent a general spin state as a complex *column vector* in some two-dimensional space: *i.e.*,

$$\chi \equiv \left(\begin{array}{c} c_+ \\ c_- \end{array}\right). \tag{10.39}$$

The corresponding dual vector is represented as a *row vector*: *i.e.*,

$$\chi^\dagger \equiv (c_+^\ast, c_-^\ast). \tag{10.40}$$

Furthermore, the product $\chi^\dagger\,\chi$ is obtained according to the ordinary rules of matrix multiplication: *i.e.*,

$$\chi^\dagger\,\chi = (c_+^\ast, c_-^\ast)\left(\begin{array}{c} c_+ \\ c_- \end{array}\right) = c_+^\ast\,c_+ + c_-^\ast\,c_- = |c_+|^2 + |c_-|^2 \geq 0. \tag{10.41}$$

Likewise, the product $\chi^\dagger\,\chi'$ of two different spin states is also obtained from the rules of matrix multiplication: *i.e.*,

$$\chi^\dagger\,\chi' = (c_+^\ast, c_-^\ast)\left(\begin{array}{c} c_+' \\ c_-' \end{array}\right) = c_+^\ast\,c_+' + c_-^\ast\,c_-'. \tag{10.42}$$

Note that this particular representation of spin space is in complete accordance with the discussion in Sect. 10.3. For obvious reasons, a vector used to represent a spin state is generally known as *spinor*.

A general spin operator A is represented as a 2×2 matrix which operates on a spinor: *i.e.*,

$$A\,\chi \equiv \left(\begin{array}{cc} A_{11}, & A_{12} \\ A_{21}, & A_{22} \end{array}\right)\left(\begin{array}{c} c_+ \\ c_- \end{array}\right). \tag{10.43}$$

As is easily demonstrated, the Hermitian conjugate of A is represented by the transposed complex conjugate of the matrix used to represent A: *i.e.*,

$$A^\dagger \equiv \left(\begin{array}{cc} A_{11}^\ast, & A_{21}^\ast \\ A_{12}^\ast, & A_{22}^\ast \end{array}\right). \tag{10.44}$$

Let us represent the spin eigenstates χ_+ and χ_- as

$$\chi_+ \equiv \left(\begin{array}{c} 1 \\ 0 \end{array}\right), \tag{10.45}$$

and

$$\chi_- \equiv \left(\begin{array}{c} 0 \\ 1 \end{array}\right), \tag{10.46}$$

respectively. Note that these forms automatically satisfy the orthonormality constraints (10.36) and (10.37). It is convenient to write the spin operators S_i (where $i = 1, 2, 3$ corresponds to x, y, z) as

$$S_i = \frac{\hbar}{2}\,\sigma_i. \tag{10.47}$$

Here, the σ_i are dimensionless 2×2 matrices. According to Eqs. (10.1)–(10.3), the σ_i satisfy the commutation relations

$$[\sigma_x, \sigma_y] = 2\,i\,\sigma_z, \tag{10.48}$$

$$[\sigma_y, \sigma_z] = 2\,i\,\sigma_x, \tag{10.49}$$

$$[\sigma_z, \sigma_x] = 2\,i\,\sigma_y. \tag{10.50}$$

Furthermore, Eq. (10.34) yields

$$\sigma_z\,\chi_\pm = \pm\chi_\pm. \tag{10.51}$$

It is easily demonstrated, from the above expressions, that the σ_i are represented by the following matrices:

$$\sigma_x \equiv \left(\begin{array}{cc} 0, & 1 \\ 1, & 0 \end{array}\right), \tag{10.52}$$

$$\sigma_y \equiv \left(\begin{array}{cc} 0, & -i \\ i, & 0 \end{array}\right), \tag{10.53}$$

$$\sigma_z \equiv \left(\begin{array}{cc} 1, & 0 \\ 0, & -1 \end{array}\right). \tag{10.54}$$

Incidentally, these matrices are generally known as the *Pauli matrices*.

Finally, a general spinor takes the form

$$\chi = c_+\chi_+ + c_-\chi_- = \left(\begin{array}{c} c_+ \\ c_- \end{array}\right). \tag{10.55}$$

If the spinor is properly normalized then

$$\chi^\dagger\,\chi = |c_+|^2 + |c_-|^2 = 1. \tag{10.56}$$

In this case, we can interpret $|c_+|^2$ as the probability that an observation of S_z will yield the result $+\hbar/2$, and $|c_-|^2$ as the probability that an observation of S_z will yield the result $-\hbar/2$.

10.6 Spin Precession

According to classical physics, a small current loop possesses a *magnetic moment* of magnitude $\mu = I\,A$, where I is the current circulating around the loop, and A the area of the loop. The direction of the magnetic moment is conventionally taken to be normal to the plane of the loop, in the sense given by a standard right-hand circulation rule. Consider a small current loop consisting of an electron in uniform circular motion. It is easily demonstrated that the electron's orbital angular momentum $\mathbf{L}$ is related to the magnetic moment $\boldsymbol{\mu}$ of the loop via

$$\boldsymbol{\mu} = -\frac{e}{2\,m_e}\,\mathbf{L}, \tag{10.57}$$

where e is the magnitude of the electron charge, and m_e the electron mass.

The above expression suggests that there may be a similar relationship between magnetic moment and spin angular momentum. We can write

$$\boldsymbol{\mu} = -\frac{g\,e}{2\,m_e}\,\mathbf{S}, \tag{10.58}$$

where g is called the *gyromagnetic ratio*. Classically, we would expect $g = 1$. In fact,

$$g = 2\left(1 + \frac{\alpha}{2\pi} + \cdots\right) = 2.0023192, \tag{10.59}$$

where $\alpha = e^2/(2\,\epsilon_0\,h\,c) \simeq 1/137$ is the so-called *fine-structure constant*. The fact that the gyromagnetic ratio is (almost) twice that expected from classical physics is only explicable using relativistic quantum mechanics. Furthermore, the small corrections to the relativistic result $g = 2$ come from quantum field theory.

The energy of a classical magnetic moment $\boldsymbol{\mu}$ in a uniform magnetic field $\mathbf{B}$ is

$$H = -\boldsymbol{\mu}\cdot\mathbf{B}. \tag{10.60}$$

Assuming that the above expression also holds good in quantum mechanics, the Hamiltonian of an electron in a z-directed magnetic field of magnitude B takes the form

$$H = \Omega\,S_z, \tag{10.61}$$

where

$$\Omega = \frac{g\,e\,B}{2\,m_e}. \tag{10.62}$$

Here, for the sake of simplicity, we are neglecting the electron's translational degrees of freedom.

Schrödinger's equation can be written [see Eq. (4.63)]

$$i\,\hbar\,\frac{\partial\chi}{\partial t} = H\,\chi, \tag{10.63}$$

where the spin state of the electron is characterized by the spinor χ. Adopting the Pauli representation, we obtain

$$\chi = \begin{pmatrix} c_+(t) \\ c_-(t) \end{pmatrix}, \tag{10.64}$$

where $|c_+|^2 + |c_-|^2 = 1$. Here, $|c_+|^2$ is the probability of observing the spin-up state, and $|c_-|^2$ the probability of observing the spin-down state. It follows from Eqs. (10.47), (10.54), (10.61), (10.63), and (10.64) that

$$i\,\hbar \begin{pmatrix} \dot{c}_+ \\ \dot{c}_- \end{pmatrix} = \frac{\Omega\,\hbar}{2} \begin{pmatrix} 1, & 0 \\ 0, & -1 \end{pmatrix} \begin{pmatrix} c_+ \\ c_- \end{pmatrix}, \tag{10.65}$$

where $\dot{} \equiv d/dt$. Hence,

$$\dot{c}_\pm = \mp i\,\frac{\Omega}{2}\,c_\pm. \tag{10.66}$$

Let

$$c_+(0) = \cos(\alpha/2), \tag{10.67}$$

$$c_-(0) = \sin(\alpha/2). \tag{10.68}$$

The significance of the angle α will become apparent presently. Solving Eq. (10.66), subject to the initial conditions (10.67) and (10.68), we obtain

$$c_+(t) = \cos(\alpha/2)\,\exp(-i\,\Omega\,t/2), \tag{10.69}$$

$$c_-(t) = \sin(\alpha/2)\,\exp(+i\,\Omega\,t/2). \tag{10.70}$$

We can most easily visualize the effect of the time dependence in the above expressions for $c_\pm$ by calculating the expectation values of the three Cartesian components of the electron's spin angular momentum. By analogy with Eq. (4.56), the expectation value of a general spin operator A is simply

$$\langle A\rangle = \chi^\dagger\,A\,\chi. \tag{10.71}$$

Hence, the expectation value of S_z is

$$\langle S_z\rangle = \frac{\hbar}{2}\,(c_+^\ast, c_-^\ast) \begin{pmatrix} 1, & 0 \\ 0, & -1 \end{pmatrix} \begin{pmatrix} c_+ \\ c_- \end{pmatrix}, \tag{10.72}$$

which reduces to

$$\langle S_z\rangle = \frac{\hbar}{2}\,\cos\alpha \tag{10.73}$$

with the help of Eqs. (10.69) and (10.70). Likewise, the expectation value of S_x is

$$\langle S_x\rangle = \frac{\hbar}{2}\,(c_+^\ast, c_-^\ast) \begin{pmatrix} 0, & 1 \\ 1, & 0 \end{pmatrix} \begin{pmatrix} c_+ \\ c_- \end{pmatrix}, \tag{10.74}$$

which reduces to

$$\langle S_x\rangle = \frac{\hbar}{2}\,\sin\alpha\,\cos(\Omega\,t). \tag{10.75}$$

Finally, the expectation value of S_y is

$$\langle S_y\rangle = \frac{\hbar}{2}\,\sin\alpha\,\sin(\Omega\,t). \tag{10.76}$$

According to Eqs. (10.73), (10.75), and (10.76), the *expectation value* of the spin angular momentum vector subtends a constant angle α with the z-axis, and *precesses* about this axis at the frequency

$$\Omega \simeq \frac{e\,B}{m_e}. \tag{10.77}$$

This behaviour is actually equivalent to that predicted by classical physics. Note, however, that a *measurement* of S_x, S_y, or S_z will always yield either $+\hbar/2$ or $-\hbar/2$. It is the *relative probabilities* of obtaining these two results which varies as the expectation value of a given component of the spin varies.

Exercises

1. Find the Pauli representations of S_x, S_y, and S_z for a spin-1 particle.
2. Find the Pauli representations of the normalized eigenstates of S_x and S_y for a spin-1/2 particle.
3. Suppose that a spin-1/2 particle has a spin vector which lies in the x-z plane, making an angle θ with the z-axis. Demonstrate that a measurement of S_z yields $\hbar/2$ with probability $\cos^2(\theta/2)$, and $-\hbar/2$ with probability $\sin^2(\theta/2)$.
4. An electron is in the spin-state

$$\chi = A\begin{pmatrix} 1-2\,i \\ 2 \end{pmatrix}$$

in the Pauli representation. Determine the constant A by normalizing χ. If a measurement of S_z is made, what values will be obtained, and with what probabilities? What is the expectation value of S_z? Repeat the above calculations for S_x and S_y.

5. Consider a spin-1/2 system represented by the normalized spinor

$$\chi = \begin{pmatrix} \cos\alpha \\ \sin\alpha\,\exp(\,i\,\beta) \end{pmatrix}$$

in the Pauli representation, where α and β are real. What is the probability that a measurement of S_y yields $-\hbar/2$?

6. An electron is at rest in an oscillating magnetic field

$$\mathbf{B} = B_0\,\cos(\omega\,t)\,\mathbf{e}_z,$$

where B_0 and ω are real positive constants.

(a) Find the Hamiltonian of the system.

(b) If the electron starts in the spin-up state with respect to the x-axis, determine the spinor $\chi(t)$ which represents the state of the system in the Pauli representation at all subsequent times.

(c) Find the probability that a measurement of S_x yields the result $-\hbar/2$ as a function of time.

(d) What is the minimum value of B_0 required to force a complete flip in S_x?

11 Addition of Angular Momentum

11.1 Introduction

Consider an electron in a hydrogen atom. As we have already seen, the electron's motion through space is parameterized by the three quantum numbers n, l, and m (see Sect. 9.4). To these we must now add the two quantum numbers s and m_s which parameterize the electron's internal motion (see the previous chapter). Now, the quantum numbers l and m specify the electron's orbital angular momentum vector, **L**, (as much as it can be specified) whereas the quantum numbers s and m_s specify its spin angular momentum vector, **S**. But, if the electron possesses both orbital and spin angular momentum then what is its total angular momentum?

11.2 General Principles

The three basic orbital angular momentum operators, L_x, L_y, and L_z, obey the commutation relations (8.6)–(8.8), which can be written in the convenient vector form:

$$\mathbf{L}\times\mathbf{L} = i\,\hbar\,\mathbf{L}. \tag{11.1}$$

Likewise, the three basic spin angular momentum operators, S_x, S_y, and S_z, obey the commutation relations (10.1)–(10.3), which can also be written in vector form: *i.e.*,

$$\mathbf{S}\times\mathbf{S} = i\,\hbar\,\mathbf{S}. \tag{11.2}$$

Now, since the orbital angular momentum operators are associated with the electron's motion through space, whilst the spin angular momentum operators are associated with its internal motion, and these two types of motion are completely unrelated (*i.e.*, they correspond to different degrees of freedom—see Sect. 6.2), it is reasonable to suppose that the two sets of operators commute with one another: *i.e.*,

$$[L_i, S_j] = 0, \tag{11.3}$$

where $i, j = 1, 2, 3$ corresponds to x, y, z.

Let us now consider the electron's total angular momentum vector

$$\mathbf{J} = \mathbf{L} + \mathbf{S}. \tag{11.4}$$

We have

$$\begin{aligned}\mathbf{J}\times\mathbf{J} &= (\mathbf{L}+\mathbf{S})\times(\mathbf{L}+\mathbf{S})\\ &= \mathbf{L}\times\mathbf{L}+\mathbf{S}\times\mathbf{S}+\mathbf{L}\times\mathbf{S}+\mathbf{S}\times\mathbf{L} = \mathbf{L}\times\mathbf{L}+\mathbf{S}\times\mathbf{S}\\ &= i\,\hbar\,\mathbf{L}+i\,\hbar\,\mathbf{S}\\ &= i\,\hbar\,\mathbf{J}.\end{aligned} \tag{11.5}$$

In other words,

$$\mathbf{J}\times\mathbf{J} = i\,\hbar\,\mathbf{J}. \tag{11.6}$$

It is thus evident that the three basic total angular momentum operators, J_x, J_y, and J_z, obey analogous commutation relations to the corresponding orbital and spin angular momentum operators. It therefore follows that the total angular momentum has similar properties to the orbital and spin angular momenta. For instance, it is only possible to simultaneously measure the magnitude squared of the total angular momentum vector,

$$J^2 = J_x^{\,2} + J_y^{\,2} + J_z^{\,2}, \tag{11.7}$$

together with a single Cartesian component. By convention, we shall always choose to measure J_z. A simultaneous eigenstate of J_z and J^2 satisfies

$$\begin{aligned} J_z\,\psi_{j,m_j} &= m_j\,\hbar\,\psi_{j,m_j}, && (11.8)\\ J^2\,\psi_{j,m_j} &= j\,(j+1)\,\hbar^2\,\psi_{j,m_j}, && (11.9) \end{aligned}$$

where the quantum number j can take positive integer, or half-integer, values, and the quantum number m_j is restricted to the following range of values:

$$-j, -j+1, \cdots, j-1, j. \tag{11.10}$$

Now

$$J^2 = (\mathbf{L}+\mathbf{S})\cdot(\mathbf{L}+\mathbf{S}) = L^2 + S^2 + 2\,\mathbf{L}\cdot\mathbf{S}, \tag{11.11}$$

which can also be written as

$$J^2 = L^2 + S^2 + 2\,L_z\,S_z + L_+\,S_- + L_-\,S_+. \tag{11.12}$$

We know that the operator L^2 commutes with itself, with all of the Cartesian components of $\mathbf{L}$ (and, hence, with the raising and lowering operators $L_\pm$), and with all of the spin angular momentum operators (see Sect. 8.2). It is therefore clear that

$$[J^2, L^2] = 0. \tag{11.13}$$

A similar argument allows us to also conclude that

$$[J^2, S^2] = 0. \tag{11.14}$$

Now, the operator L_z commutes with itself, with L^2, with all of the spin angular momentum operators, but *not* with the raising and lowering operators $L_\pm$ (see Sect. 8.2). It follows that

$$[J^2, L_z] \neq 0. \tag{11.15}$$

Likewise, we can also show that

$$[J^2, S_z] \neq 0. \tag{11.16}$$

Finally, we have

$$J_z = L_z + S_z, \tag{11.17}$$

where $[J_z, L_z] = [J_z, S_z] = 0$.

Recalling that only *commuting* operators correspond to physical quantities which can be simultaneously measured (see Sect. 4.10), it follows, from the above discussion, that there are *two* alternative sets of physical variables associated with angular momentum which we can measure simultaneously. The first set correspond to the operators L^2, S^2, L_z, S_z, and J_z. The second set correspond to the operators L^2, S^2, J^2, and J_z. In other words, we can always measure the magnitude squared of the orbital and spin angular momentum vectors, together with the z-component of the total angular momentum vector. In addition, we can either choose to measure the z-components of the orbital and spin angular momentum vectors, or the magnitude squared of the total angular momentum vector.

Let $\psi^{(1)}_{l,s;m,m_s}$ represent a simultaneous eigenstate of L^2, S^2, L_z, and S_z corresponding to the following eigenvalues:

$$L^2\,\psi^{(1)}_{l,s;m,m_s} = l\,(l+1)\,\hbar^2\,\psi^{(1)}_{l,s;m,m_s}, \tag{11.18}$$

$$S^2\,\psi^{(1)}_{l,s;m,m_s} = s\,(s+1)\,\hbar^2\,\psi^{(1)}_{l,s;m,m_s}, \tag{11.19}$$

$$L_z\,\psi^{(1)}_{l,s;m,m_s} = m\,\hbar\,\psi^{(1)}_{l,s;m,m_s}, \tag{11.20}$$

$$S_z\,\psi^{(1)}_{l,s;m,m_s} = m_s\,\hbar\,\psi^{(1)}_{l,s;m,m_s}. \tag{11.21}$$

It is easily seen that

$$\begin{aligned} J_z\,\psi^{(1)}_{l,s;m,m_s} &= (L_z + S_z)\,\psi^{(1)}_{l,s;m,m_s} = (m + m_s)\,\hbar\,\psi^{(1)}_{l,s;m,m_s} \\ &= m_j\,\hbar\,\psi^{(1)}_{l,s;m,m_s}. \end{aligned} \tag{11.22}$$

Hence,

$$m_j = m + m_s. \tag{11.23}$$

In other words, the quantum numbers controlling the z-components of the various angular momentum vectors can simply be added algebraically.

Finally, let $\psi^{(2)}_{l,s;j,m_j}$ represent a simultaneous eigenstate of L^2, S^2, J^2, and J_z corresponding to the following eigenvalues:

$$L^2\,\psi^{(2)}_{l,s;j,m_j} = l\,(l+1)\,\hbar^2\,\psi^{(2)}_{l,s;j,m_j}, \tag{11.24}$$

$$S^2\,\psi^{(2)}_{l,s;j,m_j} = s\,(s+1)\,\hbar^2\,\psi^{(2)}_{l,s;j,m_j}, \tag{11.25}$$

$$J^2\,\psi^{(2)}_{l,s;j,m_j} = j\,(j+1)\,\hbar^2\,\psi^{(2)}_{l,s;j,m_j}, \tag{11.26}$$

$$J_z\,\psi^{(2)}_{l,s;j,m_j} = m_j\,\hbar\,\psi^{(2)}_{l,s;j,m_j}. \tag{11.27}$$

11.3 Angular Momentum in the Hydrogen Atom

In a hydrogen atom, the wavefunction of an electron in a simultaneous eigenstate of L^2 and L_z has an angular dependence specified by the spherical harmonic $Y_{l,m}(\theta,\phi)$ (see Sect. 8.7). If the electron is also in an eigenstate of S^2 and S_z then the quantum numbers s and m_s take the values $1/2$ and $\pm 1/2$, respectively, and the internal state of the electron is specified by the spinors $\chi_\pm$ (see Sect. 10.5). Hence, the simultaneous eigenstates of L^2, S^2, L_z, and S_z can be written in the separable form

$$\psi^{(1)}_{l,1/2;m,\pm 1/2} = Y_{l,m}\,\chi_\pm. \tag{11.28}$$

Here, it is understood that orbital angular momentum operators act on the spherical harmonic functions, $Y_{l,m}$, whereas spin angular momentum operators act on the spinors, $\chi_\pm$.

Since the eigenstates $\psi^{(1)}_{l,1/2;m,\pm 1/2}$ are (presumably) orthonormal, and form a complete set, we can express the eigenstates $\psi^{(2)}_{l,1/2;j,m_j}$ as linear combinations of them. For instance,

$$\psi^{(2)}_{l,1/2;j,m+1/2} = \alpha\,\psi^{(1)}_{l,1/2;m,1/2} + \beta\,\psi^{(1)}_{l,1/2;m+1,-1/2}, \tag{11.29}$$

where α and β are, as yet, unknown coefficients. Note that the number of $\psi^{(1)}$ states which can appear on the right-hand side of the above expression is limited to two by the constraint that $m_j = m + m_s$ [see Eq. (11.23)], and the fact that m_s can only take the values $\pm 1/2$. Assuming that the $\psi^{(2)}$ eigenstates are properly normalized, we have

$$\alpha^2 + \beta^2 = 1. \tag{11.30}$$

Now, it follows from Eq. (11.26) that

$$J^2\,\psi^{(2)}_{l,1/2;j,m+1/2} = j\,(j+1)\,\hbar^2\,\psi^{(2)}_{l,1/2;j,m+1/2}, \tag{11.31}$$

where [see Eq. (11.12)]

$$J^2 = L^2 + S^2 + 2\,L_z\,S_z + L_+\,S_- + L_-\,S_+. \tag{11.32}$$

Moreover, according to Eqs. (11.28) and (11.29), we can write

$$\psi^{(2)}_{l,1/2;j,m+1/2} = \alpha\,Y_{l,m}\,\chi_+ + \beta\,Y_{l,m+1}\,\chi_-. \tag{11.33}$$

Recall, from Eqs. (8.43) and (8.44), that

$$L_+\,Y_{l,m} = [l\,(l+1) - m\,(m+1)]^{1/2}\,\hbar\,Y_{l,m+1}, \tag{11.34}$$

$$L_-\,Y_{l,m} = [l\,(l+1) - m\,(m-1)]^{1/2}\,\hbar\,Y_{l,m-1}. \tag{11.35}$$

By analogy, when the spin raising and lowering operators, $S_\pm$, act on a general spinor, χ_{s,m_s}, we obtain

$$S_+\,\chi_{s,m_s} = [s\,(s+1) - m_s\,(m_s+1)]^{1/2}\,\hbar\,\chi_{s,m_s+1}, \tag{11.36}$$

$$S_-\,\chi_{s,m_s} = [s\,(s+1) - m_s\,(m_s-1)]^{1/2}\,\hbar\,\chi_{s,m_s-1}. \tag{11.37}$$

For the special case of spin one-half spinors (*i.e.*, $s = 1/2, m_s = \pm 1/2$), the above expressions reduce to

$$S_+\,\chi_+ = S_-\,\chi_- = 0, \tag{11.38}$$

and

$$S_\pm\,\chi_\mp = \hbar\,\chi_\pm. \tag{11.39}$$

It follows from Eqs. (11.32) and (11.34)–(11.39) that

$$\begin{aligned} J^2\,Y_{l,m}\,\chi_+ &= [l\,(l+1) + 3/4 + m]\,\hbar^2\,Y_{l,m}\,\chi_+ \\ &\quad +[l\,(l+1) - m\,(m+1)]^{1/2}\,\hbar^2\,Y_{l,m+1}\,\chi_-, \end{aligned} \tag{11.40}$$

and

$$\begin{aligned} J^2\,Y_{l,m+1}\,\chi_- &= [l\,(l+1) + 3/4 - m - 1]\,\hbar^2\,Y_{l,m+1}\,\chi_- \\ &\quad +[l\,(l+1) - m\,(m+1)]^{1/2}\,\hbar^2\,Y_{l,m}\,\chi_+. \end{aligned} \tag{11.41}$$

Hence, Eqs. (11.31) and (11.33) yield

$$(x-m)\,\alpha - [l\,(l+1) - m\,(m+1)]^{1/2}\,\beta = 0, \tag{11.42}$$

$$-[l\,(l+1) - m\,(m+1)]^{1/2}\,\alpha + (x+m+1)\,\beta = 0, \tag{11.43}$$

where

$$x = j\,(j+1) - l\,(l+1) - 3/4. \tag{11.44}$$

Equations (11.42) and (11.43) can be solved to give

$$x\,(x+1) = l\,(l+1), \tag{11.45}$$

and

$$\frac{\alpha}{\beta} = \frac{[(l-m)\,(l+m+1)]^{1/2}}{x-m}. \tag{11.46}$$

It follows that $x = l$ or $x = -l-1$, which corresponds to $j = l + 1/2$ or $j = l - 1/2$, respectively. Once x is specified, Eqs. (11.30) and (11.46) can be solved for α and β. We obtain

$$\psi^{(2)}_{l+1/2,m+1/2} = \left(\frac{l+m+1}{2\,l+1}\right)^{1/2}\psi^{(1)}_{m,1/2} + \left(\frac{l-m}{2\,l+1}\right)^{1/2}\psi^{(1)}_{m+1,-1/2}, \tag{11.47}$$

and

$$\psi^{(2)}_{l-1/2,m+1/2} = \left(\frac{l-m}{2\,l+1}\right)^{1/2}\psi^{(1)}_{m,1/2} - \left(\frac{l+m+1}{2\,l+1}\right)^{1/2}\psi^{(1)}_{m+1,-1/2}. \tag{11.48}$$

	$m, 1/2$	$m+1, -1/2$	m, m_s
$l+1/2, m+1/2$	$\sqrt{(l+m+1)/(2\,l+1)}$	$\sqrt{(l-m)/(2\,l+1)}$	
$l-1/2, m+1/2$	$\sqrt{(l-m)/(2\,l+1)}$	$-\sqrt{(l+m+1)/(2\,l+1)}$	
j, m_j			

Table 11.1: *Clebsch-Gordon coefficients for adding spin one-half to spin* l.

Here, we have neglected the common subscripts $l, 1/2$ for the sake of clarity: *i.e.*, $\psi^{(2)}_{l+1/2,m+1/2} \equiv \psi^{(2)}_{l,1/2;l+1/2,m+1/2}$, *etc.* The above equations can easily be inverted to give the $\psi^{(1)}$ eigenstates in terms of the $\psi^{(2)}$ eigenstates:

$$\psi^{(1)}_{m,1/2} = \left(\frac{l+m+1}{2\,l+1}\right)^{1/2} \psi^{(2)}_{l+1/2,m+1/2} + \left(\frac{l-m}{2\,l+1}\right)^{1/2} \psi^{(2)}_{l-1/2,m+1/2}, \tag{11.49}$$

$$\psi^{(1)}_{m+1,-1/2} = \left(\frac{l-m}{2\,l+1}\right)^{1/2} \psi^{(2)}_{l+1/2,m+1/2} - \left(\frac{l+m+1}{2\,l+1}\right)^{1/2} \psi^{(2)}_{l-1/2,m+1/2}. \tag{11.50}$$

The information contained in Eqs. (11.47)–(11.50) is neatly summarized in Table 11.1. For instance, Eq. (11.47) is obtained by reading the first row of this table, whereas Eq. (11.50) is obtained by reading the second column. The coefficients in this type of table are generally known as *Clebsch-Gordon coefficients*.

As an example, let us consider the $l = 1$ states of a hydrogen atom. The eigenstates of L^2, S^2, L_z, and S_z, are denoted $\psi^{(1)}_{m,m_s}$. Since m can take the values $-1, 0, 1$, whereas m_s can take the values $\pm 1/2$, there are clearly six such states: *i.e.*, $\psi^{(1)}_{1,\pm 1/2}$, $\psi^{(1)}_{0,\pm 1/2}$, and $\psi^{(1)}_{-1,\pm 1/2}$. The eigenstates of L^2, S^2, J^2, and J_z, are denoted $\psi^{(2)}_{j,m_j}$. Since $l = 1$ and $s = 1/2$ can be combined together to form either $j = 3/2$ or $j = 1/2$ (see earlier), there are also six such states: *i.e.*, $\psi^{(2)}_{3/2,\pm 3/2}$, $\psi^{(2)}_{3/2,\pm 1/2}$, and $\psi^{(2)}_{1/2,\pm 1/2}$. According to Table 11.1, the various different eigenstates are interrelated as follows:

$$\psi^{(2)}_{3/2,\pm 3/2} = \psi^{(1)}_{\pm 1,\pm 1/2}, \tag{11.51}$$

$$\psi^{(2)}_{3/2,1/2} = \sqrt{\frac{2}{3}}\,\psi^{(1)}_{0,1/2} + \sqrt{\frac{1}{3}}\,\psi^{(1)}_{1,-1/2}, \tag{11.52}$$

$$\psi^{(2)}_{1/2,1/2} = \sqrt{\frac{1}{3}}\,\psi^{(1)}_{0,1/2} - \sqrt{\frac{2}{3}}\,\psi^{(1)}_{1,-1/2}, \tag{11.53}$$

$$\psi^{(2)}_{1/2,-1/2} = \sqrt{\frac{2}{3}}\,\psi^{(1)}_{-1,1/2} - \sqrt{\frac{1}{3}}\,\psi^{(1)}_{0,-1/2}, \tag{11.54}$$

$$\psi^{(2)}_{3/2,-1/2} = \sqrt{\frac{1}{3}}\,\psi^{(1)}_{-1,1/2} + \sqrt{\frac{2}{3}}\,\psi^{(1)}_{0,-1/2}, \tag{11.55}$$

and

$$\psi^{(1)}_{\pm 1,\pm 1/2} = \psi^{(2)}_{3/2,\pm 3/2}, \tag{11.56}$$

	$-1,-1/2$	$-1,1/2$	$0,-1/2$	$0,1/2$	$1,-1/2$	$1,1/2$	m, m_s
$3/2,-3/2$	1						
$3/2,-1/2$		$\sqrt{1/3}$	$\sqrt{2/3}$				
$1/2,-1/2$		$\sqrt{2/3}$	$-\sqrt{1/3}$				
$3/2,1/2$				$\sqrt{2/3}$	$\sqrt{1/3}$		
$1/2,1/2$				$\sqrt{1/3}$	$-\sqrt{2/3}$		
$3/2,3/2$						1	
j, m_j							

Table 11.2: *Clebsch-Gordon coefficients for adding spin one-half to spin one. Only non-zero coefficients are shown.*

$$\psi^{(1)}_{1,-1/2} = \sqrt{\frac{1}{3}}\,\psi^{(2)}_{3/2,1/2} - \sqrt{\frac{2}{3}}\,\psi^{(2)}_{1/2,1/2}, \tag{11.57}$$

$$\psi^{(1)}_{0,1/2} = \sqrt{\frac{2}{3}}\,\psi^{(2)}_{3/2,1/2} + \sqrt{\frac{1}{3}}\,\psi^{(2)}_{1/2,1/2}, \tag{11.58}$$

$$\psi^{(1)}_{0,-1/2} = \sqrt{\frac{2}{3}}\,\psi^{(2)}_{3/2,-1/2} - \sqrt{\frac{1}{3}}\,\psi^{(2)}_{1/2,-1/2}, \tag{11.59}$$

$$\psi^{(1)}_{-1,1/2} = \sqrt{\frac{1}{3}}\,\psi^{(2)}_{3/2,-1/2} + \sqrt{\frac{2}{3}}\,\psi^{(2)}_{1/2,-1/2}, \tag{11.60}$$

Thus, if we know that an electron in a hydrogen atom is in an $l = 1$ state characterized by $m = 0$ and $m_s = 1/2$ [*i.e.*, the state represented by $\psi^{(1)}_{0,1/2}$] then, according to Eq. (11.58), a measurement of the total angular momentum will yield $j = 3/2$, $m_j = 1/2$ with probability 2/3, and $j = 1/2$, $m_j = 1/2$ with probability 1/3. Suppose that we make such a measurement, and obtain the result $j = 3/2$, $m_j = 1/2$. As a result of the measurement, the electron is thrown into the corresponding eigenstate, $\psi^{(2)}_{3/2,1/2}$. It thus follows from Eq. (11.52) that a subsequent measurement of L_z and S_z will yield $m = 0$, $m_s = 1/2$ with probability 2/3, and $m = 1$, $m_s = -1/2$ with probability 1/3.

The information contained in Eqs. (11.51)–(11.59) is neatly summed up in Table 11.2. Note that each row and column of this table has unit norm, and also that the different rows and different columns are mutually orthogonal. Of course, this is because the $\psi^{(1)}$ and $\psi^{(2)}$ eigenstates are orthonormal.

11.4 Two Spin One-Half Particles

Consider a system consisting of two spin one-half particles. Suppose that the system does not possess any orbital angular momentum. Let $\mathbf{S}_1$ and $\mathbf{S}_2$ be the spin angular momentum operators of the first and second particles, respectively, and let

$$\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2 \tag{11.61}$$

be the total spin angular momentum operator. By analogy with the previous analysis, we conclude that it is possible to simultaneously measure either S_1^2, S_2^2, S^2, and S_z, or S_1^2, S_2^2, S_{1z}, S_{2z}, and S_z. Let the quantum numbers associated with measurements of S_1^2, S_{1z}, S_2^2, S_{2z}, S^2, and S_z be s_1, m_{s_1}, s_2, m_{s_2}, s, and m_s, respectively. In other words, if the spinor $\chi^{(1)}_{s_1,s_2;m_{s_1},m_{s_2}}$ is a simultaneous eigenstate of S_1^2, S_2^2, S_{1z}, and S_{2z}, then

$$\begin{aligned}
S_1^2\,\chi^{(1)}_{s_1,s_2;m_{s_1},m_{s_2}} &= s_1\,(s_1+1)\,\hbar^2\,\chi^{(1)}_{s_1,s_2;m_{s_1},m_{s_2}}, && (11.62)\\
S_2^2\,\chi^{(1)}_{s_1,s_2;m_{s_1},m_{s_2}} &= s_2\,(s_2+1)\,\hbar^2\,\chi^{(1)}_{s_1,s_2;m_{s_1},m_{s_2}}, && (11.63)\\
S_{1z}\,\chi^{(1)}_{s_1,s_2;m_{s_1},m_{s_2}} &= m_{s_1}\,\hbar\,\chi^{(1)}_{s_1,s_2;m_{s_1},m_{s_2}}, && (11.64)\\
S_{2z}\,\chi^{(1)}_{s_1,s_2;m_{s_1},m_{s_2}} &= m_{s_2}\,\hbar\,\chi^{(1)}_{s_1,s_2;m_{s_1},m_{s_2}}, && (11.65)\\
S_z\,\chi^{(1)}_{s_1,s_2;m_{s_1},m_{s_2}} &= m_s\,\hbar\,\chi^{(1)}_{s_1,s_2;m_{s_1},m_{s_2}}. && (11.66)
\end{aligned}$$

Likewise, if the spinor $\chi^{(2)}_{s_1,s_2;s,m_s}$ is a simultaneous eigenstate of S_1^2, S_2^2, S^2, and S_z, then

$$\begin{aligned}
S_1^2\,\chi^{(2)}_{s_1,s_2;s,m_s} &= s_1\,(s_1+1)\,\hbar^2\,\chi^{(2)}_{s_1,s_2;s,m_s}, && (11.67)\\
S_2^2\,\chi^{(2)}_{s_1,s_2;s,m_s} &= s_2\,(s_2+1)\,\hbar^2\,\chi^{(2)}_{s_1,s_2;s,m_s}, && (11.68)\\
S^2\,\chi^{(2)}_{s_1,s_2;s,m_s} &= s\,(s+1)\,\hbar^2\,\chi^{(2)}_{s_1,s_2;s,m_s}, && (11.69)\\
S_z\,\chi^{(2)}_{s_1,s_2;s,m_s} &= m_s\,\hbar\,\chi^{(2)}_{s_1,s_2;s,m_s}. && (11.70)
\end{aligned}$$

Of course, since both particles have spin one-half, $s_1 = s_2 = 1/2$, and $s_{1z}, s_{2z} = \pm 1/2$. Furthermore, by analogy with previous analysis,

$$m_s = m_{s_1} + m_{s_2}. \tag{11.71}$$

Now, we saw, in the previous section, that when spin l is added to spin one-half then the possible values of the total angular momentum quantum number are $j = l \pm 1/2$. By analogy, when spin one-half is added to spin one-half then the possible values of the total spin quantum number are $s = 1/2 \pm 1/2$. In other words, when two spin one-half particles are combined, we either obtain a state with overall spin $s = 1$, or a state with overall spin $s = 0$. To be more exact, there are three possible $s = 1$ states (corresponding to $m_s = -1$, 0, 1), and one possible $s = 0$ state (corresponding to $m_s = 0$). The three $s = 1$ states are generally known as the *triplet* states, whereas the $s = 0$ state is known as the *singlet* state.

The Clebsch-Gordon coefficients for adding spin one-half to spin one-half can easily be inferred from Table 11.1 (with $l = 1/2$), and are listed in Table 11.3. It follows from this table that the three triplet states are:

$$\chi^{(2)}_{1,-1} = \chi^{(1)}_{-1/2,-1.2}, \tag{11.72}$$

	$-1/2,-1/2$	$-1/2,1/2$	$1/2,-1/2$	$1/2,1/2$	m_{s_1}, m_{s_2}
$1,-1$	1				
$1,0$		$1/\sqrt{2}$	$1/\sqrt{2}$		
$0,0$		$1/\sqrt{2}$	$-1/\sqrt{2}$		
$1,1$				1	
s, m_s					

Table 11.3: *Clebsch-Gordon coefficients for adding spin one-half to spin one-half. Only non-zero coefficients are shown.*

$$\chi^{(2)}_{1,0} = \frac{1}{\sqrt{2}}\left(\chi^{(1)}_{-1/2,1/2} + \chi^{(1)}_{1/2,-1/2}\right), \tag{11.73}$$

$$\chi^{(2)}_{1,1} = \chi^{(1)}_{1/2,1/2}, \tag{11.74}$$

where $\chi^{(2)}_{s,m_s}$ is shorthand for $\chi^{(2)}_{s_1,s_2;s,m_s}$, *etc.* Likewise, the singlet state is written:

$$\chi^{(2)}_{0,0} = \frac{1}{\sqrt{2}}\left(\chi^{(1)}_{-1/2,1/2} - \chi^{(1)}_{1/2,-1/2}\right). \tag{11.75}$$

Exercises

1. An electron in a hydrogen atom occupies the combined spin and position state

$$R_{2,1}\left(\sqrt{1/3}\,Y_{1,0}\,\chi_+ + \sqrt{2/3}\,Y_{1,1}\,\chi_-\right).$$

   (a) What values would a measurement of L^2 yield, and with what probabilities?
   (b) Same for L_z.
   (c) Same for S^2.
   (d) Same for S_z.
   (e) Same for J^2.
   (f) Same for J_z.
   (g) What is the probability density for finding the electron at r, θ, ϕ?
   (h) What is the probability density for finding the electron in the spin up state (with respect to the z-axis) at radius r?

2. In a low energy neutron-proton system (with zero orbital angular momentum) the potential energy is given by

$$V(r) = V_1(r) + V_2(r)\left(3\,\frac{(\boldsymbol{\sigma}_1\cdot\mathbf{r})\,(\boldsymbol{\sigma}_2\cdot\mathbf{r})}{r^2} - \boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2\right) + V_3(r)\,\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2,$$

   where $\boldsymbol{\sigma}_1$ denotes the vector of the Pauli matrices of the neutron, and $\boldsymbol{\sigma}_2$ denotes the vector of the Pauli matrices of the proton. Calculate the potential energy for the neutron-proton system:

(a) In the spin singlet state.

(b) In the spin triplet state.

3. Consider two electrons in a spin singlet state.

   (a) If a measurement of the spin of one of the electrons shows that it is in the state with $S_z = \hbar/2$, what is the probability that a measurement of the z-component of the spin of the other electron yields $S_z = \hbar/2$?

   (b) If a measurement of the spin of one of the electrons shows that it is in the state with $S_y = \hbar/2$, what is the probability that a measurement of the x-component of the spin of the other electron yields $S_x = -\hbar/2$?

   Finally, if electron 1 is in a spin state described by $\cos\alpha_1\,\chi_+ + \sin\alpha_1\,e^{\,i\,\beta_1}\,\chi_-$, and electron 2 is in a spin state described by $\cos\alpha_2\,\chi_+ + \sin\alpha_2\,e^{\,i\,\beta_2}\,\chi_-$, what is the probability that the two-electron spin state is a triplet state?

12 Time-Independent Perturbation Theory

12.1 Introduction

Consider the following very commonly occurring problem. The Hamiltonian of a quantum mechanical system is written

$$H = H_0 + H_1. \tag{12.1}$$

Here, H_0 is a simple Hamiltonian whose eigenvalues and eigenstates are known *exactly*. H_1 introduces some interesting additional physics into the problem, but is sufficiently complicated that when we add it to H_0 we can no longer find the exact energy eigenvalues and eigenstates. However, H_1 can, in some sense (which we shall specify more precisely later on), be regarded as being *small* compared to H_0. Can we find approximate eigenvalues and eigenstates of the modified Hamiltonian, $H_0 + H_1$, by performing some sort of perturbation expansion about the eigenvalues and eigenstates of the original Hamiltonian, H_0? Let us investigate.

Incidentally, in this chapter, we shall only discuss so-called *time-independent perturbation theory*, in which the modification to the Hamiltonian, H_1, has no explicit dependence on time. It is also assumed that the unperturbed Hamiltonian, H_0, is time-independent.

12.2 Improved Notation

Before commencing our investigation, it is helpful to introduce some improved notation. Let the ψ_i be a complete set of eigenstates of the Hamiltonian, H, corresponding to the eigenvalues E_i: *i.e.*,

$$H\,\psi_i = E_i\,\psi_i. \tag{12.2}$$

Now, we expect the ψ_i to be orthonormal (see Sect. 4.9). In one dimension, this implies that

$$\int_{-\infty}^{\infty} \psi_i^\ast\,\psi_j\,dx = \delta_{ij}. \tag{12.3}$$

In three dimensions (see Cha. 7), the above expression generalizes to

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi_i^\ast\,\psi_j\,dx\,dy\,dz = \delta_{ij}. \tag{12.4}$$

Finally, if the ψ_i are spinors (see Cha. 10) then we have

$$\psi_i^\dagger\,\psi_j = \delta_{ij}. \tag{12.5}$$

The generalization to the case where ψ is a product of a regular wavefunction and a spinor is fairly obvious. We can represent all of the above possibilities by writing

$$\langle\psi_i|\psi_j\rangle \equiv \langle i|j\rangle = \delta_{ij}. \tag{12.6}$$

Here, the term in angle brackets represents the integrals in Eqs. (12.3) and (12.4) in one- and three-dimensional regular space, respectively, and the spinor product (12.5) in spin-space. The advantage of our new notation is its great generality: *i.e.*, it can deal with one-dimensional wavefunctions, three-dimensional wavefunctions, spinors, *etc.*

Expanding a general wavefunction, ψ_a, in terms of the energy eigenstates, ψ_i, we obtain

$$\psi_a = \sum_i c_i\,\psi_i. \tag{12.7}$$

In one dimension, the expansion coefficients take the form (see Sect. 4.9)

$$c_i = \int_{-\infty}^{\infty} \psi_i^*\,\psi_a\,dx, \tag{12.8}$$

whereas in three dimensions we get

$$c_i = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi_i^*\,\psi_a\,dx\,dy\,dz. \tag{12.9}$$

Finally, if ψ is a spinor then we have

$$c_i = \psi_i^\dagger\,\psi_a. \tag{12.10}$$

We can represent all of the above possibilities by writing

$$c_i = \langle\psi_i|\psi_a\rangle \equiv \langle i|a\rangle. \tag{12.11}$$

The expansion (12.7) thus becomes

$$\psi_a = \sum_i \langle\psi_i|\psi_a\rangle\,\psi_i \equiv \sum_i \langle i|a\rangle\,\psi_i. \tag{12.12}$$

Incidentally, it follows that

$$\langle i|a\rangle^* = \langle a|i\rangle. \tag{12.13}$$

Finally, if A is a general operator, and the wavefunction ψ_a is expanded in the manner shown in Eq. (12.7), then the expectation value of A is written (see Sect. 4.9)

$$\langle A\rangle = \sum_{i,j} c_i^*\,c_j\,A_{ij}. \tag{12.14}$$

Here, the A_{ij} are unsurprisingly known as the *matrix elements* of A. In one dimension, the matrix elements take the form

$$A_{ij} = \int_{-\infty}^{\infty} \psi_i^*\,A\,\psi_j\,dx, \tag{12.15}$$

whereas in three dimensions we get

$$A_{ij} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi_i^*\,A\,\psi_j\,dx\,dy\,dz. \tag{12.16}$$

Finally, if ψ is a spinor then we have

$$A_{ij} = \psi_i^\dagger\, A\, \psi_j. \tag{12.17}$$

We can represent all of the above possibilities by writing

$$A_{ij} = \langle\psi_i|A|\psi_j\rangle \equiv \langle i|A|j\rangle. \tag{12.18}$$

The expansion (12.14) thus becomes

$$\langle A\rangle \equiv \langle a|A|a\rangle = \sum_{i,j}\langle a|i\rangle\langle i|A|j\rangle\langle j|a\rangle. \tag{12.19}$$

Incidentally, it follows that [see Eq. (4.58)]

$$\langle i|A|j\rangle^* = \langle j|A^\dagger|i\rangle. \tag{12.20}$$

Finally, it is clear from Eq. (12.19) that

$$\sum_i |i\rangle\langle i| \equiv 1, \tag{12.21}$$

where the ψ_i are a *complete* set of eigenstates, and 1 is the identity operator.

12.3 Two-State System

Consider the simplest possible non-trivial quantum mechanical system. In such a system, there are only *two* independent eigenstates of the unperturbed Hamiltonian: *i.e.*,

$$H_0\,\psi_1 = E_1\,\psi_1, \tag{12.22}$$

$$H_0\,\psi_2 = E_2\,\psi_2. \tag{12.23}$$

It is assumed that these states, and their associated eigenvalues, are known. We also expect the states to be orthonormal, and to form a complete set.

Let us now try to solve the modified energy eigenvalue problem

$$(H_0 + H_1)\,\psi_E = E\,\psi_E. \tag{12.24}$$

We can, in fact, solve this problem exactly. Since the eigenstates of H_0 form a complete set, we can write [see Eq. (12.12)]

$$\psi_E = \langle 1|E\rangle\,\psi_1 + \langle 2|E\rangle\,\psi_2. \tag{12.25}$$

It follows from (12.24) that

$$\langle i|H_0 + H_1|E\rangle = E\,\langle i|E\rangle, \tag{12.26}$$

where $i = 1$ or 2. Equations (12.22), (12.23), (12.25), (12.26), and the orthonormality condition

$$\langle i|j\rangle = \delta_{ij}, \tag{12.27}$$

yield two coupled equations which can be written in matrix form:

$$\begin{pmatrix} E_1 - E + e_{11} & e_{12} \\ e_{12}^* & E_2 - E + e_{22} \end{pmatrix} \begin{pmatrix} \langle 1|E\rangle \\ \langle 2|E\rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{12.28}$$

where

$$\begin{aligned} e_{11} &= \langle 1|H_1|1\rangle, & (12.29) \\ e_{22} &= \langle 2|H_1|2\rangle, & (12.30) \\ e_{12} &= \langle 1|H_1|2\rangle = \langle 2|H_1|1\rangle^*. & (12.31) \end{aligned}$$

Here, use has been made of the fact that H_1 is an Hermitian operator.

Consider the special (but not uncommon) case of a perturbing Hamiltonian whose diagonal matrix elements are zero, so that

$$e_{11} = e_{22} = 0. \tag{12.32}$$

The solution of Eq. (12.28) (obtained by setting the determinant of the matrix to zero) is

$$E = \frac{(E_1 + E_2) \pm \sqrt{(E_1 - E_2)^2 + 4\,|e_{12}|^2}}{2}. \tag{12.33}$$

Let us expand in the supposedly small parameter

$$\epsilon = \frac{|e_{12}|}{|E_1 - E_2|}. \tag{12.34}$$

We obtain

$$E \simeq \frac{1}{2}\,(E_1 + E_2) \pm \frac{1}{2}\,(E_1 - E_2)(1 + 2\,\epsilon^2 + \cdots). \tag{12.35}$$

The above expression yields the modification of the energy eigenvalues due to the perturbing Hamiltonian:

$$\begin{aligned} E_1' &= E_1 + \frac{|e_{12}|^2}{E_1 - E_2} + \cdots, & (12.36) \\ E_2' &= E_2 - \frac{|e_{12}|^2}{E_1 - E_2} + \cdots. & (12.37) \end{aligned}$$

Note that H_1 causes the upper eigenvalue to rise, and the lower to fall. It is easily demonstrated that the modified eigenstates take the form

$$\begin{aligned} \psi_1' &= \psi_1 + \frac{e_{12}^*}{E_1 - E_2}\,\psi_2 + \cdots, & (12.38) \\ \psi_2' &= \psi_2 - \frac{e_{12}}{E_1 - E_2}\,\psi_1 + \cdots. & (12.39) \end{aligned}$$

Thus, the modified energy eigenstates consist of one of the unperturbed eigenstates, plus a slight admixture of the other. Now our expansion procedure is only valid when $\epsilon \ll 1$. This suggests that the condition for the validity of the perturbation method as a whole is

$$|e_{12}| \ll |E_1 - E_2|. \tag{12.40}$$

In other words, when we say that H_1 needs to be small compared to H_0, what we are really saying is that the above inequality must be satisfied.

12.4 Non-Degenerate Perturbation Theory

Let us now generalize our perturbation analysis to deal with systems possessing more than two energy eigenstates. Consider a system in which the energy eigenstates of the unperturbed Hamiltonian, H_0, are denoted

$$H_0\,\psi_n = E_n\,\psi_n, \tag{12.41}$$

where n runs from 1 to N. The eigenstates are assumed to be orthonormal, so that

$$\langle m|n\rangle = \delta_{nm}, \tag{12.42}$$

and to form a complete set. Let us now try to solve the energy eigenvalue problem for the perturbed Hamiltonian:

$$(H_0 + H_1)\,\psi_E = E\,\psi_E. \tag{12.43}$$

If follows that

$$\langle m|H_0 + H_1|E\rangle = E\,\langle m|E\rangle, \tag{12.44}$$

where m can take any value from 1 to N. Now, we can express ψ_E as a linear superposition of the unperturbed energy eigenstates:

$$\psi_E = \sum_k \langle k|E\rangle\,\psi_k, \tag{12.45}$$

where k runs from 1 to N. We can combine the above equations to give

$$(E_m - E + e_{mm})\,\langle m|E\rangle + \sum_{k\neq m} e_{mk}\,\langle k|E\rangle = 0, \tag{12.46}$$

where

$$e_{mk} = \langle m|H_1|k\rangle. \tag{12.47}$$

Let us now develop our perturbation expansion. We assume that

$$\frac{e_{mk}}{E_m - E_k} \sim \mathcal{O}(\epsilon) \tag{12.48}$$

for all $m \neq k$, where $\epsilon \ll 1$ is our expansion parameter. We also assume that

$$\frac{e_{mm}}{E_m} \sim \mathcal{O}(\epsilon) \tag{12.49}$$

for all m. Let us search for a modified version of the nth unperturbed energy eigenstate for which

$$E = E_n + \mathcal{O}(\epsilon), \tag{12.50}$$

and

$$\langle n|E\rangle = 1, \tag{12.51}$$

$$\langle m|E\rangle = \mathcal{O}(\epsilon) \tag{12.52}$$

for $m \neq n$. Suppose that we write out Eq. (12.46) for $m \neq n$, neglecting terms which are $\mathcal{O}(\epsilon^2)$ according to our expansion scheme. We find that

$$(E_m - E_n)\,\langle m|E\rangle + e_{mn} \simeq 0, \tag{12.53}$$

giving

$$\langle m|E\rangle \simeq -\frac{e_{mn}}{E_m - E_n}. \tag{12.54}$$

Substituting the above expression into Eq. (12.46), evaluated for $m = n$, and neglecting $\mathcal{O}(\epsilon^3)$ terms, we obtain

$$(E_n - E + e_{nn}) - \sum_{k\neq n} \frac{|e_{nk}|^2}{E_k - E_n} \simeq 0. \tag{12.55}$$

Thus, the modified nth energy eigenstate possesses an eigenvalue

$$E_n' = E_n + e_{nn} + \sum_{k\neq n} \frac{|e_{nk}|^2}{E_n - E_k} + \mathcal{O}(\epsilon^3), \tag{12.56}$$

and a wavefunction

$$\psi_n' = \psi_n + \sum_{k\neq n} \frac{e_{kn}}{E_n - E_k}\,\psi_k + \mathcal{O}(\epsilon^2). \tag{12.57}$$

Incidentally, it is easily demonstrated that the modified eigenstates remain orthonormal to $\mathcal{O}(\epsilon^2)$.

12.5 Quadratic Stark Effect

Suppose that a hydrogen atom is subject to a uniform external electric field, of magnitude $|\mathbf{E}|$, directed along the z-axis. The Hamiltonian of the system can be split into two parts. Namely, the unperturbed Hamiltonian,

$$H_0 = \frac{p^2}{2\,m_e} - \frac{e^2}{4\pi\epsilon_0\,r}, \tag{12.58}$$

and the perturbing Hamiltonian

$$H_1 = e\,|\mathbf{E}|\,z. \tag{12.59}$$

Note that the electron spin is irrelevant to this problem (since the spin operators all commute with H_1), so we can ignore the spin degrees of freedom of the system. Hence, the energy eigenstates of the unperturbed Hamiltonian are characterized by three quantum numbers—the radial quantum number n, and the two angular quantum numbers l and m (see Cha. 9). Let us denote these states as the ψ_{nlm}, and let their corresponding energy eigenvalues be the E_{nlm}. According to the analysis in the previous section, the change in energy of the eigenstate characterized by the quantum numbers n, l, m in the presence of a *small* electric field is given by

$$\begin{aligned}\Delta E_{nlm} &= e\,|\mathbf{E}|\,\langle n, l, m|z|n, l, m\rangle \\ &\quad + e^2\,|\mathbf{E}|^2 \sum_{n', l', m' \neq n, l, m} \frac{|\langle n, l, m|z|n', l', m'\rangle|^2}{E_{nlm} - E_{n'l'm'}}.\end{aligned} \tag{12.60}$$

This energy-shift is known as the *Stark effect*.

The sum on the right-hand side of the above equation seems very complicated. However, it turns out that most of the terms in this sum are zero. This follows because the matrix elements $\langle n, l, m|z|n', l', m'\rangle$ are zero for virtually all choices of the two sets of quantum number, n, l, m and n', l', m'. Let us try to find a set of rules which determine when these matrix elements are non-zero. These rules are usually referred to as the *selection rules* for the problem in hand.

Now, since [see Eq. (8.4)]

$$L_z = x\,p_y - y\,p_x, \tag{12.61}$$

it follows that [see Eqs. (7.15)–(7.17)]

$$[L_z, z] = 0. \tag{12.62}$$

Thus,

$$\begin{aligned}\langle n, l, m|[L_z, z]|n', l', m'\rangle &= \langle n, l, m|L_z\,z - z\,L_z|n', l', m'\rangle \\ &= \hbar\,(m - m')\,\langle n, l, m|z|n', l', m'\rangle = 0,\end{aligned} \tag{12.63}$$

since ψ_{nlm} is, by definition, an eigenstate of L_z corresponding to the eigenvalue $m\,\hbar$. Hence, it is clear, from the above equation, that one of the selection rules is that the matrix element $\langle n, l, m|z|n', l', m'\rangle$ is zero unless

$$m' = m. \tag{12.64}$$

Let us now determine the selection rule for l. We have

$$[L^2, z] = [L_x^2, z] + [L_y^2, z]$$

$$\begin{aligned}
&= L_x\,[L_x, z] + [L_x, z]\,L_x + L_y\,[L_y, z] + [L_y, z]\,L_y \\
&= i\,\hbar\,(-L_x\,y - y\,L_x + L_y\,x + x\,L_y) \\
&= 2\,i\,\hbar\,(L_y\,x - L_x\,y + i\,\hbar\,z) \\
&= 2\,i\,\hbar\,(L_y\,x - y\,L_x) = 2\,i\,\hbar\,(x\,L_y - L_x\,y),
\end{aligned} \tag{12.65}$$

where use has been made of Eqs. (7.15)–(7.17), (8.2)–(8.4), and (8.10). Thus,

$$\begin{aligned}
[L^2, [L^2, z]] &= 2\,i\,\hbar\,\left(L^2, L_y\,x - L_x\,y + i\,\hbar\,z\right) \\
&= 2\,i\,\hbar\,\left(L_y\,[L^2, x] - L_x\,[L^2, y] + i\,\hbar\,[L^2, z]\right) \\
&= -4\,\hbar^2\,L_y\,(y\,L_z - L_y\,z) + 4\,\hbar^2\,L_x\,(L_x\,z - x\,L_z) \\
&\quad -2\,\hbar^2\,(L^2\,z - z\,L^2),
\end{aligned} \tag{12.66}$$

which reduces to

$$\begin{aligned}
[L^2, [L^2, z]] &= -\hbar^2\,\left\{4\,(L_x\,x + L_y\,y + L_z\,z)\,L_z - 4\,(L_x^2 + L_y^2 + L_z^2)\,z \right. \\
&\quad \left. +2\,(L^2\,z - z\,L^2)\right\} \\
&= -\hbar^2\,\left\{4\,(L_x\,x + L_y\,y + L_z\,z)\,L_z - 2\,(L^2\,z + z\,L^2)\right\}.
\end{aligned} \tag{12.67}$$

However, it is clear from Eqs. (8.2)–(8.4) that

$$L_x\,x + L_y\,y + L_z\,z = 0. \tag{12.68}$$

Hence, we obtain

$$[L^2, [L^2, z]] = 2\,\hbar^2\,(L^2\,z + z\,L^2). \tag{12.69}$$

Finally, the above expression expands to give

$$L^4\,z - 2\,L^2\,z\,L^2 + z\,L^4 - 2\,\hbar^2\,(L^2\,z + z\,L^2) = 0. \tag{12.70}$$

Equation (12.70) implies that

$$\langle n, l, m|L^4\,z - 2\,L^2\,z\,L^2 + z\,L^4 - 2\,\hbar^2\,(L^2\,z + z\,L^2)|n', l', m\rangle = 0. \tag{12.71}$$

Since, by definition, ψ_{nlm} is an eigenstate of L^2 corresponding to the eigenvalue $l\,(l+1)\,\hbar^2$, this expression yields

$$\begin{aligned}
&\left\{l^2\,(l+1)^2 - 2\,l\,(l+1)\,l'\,(l'+1) + l'^2\,(l'+1)^2 \right. \\
&\quad \left. -2\,l\,(l+1) - 2\,l'\,(l'+1)\right\}\langle n, l, m|z|n', l', m\rangle = 0,
\end{aligned} \tag{12.72}$$

which reduces to

$$(l + l' + 2)\,(l + l')\,(l - l' + 1)\,(l - l' - 1)\,\langle n, l, m|z|n', l', m\rangle = 0. \tag{12.73}$$

According to the above formula, the matrix element $\langle n, l, m|z|n', l', m\rangle$ vanishes unless $l = l' = 0$ or $l' = l \pm 1$. [Of course, the factor $l + l' + 2$, in the above equation, can never be zero, since l and l' can never be negative.] Recall, however, from Cha. 9, that an $l = 0$ wavefunction is *spherically symmetric*. It, therefore, follows, from symmetry, that the matrix element $\langle n, l, m|z|n', l', m\rangle$ is zero when $l = l' = 0$. In conclusion, the selection rule for l is that the matrix element $\langle n, l, m|z|n', l', m\rangle$ is zero unless

$$l' = l \pm 1. \tag{12.74}$$

Application of the selection rules (12.64) and (12.74) to Eq. (12.60) yields

$$\Delta E_{nlm} = e^2\,|\mathbf{E}|^2 \sum_{n',l'=l\pm1} \frac{|\langle n, l, m|z|n', l', m\rangle|^2}{E_{nlm} - E_{n'l'm}}. \tag{12.75}$$

Note that, according to the selection rules, all of the terms in Eq. (12.60) which vary linearly with the electric field-strength vanish. Only those terms which vary quadratically with the field-strength survive. Hence, this type of energy-shift of an atomic state in the presence of a small electric field is known as the *quadratic* Stark effect. Now, the *electric polarizability* of an atom is defined in terms of the energy-shift of the atomic state as follows:

$$\Delta E = -\frac{1}{2}\,\alpha\,|\mathbf{E}|^2. \tag{12.76}$$

Hence, we can write

$$\alpha_{nlm} = 2\,e^2 \sum_{n',l'=l\pm1} \frac{|\langle n, l, m|z|n', l', m\rangle|^2}{E_{n'l'm} - E_{nlm}}. \tag{12.77}$$

Unfortunately, there is one fairly obvious problem with Eq. (12.75). Namely, it predicts an *infinite* energy-shift if there exists some non-zero matrix element $\langle n, l, m|z|n', l', m\rangle$ which couples two *degenerate* unperturbed energy eigenstates: *i.e.*, if $\langle n, l, m|z|n', l', m\rangle \neq 0$ and $E_{nlm} = E_{n'l'm}$. Clearly, our perturbation method breaks down completely in this situation. Hence, we conclude that Eqs. (12.75) and (12.77) are only applicable to cases where the coupled eigenstates are *non-degenerate*. For this reason, the type of perturbation theory employed here is known as *non-degenerate perturbation theory*. Now, the unperturbed eigenstates of a hydrogen atom have energies which only depend on the radial quantum number n (see Cha. 9). It follows that we can only apply the above results to the $n = 1$ eigenstate (since for $n > 1$ there will be coupling to degenerate eigenstates with the same value of n but different values of l).

Thus, according to non-degenerate perturbation theory, the polarizability of the ground-state (*i.e.*, $n = 1$) of a hydrogen atom is given by

$$\alpha = 2\,e^2 \sum_{n>1} \frac{|\langle 1, 0, 0|z|n, 1, 0\rangle|^2}{E_{n00} - E_{100}}. \tag{12.78}$$

Here, we have made use of the fact that $E_{n10} = E_{n00}$. The sum in the above expression can be evaluated approximately by noting that (see Sect. 9.4)

$$E_{n00} = -\frac{e^2}{8\pi\,\epsilon_0\,a_0\,n^2}, \tag{12.79}$$

where

$$a_0 = \frac{4\pi\epsilon_0\,\hbar^2}{m_e\,e^2} \tag{12.80}$$

is the Bohr radius. Hence, we can write

$$E_{n00} - E_{100} \geq E_{200} - E_{100} = \frac{3}{4}\,\frac{e^2}{8\pi\,\epsilon_0\,a_0}, \tag{12.81}$$

which implies that

$$\alpha < \frac{16}{3}\,4\pi\epsilon_0\,a_0 \sum_{n>1} |\langle 1,0,0|z|n,1,0\rangle|^2. \tag{12.82}$$

However, [see Eq. (12.21)]

$$\begin{aligned}
\sum_{n>1} |\langle 1,0,0|z|n,1,0\rangle|^2 &= \sum_{n>1} \langle 1,0,0|z|n,1,0\rangle\,\langle n,1,0|z|1,0,0\rangle \\
&= \sum_{n',l',m'} \langle 1,0,0|z|n',l',m'\rangle\,\langle n',l',m'|z|1,0,0\rangle \\
&= \langle 1,0,0|z^2|1,0,0\rangle = \frac{1}{3}\,\langle 1,0,0|r^2|1,0,0\rangle,
\end{aligned} \tag{12.83}$$

where we have made use of the selection rules, the fact that the $\psi_{n',l',m'}$ form a complete set, and the fact the the ground-state of hydrogen is spherically symmetric. Finally, it follows from Eq. (9.72) that

$$\langle 1,0,0|r^2|1,0,0\rangle = 3\,a_0^2. \tag{12.84}$$

Hence, we conclude that

$$\alpha < \frac{16}{3}\,4\pi\epsilon_0\,a_0^3 \simeq 5.3\;4\pi\epsilon_0\,a_0^3. \tag{12.85}$$

The exact result (which can be obtained by solving Schrödinger's equation in parabolic coordinates) is

$$\alpha = \frac{9}{2}\,4\pi\epsilon_0\,a_0^3 = 4.5\;4\pi\epsilon_0\,a_0^3. \tag{12.86}$$

12.6 Degenerate Perturbation Theory

Let us, rather naively, investigate the Stark effect in an excited (*i.e.*, $n > 1$) state of the hydrogen atom using standard non-degenerate perturbation theory. We can write

$$H_0\,\psi_{nlm} = E_n\,\psi_{nlm}, \tag{12.87}$$

since the energy eigenstates of the unperturbed Hamiltonian only depend on the quantum number n. Making use of the selection rules (12.64) and (12.74), non-degenerate perturbation theory yields the following expressions for the perturbed energy levels and eigenstates [see Eqs. (12.56) and (12.57)]:

$$E'_{nl} = E_n + e_{nlnl} + \sum_{n',l'=l\pm1} \frac{|e_{n'l'nl}|^2}{E_n - E_{n'}}, \tag{12.88}$$

and

$$\psi'_{nlm} = \psi_{nlm} + \sum_{n',l'=l\pm1} \frac{e_{n'l'nl}}{E_n - E_{n'}}\,\psi_{n'l'm}, \tag{12.89}$$

where

$$e_{n'l'nl} = \langle n', l', m|H_1|n, l, m\rangle. \tag{12.90}$$

Unfortunately, if $n > 1$ then the summations in the above expressions are not well-defined, because there exist non-zero matrix elements, $e_{nl'nl}$, which couple degenerate eigenstates: *i.e.*, there exist non-zero matrix elements which couple states with the same value of n, but different values of l. These particular matrix elements give rise to singular factors $1/(E_n - E_n)$ in the summations. This does not occur if $n = 1$ because, in this case, the selection rule $l' = l \pm 1$, and the fact that $l = 0$ (since $0 \leq l < n$), only allow l' to take the single value 1. Of course, there is no $n = 1$ state with $l' = 1$. Hence, there is only one coupled state corresponding to the eigenvalue E_1. Unfortunately, if $n > 1$ then there are multiple coupled states corresponding to the eigenvalue E_n.

Note that our problem would disappear if the matrix elements of the perturbed Hamiltonian corresponding to the same value of n, but different values of l, were all zero: *i.e.*, if

$$\langle n, l', m|H_1|n, l, m\rangle = \lambda_{nl}\,\delta_{ll'}. \tag{12.91}$$

In this case, all of the singular terms in Eqs. (12.88) and (12.89) would reduce to zero. Unfortunately, the above equation is not satisfied. Fortunately, we can always redefine the unperturbed eigenstates corresponding to the eigenvalue E_n in such a manner that Eq. (12.91) is satisfied. Suppose that there are N_n coupled eigenstates belonging to the eigenvalue E_n. Let us define N_n new states which are linear combinations of our N_n original degenerate eigenstates:

$$\psi^{(1)}_{nlm} = \sum_{k=1,N_n} \langle n, k, m|n, l^{(1)}, m\rangle\,\psi_{nkm}. \tag{12.92}$$

Note that these new states are also degenerate energy eigenstates of the unperturbed Hamiltonian, H_0, corresponding to the eigenvalue E_n. The $\psi^{(1)}_{nlm}$ are chosen in such a manner that they are also eigenstates of the perturbing Hamiltonian, H_1: *i.e.*, they are *simultaneous eigenstates* of H_0 and H_1. Thus,

$$H_1\,\psi^{(1)}_{nlm} = \lambda_{nl}\,\psi^{(1)}_{nlm}. \tag{12.93}$$

The $\psi_{nlm}^{(1)}$ are also chosen so as to be orthonormal: *i.e.*,

$$\langle n, l'^{(1)}, m|n, l^{(1)}, m\rangle = \delta_{ll'}. \tag{12.94}$$

It follows that

$$\langle n, l'^{(1)}, m|H_1|n, l^{(1)}, m\rangle = \lambda_{nl}\,\delta_{ll'}. \tag{12.95}$$

Thus, if we use the new eigenstates, instead of the old ones, then we can employ Eqs. (12.88) and (12.89) directly, since all of the singular terms vanish. The only remaining difficulty is to determine the new eigenstates in terms of the original ones.

Now [see Eq. (12.21)]

$$\sum_{l=1,N_n} |n, l, m\rangle\langle n, l, m| \equiv 1, \tag{12.96}$$

where 1 denotes the identity operator in the sub-space of all coupled unperturbed eigenstates corresponding to the eigenvalue E_n. Using this completeness relation, the eigenvalue equation (12.93) can be transformed into a straightforward matrix equation:

$$\sum_{l''=1,N_n} \langle n, l', m|H_1|n, l'', m\rangle\,\langle n, l'', m|n, l^{(1)}, m\rangle = \lambda_{nl}\,\langle n, l', m|n, l^{(1)}, m\rangle. \tag{12.97}$$

This can be written more transparently as

$$\mathbf{U}\,\mathbf{x} = \lambda\,\mathbf{x}, \tag{12.98}$$

where the elements of the $N_n \times N_n$ Hermitian matrix $\mathbf{U}$ are

$$U_{jk} = \langle n, j, m|H_1|n, k, m\rangle. \tag{12.99}$$

Provided that the determinant of $\mathbf{U}$ is non-zero, Eq. (12.98) can always be solved to give N_n eigenvalues λ_{nl} (for $l = 1$ to N_n), with N_n corresponding eigenvectors $\mathbf{x}_{nl}$. The normalized eigenvectors specify the weights of the new eigenstates in terms of the original eigenstates: *i.e.*,

$$(\mathbf{x}_{nl})_k = \langle n, k, m|n, l^{(1)}, m\rangle, \tag{12.100}$$

for $k = 1$ to N_n. In our new scheme, Eqs. (12.88) and (12.89) yield

$$E_{nl}' = E_n + \lambda_{nl} + \sum_{n'\neq n, l'=l\pm 1} \frac{|e_{n'l'nl}|^2}{E_n - E_{n'}}, \tag{12.101}$$

and

$$\psi_{nlm}^{(1)'} = \psi_{nlm}^{(1)} + \sum_{n'\neq n, l'=l\pm 1} \frac{e_{n'l'nl}}{E_n - E_{n'}}\,\psi_{n'l'm}. \tag{12.102}$$

There are no singular terms in these expressions, since the summations are over $n' \neq n$: *i.e.*, they specifically exclude the problematic, degenerate, unperturbed energy eigenstates corresponding to the eigenvalue E_n. Note that the first-order energy shifts are equivalent to the eigenvalues of the matrix equation (12.98).

12.7 Linear Stark Effect

Returning to the Stark effect, let us examine the effect of an external electric field on the energy levels of the $n = 2$ states of a hydrogen atom. There are four such states: an $l = 0$ state, usually referred to as 2S, and three $l = 1$ states (with $m = -1, 0, 1$), usually referred to as 2P. All of these states possess the same unperturbed energy, $E_{200} = -e^2/(32\pi\,\epsilon_0\,a_0)$. As before, the perturbing Hamiltonian is

$$H_1 = e\,|\mathbf{E}|\,z. \tag{12.103}$$

According to the previously determined selection rules (*i.e.*, $m' = m$, and $l' = l \pm 1$), this Hamiltonian couples ψ_{200} and ψ_{210}. Hence, non-degenerate perturbation theory breaks down when applied to these two states. On the other hand, non-degenerate perturbation theory works fine for the ψ_{211} and ψ_{21-1} states, since these are not coupled to any other $n = 2$ states by the perturbing Hamiltonian.

In order to apply perturbation theory to the ψ_{200} and ψ_{210} states, we have to solve the matrix eigenvalue equation

$$\mathbf{U}\,\mathbf{x} = \lambda\,\mathbf{x}, \tag{12.104}$$

where $\mathbf{U}$ is the matrix of the matrix elements of H_1 between these states. Thus,

$$\mathbf{U} = e\,|\mathbf{E}|\begin{pmatrix} 0 & \langle 2,0,0|z|2,1,0\rangle \\ \langle 2,1,0|z|2,0,0\rangle & 0 \end{pmatrix}, \tag{12.105}$$

where the rows and columns correspond to ψ_{200} and ψ_{210}, respectively. Here, we have again made use of the selection rules, which tell us that the matrix element of z between two hydrogen atom states is zero unless the states possess l quantum numbers which differ by unity. It is easily demonstrated, from the exact forms of the 2S and 2P wavefunctions, that

$$\langle 2,0,0|z|2,1,0\rangle = \langle 2,1,0|z|2,0,0\rangle = 3\,a_0. \tag{12.106}$$

It can be seen, by inspection, that the eigenvalues of $\mathbf{U}$ are $\lambda_1 = 3\,e\,a_0\,|\mathbf{E}|$ and $\lambda_2 = -3\,e\,a_0\,|\mathbf{E}|$. The corresponding normalized eigenvectors are

$$\mathbf{x}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, \tag{12.107}$$

$$\mathbf{x}_2 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}. \tag{12.108}$$

It follows that the simultaneous eigenstates of H_0 and H_1 take the form

$$\psi_1 = \frac{\psi_{200} + \psi_{210}}{\sqrt{2}}, \tag{12.109}$$

$$\psi_2 = \frac{\psi_{200} - \psi_{210}}{\sqrt{2}}. \tag{12.110}$$

In the absence of an external electric field, both of these states possess the same energy, E_{200}. The first-order energy shifts induced by an external electric field are given by

$$\Delta E_1 = +3\,e\,a_0\,|\mathbf{E}|, \tag{12.111}$$

$$\Delta E_2 = -3\,e\,a_0\,|\mathbf{E}|. \tag{12.112}$$

Thus, in the presence of an electric field, the energies of states 1 and 2 are shifted upwards and downwards, respectively, by an amount $3\,e\,a_0\,|\mathbf{E}|$. These states are orthogonal linear combinations of the original ψ_{200} and ψ_{210} states. Note that the energy shifts are *linear* in the electric field-strength, so this effect—which is known as the *linear* Stark effect—is much larger than the quadratic effect described in Sect. 12.5. Note, also, that the energies of the ψ_{211} and ψ_{21-1} states are not affected by the electric field to first-order. Of course, to second-order the energies of these states are shifted by an amount which depends on the square of the electric field-strength (see Sect. 12.5).

12.8 Fine Structure of Hydrogen

According to special relativity, the kinetic energy (*i.e.*, the difference between the total energy and the rest mass energy) of a particle of rest mass m and momentum p is

$$T = \sqrt{p^2\,c^2 + m^2\,c^4} - m\,c^2. \tag{12.113}$$

In the non-relativistic limit $p \ll m\,c$, we can expand the square-root in the above expression to give

$$T = \frac{p^2}{2\,m}\left[1 - \frac{1}{4}\left(\frac{p}{m\,c}\right)^2 + \mathcal{O}\left(\frac{p}{m\,c}\right)^4\right]. \tag{12.114}$$

Hence,

$$T \simeq \frac{p^2}{2\,m} - \frac{p^4}{8\,m^3\,c^2}. \tag{12.115}$$

Of course, we recognize the first term on the right-hand side of this equation as the standard non-relativistic expression for the kinetic energy. The second term is the lowest-order relativistic correction to this energy. Let us consider the effect of this type of correction on the energy levels of a hydrogen atom. So, the unperturbed Hamiltonian is given by Eq. (12.58), and the perturbing Hamiltonian takes the form

$$H_1 = -\frac{p^4}{8\,m_e^3\,c^2}. \tag{12.116}$$

Now, according to standard first-order perturbation theory (see Sect. 12.4), the lowest-order relativistic correction to the energy of a hydrogen atom state characterized by the standard quantum numbers n, l, and m is given by

$$\begin{aligned}\Delta E_{nlm} &= \langle n, l, m|H_1|n, l, m\rangle = -\frac{1}{8\,m_e^3\,c^2}\,\langle n, l, m|p^4|n, l, m\rangle \\ &= -\frac{1}{8\,m_e^3\,c^2}\,\langle n, l, m|p^2\,p^2|n, l, m\rangle.\end{aligned} \tag{12.117}$$

However, Schrödinger's equation for a unperturbed hydrogen atom can be written

$$p^2\,\psi_{n,l,m} = 2\,m_e\,(E_n - V)\,\psi_{n,l,m}, \tag{12.118}$$

where $V = -e^2/(4\pi\epsilon_0\,r)$. Since p^2 is an Hermitian operator, it follows that

$$\begin{aligned}
\Delta E_{nlm} &= -\frac{1}{2\,m_e\,c^2}\,\langle n,l,m|(E_n - V)^2|n,l,m\rangle \\
&= -\frac{1}{2\,m_e\,c^2}\left(E_n^{\,2} - 2\,E_n\,\langle n,l,m|V|n,l,m\rangle + \langle n,l,m|V^2|n,l,m\rangle\right) \\
&= -\frac{1}{2\,m_e\,c^2}\left[E_n^{\,2} + 2\,E_n\left(\frac{e^2}{4\pi\epsilon_0}\right)\left\langle\frac{1}{r}\right\rangle + \left(\frac{e^2}{4\pi\epsilon_0}\right)^2\left\langle\frac{1}{r^2}\right\rangle\right].
\end{aligned} \tag{12.119}$$

It follows from Eqs. (9.74) and (9.75) that

$$\Delta E_{nlm} = -\frac{1}{2\,m_e\,c^2}\left[E_n^{\,2} + 2\,E_n\left(\frac{e^2}{4\pi\epsilon_0}\right)\frac{1}{n^2\,a_0} + \left(\frac{e^2}{4\pi\epsilon_0}\right)^2\frac{1}{(l+1/2)\,n^3\,a_0^2}\right]. \tag{12.120}$$

Finally, making use of Eqs. (9.55), (9.57), and (9.58), the above expression reduces to

$$\Delta E_{nlm} = E_n\,\frac{\alpha^2}{n^2}\left(\frac{n}{l+1/2} - \frac{3}{4}\right), \tag{12.121}$$

where

$$\alpha = \frac{e^2}{4\pi\epsilon_0\,\hbar\,c} \simeq \frac{1}{137} \tag{12.122}$$

is the dimensionless *fine structure constant*.

Note that the above derivation implicitly assumes that p^4 is an Hermitian operator. It turns out that this is not the case for $l = 0$ states. However, somewhat fortuitously, our calculation still gives the correct answer when $l = 0$. Note, also, that we are able to use *non-degenerate* perturbation theory in the above calculation, using the ψ_{nlm} eigenstates, because the perturbing Hamiltonian commutes with both L^2 and L_z. It follows that there is no coupling between states with different l and m quantum numbers. Hence, all coupled states have different n quantum numbers, and therefore have different energies.

Now, an electron in a hydrogen atom experiences an electric field

$$\mathbf{E} = \frac{e\,\mathbf{r}}{4\pi\epsilon_0\,r^3} \tag{12.123}$$

due to the charge on the nucleus. However, according to electromagnetic theory, a non-relativistic particle moving in a electric field **E** with velocity **v** also experiences an effective magnetic field

$$\mathbf{B} = -\frac{\mathbf{v}\times\mathbf{E}}{c^2}. \tag{12.124}$$

Recall, that an electron possesses a magnetic moment [see Eqs. (10.58) and (10.59)]

$$\mu = -\frac{e}{m_e}\,\mathbf{S} \tag{12.125}$$

due to its spin angular momentum, **S**. We, therefore, expect an additional contribution to the Hamiltonian of a hydrogen atom of the form [see Eq. (10.60)]

$$\begin{aligned} H_1 &= -\mu\cdot\mathbf{B} \\ &= -\frac{e^2}{4\pi\epsilon_0\, m_e\, c^2\, r^3}\,\mathbf{v}\times\mathbf{r}\cdot\mathbf{S} \\ &= \frac{e^2}{4\pi\epsilon_0\, m_e^2\, c^2\, r^3}\,\mathbf{L}\cdot\mathbf{S}, \end{aligned} \tag{12.126}$$

where $\mathbf{L} = m_e\,\mathbf{r}\times\mathbf{v}$ is the electron's orbital angular momentum. This effect is known as *spin-orbit coupling*. It turns out that the above expression is too large, by a factor 2, due to an obscure relativistic effect known as *Thomas precession*. Hence, the true spin-orbit correction to the Hamiltonian is

$$H_1 = \frac{e^2}{8\pi\,\epsilon_0\, m_e^2\, c^2\, r^3}\,\mathbf{L}\cdot\mathbf{S}. \tag{12.127}$$

Let us now apply perturbation theory to the hydrogen atom, using the above expression as the perturbing Hamiltonian.

Now

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \tag{12.128}$$

is the total angular momentum of the system. Hence,

$$J^2 = L^2 + S^2 + 2\,\mathbf{L}\cdot\mathbf{S}, \tag{12.129}$$

giving

$$\mathbf{L}\cdot\mathbf{S} = \frac{1}{2}\,(J^2 - L^2 - S^2). \tag{12.130}$$

Recall, from Sect. 11.2, that whilst J^2 commutes with both L^2 and S^2, it does not commute with either L_z or S_z. It follows that the perturbing Hamiltonian (12.127) also commutes with both L^2 and S^2, but does not commute with either L_z or S_z. Hence, the simultaneous eigenstates of the unperturbed Hamiltonian (12.58) and the perturbing Hamiltonian (12.127) are the same as the simultaneous eigenstates of L^2, S^2, and J^2 discussed in Sect. 11.3. It is important to know this since, according to Sect. 12.6, we can only safely apply perturbation theory to the simultaneous eigenstates of the unperturbed and perturbing Hamiltonians.

Adopting the notation introduced in Sect. 11.3, let $\psi^{(2)}_{l,s;j,m_j}$ be a simultaneous eigenstate of L^2, S^2, J^2, and J_z corresponding to the eigenvalues

$$L^2\,\psi^{(2)}_{l,s;j,m_j} = l\,(l+1)\,\hbar^2\,\psi^{(2)}_{l,s;j,m_j}, \tag{12.131}$$

$$S^2\,\psi^{(2)}_{l,s;j,m_j} = s\,(s+1)\,\hbar^2\,\psi^{(2)}_{l,s;j,m_j}, \tag{12.132}$$

$$J^2\,\psi^{(2)}_{l,s;j,m_j} = j\,(j+1)\,\hbar^2\,\psi^{(2)}_{l,s;j,m_j}, \tag{12.133}$$

$$J_z\,\psi^{(2)}_{l,s;j,m_j} = m_j\,\hbar\,\psi^{(2)}_{l,s;j,m_j}. \tag{12.134}$$

According to standard first-order perturbation theory, the energy-shift induced in such a state by spin-orbit coupling is given by

$$\begin{aligned}\Delta E_{l,1/2;j,m_j} &= \langle l,1/2;j,m_j|H_1|l,1/2;j,m_j\rangle\\ &= \frac{e^2}{16\pi\,\epsilon_0\,m_e^2\,c^2}\left\langle 1,1/2;j,m_j\left|\frac{J^2-L^2-S^2}{r^3}\right|l,1/2;j,m_j\right\rangle\\ &= \frac{e^2\,\hbar^2}{16\pi\,\epsilon_0\,m_e^2\,c^2}\,[j\,(j+1)-l\,(l+1)-3/4]\left\langle\frac{1}{r^3}\right\rangle.\end{aligned} \tag{12.135}$$

Here, we have made use of the fact that $s=1/2$ for an electron. It follows from Eq. (9.76) that

$$\Delta E_{l,1/2;j,m_j} = \frac{e^2\,\hbar^2}{16\pi\,\epsilon_0\,m_e^2\,c^2\,a_0^3}\left[\frac{j\,(j+1)-l\,(l+1)-3/4}{l\,(l+1/2)\,(l+1)\,n^3}\right], \tag{12.136}$$

where n is the radial quantum number. Finally, making use of Eqs. (9.55), (9.57), and (9.58), the above expression reduces to

$$\Delta E_{l,1/2;j,m_j} = E_n\,\frac{\alpha^2}{n^2}\left[\frac{n\,\{3/4+l\,(l+1)-j\,(j+1)\}}{2\,l\,(l+1/2)\,(l+1)}\right], \tag{12.137}$$

where α is the fine structure constant. A comparison of this expression with Eq. (12.121) reveals that the energy-shift due to spin-orbit coupling is of the same order of magnitude as that due to the lowest-order relativistic correction to the Hamiltonian. We can add these two corrections together (making use of the fact that $j=l\pm 1/2$ for a hydrogen atom—see Sect. 11.3) to obtain a net energy-shift of

$$\Delta E_{l,1/2;j,m_j} = E_n\,\frac{\alpha^2}{n^2}\left(\frac{n}{j+1/2}-\frac{3}{4}\right). \tag{12.138}$$

This modification of the energy levels of a hydrogen atom due to a combination of relativity and spin-orbit coupling is known as *fine structure*.

Now, it is conventional to refer to the energy eigenstates of a hydrogen atom which are also simultaneous eigenstates of J^2 as nL_j states, where n is the radial quantum number, $L=(S,P,D,F,\cdots)$ as $l=(0,1,2,3,\cdots)$, and j is the total angular momentum quantum number. Let us examine the effect of the fine structure energy-shift (12.138) on these eigenstates for $n=1,2$ and 3.

For $n=1$, in the absence of fine structure, there are two degenerate $1S_{1/2}$ states. According to Eq. (12.138), the fine structure induced energy-shifts of these two states are

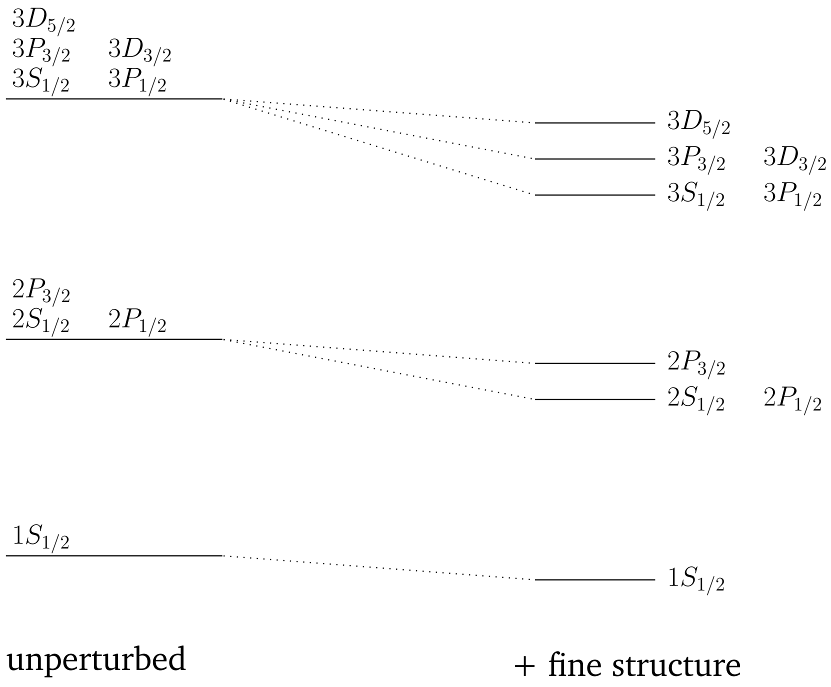

Figure 12.1: *Effect of the fine structure energy-shift on the* $n = 1, 2$ *and 3 states of a hydrogen atom. Not to scale.*

the same. Hence, fine structure does not break the degeneracy of the two $1S_{1/2}$ states of hydrogen.

For $n = 2$, in the absence of fine structure, there are two $2S_{1/2}$ states, two $2P_{1/2}$ states, and four $2P_{3/2}$ states, all of which are degenerate. According to Eq. (12.138), the fine structure induced energy-shifts of the $2S_{1/2}$ and $2P_{1/2}$ states are the same as one another, but are different from the induced energy-shift of the $2P_{3/2}$ states. Hence, fine structure does not break the degeneracy of the $2S_{1/2}$ and $2P_{1/2}$ states of hydrogen, but does break the degeneracy of these states relative to the $2P_{3/2}$ states.

For $n = 3$, in the absence of fine structure, there are two $3S_{1/2}$ states, two $3P_{1/2}$ states, four $3P_{3/2}$ states, four $3D_{3/2}$ states, and six $3D_{5/2}$ states, all of which are degenerate. According to Eq. (12.138), fine structure breaks these states into three groups: the $3S_{1/2}$ and $3P_{1/2}$ states, the $3P_{3/2}$ and $3D_{3/2}$ states, and the $3D_{5/2}$ states.

The effect of the fine structure energy-shift on the $n = 1$, 2, and 3 energy states of a hydrogen atom is illustrated in Fig. 12.1.

Note, finally, that although expression (12.137) does not have a well-defined value for $l = 0$, when added to expression (12.121) it, somewhat fortuitously, gives rise to an expression (12.138) which is both well-defined and correct when $l = 0$.

12.9 Zeeman Effect

Consider a hydrogen atom placed in a uniform z-directed external magnetic field of strength B. The modification to the Hamiltonian of the system is

$$H_1 = -\boldsymbol{\mu}\cdot\mathbf{B}, \tag{12.139}$$

where

$$\boldsymbol{\mu} = -\frac{e}{2\,m_e}\,(\mathbf{L}+2\,\mathbf{S}) \tag{12.140}$$

is the total electron magnetic moment, including both orbital and spin contributions [see Eqs. (10.57)–(10.59)]. Thus,

$$H_1 = \frac{e\,B}{2\,m_e}\,(L_z+2\,S_z). \tag{12.141}$$

Suppose that the applied magnetic field is much weaker than the atom's internal magnetic field (12.124). Since the magnitude of the internal field is about 25 tesla, this is a fairly reasonable assumption. In this situation, we can treat H_1 as a small perturbation acting on the simultaneous eigenstates of the unperturbed Hamiltonian and the fine structure Hamiltonian. Of course, these states are the simultaneous eigenstates of L^2, S^2, J^2, and J_z (see previous section). Hence, from standard perturbation theory, the first-order energy-shift induced by a weak external magnetic field is

$$\begin{aligned}\Delta E_{l,1/2;j,m_j} &= \langle l,1/2;j,m_j|H_1|l,1/2;j,m_j\rangle\\ &= \frac{e\,B}{2\,m_e}\left(m_j\,\hbar+\langle l,1/2;j,m_j|S_z|l,1/2;j,m_j\rangle\right),\end{aligned} \tag{12.142}$$

since $J_z = L_z + S_z$. Now, according to Eqs. (11.47) and (11.48),

$$\psi^{(2)}_{j,m_j} = \left(\frac{j+m_j}{2\,l+1}\right)^{1/2}\psi^{(1)}_{m_j-1/2,1/2}+\left(\frac{j-m_j}{2\,l+1}\right)^{1/2}\psi^{(1)}_{m_j+1/2,-1/2} \tag{12.143}$$

when $j = l + 1/2$, and

$$\psi^{(2)}_{j,m_j} = \left(\frac{j+1-m_j}{2\,l+1}\right)^{1/2}\psi^{(1)}_{m_j-1/2,1/2}-\left(\frac{j+1+m_j}{2\,l+1}\right)^{1/2}\psi^{(1)}_{m_j+1/2,-1/2} \tag{12.144}$$

when $j = l - 1/2$. Here, the $\psi^{(1)}_{m,m_s}$ are the simultaneous eigenstates of L^2, S^2, L_z, and S_z, whereas the $\psi^{(2)}_{j,m_j}$ are the simultaneous eigenstates of L^2, S^2, J^2, and J_z. In particular,

$$S_z\,\psi^{(1)}_{m,\pm 1/2} = \pm\frac{\hbar}{2}\,\psi^{(1)}_{m,\pm 1/2}. \tag{12.145}$$

It follows from Eqs. (12.143)–(12.145), and the orthormality of the $\psi^{(1)}$, that

$$\langle l,1/2;j,m_j|S_z|l,1/2;j,m_j\rangle = \pm\frac{m_j\,\hbar}{2\,l+1} \tag{12.146}$$

when $j = l \pm 1/2$. Thus, the induced energy-shift when a hydrogen atom is placed in an external magnetic field—which is known as the *Zeeman effect*—becomes

$$\Delta E_{l,1/2;j,m_j} = \mu_B\, B\, m_j \left[1 \pm \frac{1}{2\,l+1}\right] \tag{12.147}$$

where the ± signs correspond to $j = l \pm 1/2$. Here,

$$\mu_B = \frac{e\,\hbar}{2\,m_e} = 5.788 \times 10^{-5}\,\text{eV/T} \tag{12.148}$$

is known as the *Bohr magnetron*. Of course, the quantum number m_j takes values differing by unity in the range $-j$ to j. It, thus, follows from Eq. (12.147) that the Zeeman effect splits degenerate states characterized by $j = l + 1/2$ into $2\,j + 1$ equally spaced states of interstate spacing

$$\Delta E_{j=l+1/2} = \mu_B\, B\, \frac{2\,l+2}{2\,l+1}. \tag{12.149}$$

Likewise, the Zeeman effect splits degenerate states characterized by $j = l-1/2$ into $2\,j+1$ equally spaced states of interstate spacing

$$\Delta E_{j=l-1/2} = \mu_B\, B\, \frac{2\,l}{2\,l+1}. \tag{12.150}$$

In conclusion, in the presence of a weak external magnetic field, the two degenerate $1S_{1/2}$ states of the hydrogen atom are split by $2\,\mu_B\,B$. Likewise, the four degenerate $2S_{1/2}$ and $2P_{1/2}$ states are split by $(2/3)\,\mu_B\,B$, whereas the four degenerate $2P_{3/2}$ states are split by $(4/3)\,\mu_B\,B$. This is illustrated in Fig. 12.2. Note, finally, that since the $\psi^{(2)}_{l,m_j}$ are not simultaneous eigenstates of the unperturbed and perturbing Hamiltonians, Eqs. (12.149) and (12.150) can only be regarded as the expectation values of the magnetic-field induced energy-shifts. However, as long as the external magnetic field is much weaker than the internal magnetic field, these expectation values are almost identical to the actual measured values of the energy-shifts.

12.10 Hyperfine Structure

The proton in a hydrogen atom is a spin one-half charged particle, and therefore possesses a magnetic moment. By analogy with Eq. (10.58), we can write

$$\boldsymbol{\mu}_p = \frac{g_p\, e}{2\,m_p}\, \mathbf{S}_p, \tag{12.151}$$

where $\boldsymbol{\mu}_p$ is the proton magnetic moment, $\mathbf{S}_p$ is the proton spin, and the proton gyromagnetic ratio g_p is found experimentally to take that value 5.59. Note that the magnetic

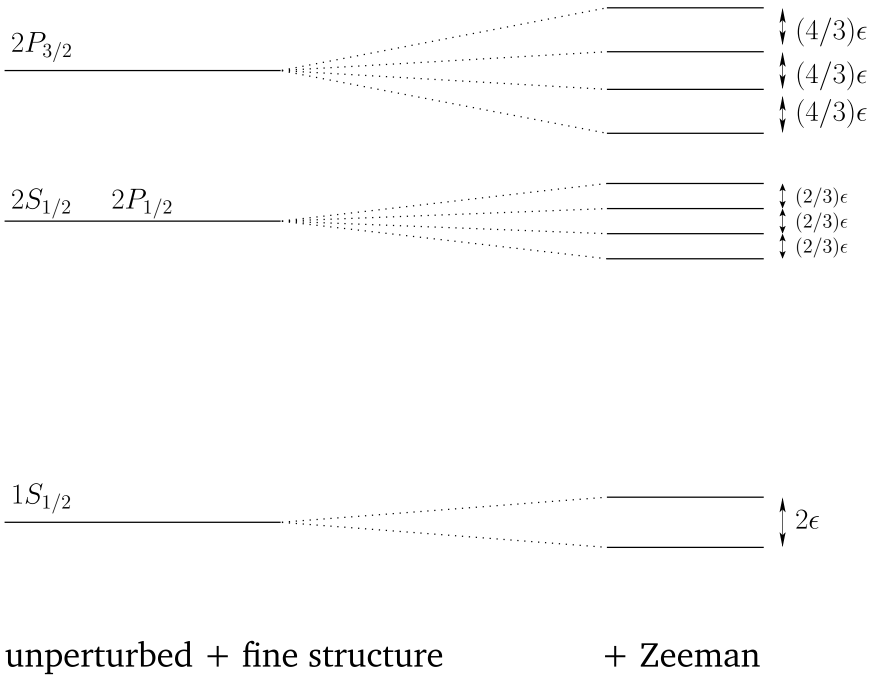

Figure 12.2: *The Zeeman effect for the* $n = 1$ *and* 2 *states of a hydrogen atom. Here,* $\epsilon = \mu_B\,B$. *Not to scale.*

moment of a proton is much smaller (by a factor of order m_e/m_p) than that of an electron. According to classical electromagnetism, the proton's magnetic moment generates a magnetic field of the form

$$\mathbf{B} = \frac{\mu_0}{4\pi\,r^3}\left[3\,(\boldsymbol{\mu}_p\cdot\mathbf{e}_r)\,\mathbf{e}_r - \boldsymbol{\mu}_p\right] + \frac{2\,\mu_0}{3}\,\boldsymbol{\mu}_p\,\delta^3(\mathbf{r}), \tag{12.152}$$

where $\mathbf{e}_r = \mathbf{r}/r$. We can understand the origin of the delta-function term in the above expression by thinking of the proton as a tiny current loop centred on the origin. All magnetic field-lines generated by the loop must pass through the loop. Hence, if the size of the loop goes to zero then the field will be infinite at the origin, and this contribution is what is reflected by the delta-function term. Now, the Hamiltonian of the electron in the magnetic field generated by the proton is simply

$$H_1 = -\boldsymbol{\mu}_e\cdot\mathbf{B}, \tag{12.153}$$

where

$$\boldsymbol{\mu}_e = -\frac{e}{m_e}\,\mathbf{S}_e. \tag{12.154}$$

Here, $\boldsymbol{\mu}_e$ is the electron magnetic moment [see Eqs. (10.58) and (10.59)], and $\mathbf{S}_e$ the electron spin. Thus, the perturbing Hamiltonian is written

$$H_1 = \frac{\mu_0\,g_p\,e^2}{8\pi\,m_p\,m_e}\,\frac{3\,(\mathbf{S}_p\cdot\mathbf{e}_r)\,(\mathbf{S}_e\cdot\mathbf{e}_r) - \mathbf{S}_p\cdot\mathbf{S}_e}{r^3} + \frac{\mu_0\,g_p\,e^2}{3\,m_p\,m_e}\,\mathbf{S}_p\cdot\mathbf{S}_e\,\delta^3(\mathbf{r}). \tag{12.155}$$

Note that, since we have neglected coupling between the proton spin and the magnetic field generated by the electron's orbital motion, the above expression is only valid for $l = 0$ states.

According to standard first-order perturbation theory, the energy-shift induced by spin-spin coupling between the proton and the electron is the expectation value of the perturbing Hamiltonian. Hence,

$$\Delta E = \frac{\mu_0\, g_p\, e^2}{8\pi\, m_p\, m_e} \left\langle \frac{3\,(\mathbf{S}_p\cdot\mathbf{e}_r)\,(\mathbf{S}_e\cdot\mathbf{e}_r) - \mathbf{S}_p\cdot\mathbf{S}_e}{r^3} \right\rangle + \frac{\mu_0\, g_p\, e^2}{3\, m_p\, m_e}\, \langle \mathbf{S}_p\cdot\mathbf{S}_e\rangle\, |\psi(0)|^2. \tag{12.156}$$

For the ground-state of hydrogen, which is spherically symmetric, the first term in the above expression vanishes by symmetry. Moreover, it is easily demonstrated that $|\psi_{000}(0)|^2 = 1/(\pi\, a_0^3)$. Thus, we obtain

$$\Delta E = \frac{\mu_0\, g_p\, e^2}{3\pi\, m_p\, m_e\, a_0^3}\, \langle \mathbf{S}_p\cdot\mathbf{S}_e\rangle. \tag{12.157}$$

Let

$$\mathbf{S} = \mathbf{S}_e + \mathbf{S}_p \tag{12.158}$$

be the total spin. We can show that

$$\mathbf{S}_p\cdot\mathbf{S}_e = \frac{1}{2}\,(S^2 - S_e^{\,2} - S_p^{\,2}). \tag{12.159}$$

Thus, the simultaneous eigenstates of the perturbing Hamiltonian and the main Hamiltonian are the simultaneous eigenstates of $S_e^{\,2}$, $S_p^{\,2}$, and S^2. However, both the proton and the electron are spin one-half particles. According to Sect. 11.4, when two spin one-half particles are combined (in the absence of orbital angular momentum) the net state has either spin 1 or spin 0. In fact, there are three spin 1 states, known as triplet states, and a single spin 0 state, known as the singlet state. For all states, the eigenvalues of $S_e^{\,2}$ and $S_p^{\,2}$ are $(3/4)\,\hbar^2$. The eigenvalue of S^2 is 0 for the singlet state, and $2\,\hbar^2$ for the triplet states. Hence,

$$\langle \mathbf{S}_p\cdot\mathbf{S}_e\rangle = -\frac{3}{4}\,\hbar^2 \tag{12.160}$$

for the singlet state, and

$$\langle \mathbf{S}_p\cdot\mathbf{S}_e\rangle = \frac{1}{4}\,\hbar^2 \tag{12.161}$$

for the triplet states.

It follows, from the above analysis, that spin-spin coupling breaks the degeneracy of the two $1S_{1/2}$ states in hydrogen, lifting the energy of the triplet configuration, and lowering that of the singlet. This splitting is known as *hyperfine structure*. The net energy difference between the singlet and the triplet states is

$$\Delta E = \frac{8}{3}\, g_p\, \frac{m_e}{m_p}\, \alpha^2\, E_0 = 5.88\times 10^{-6}\,\text{eV}, \tag{12.162}$$

where $E_0 = 13.6\,\text{eV}$ is the (magnitude of the) ground-state energy. Note that the hyperfine energy-shift is much smaller, by a factor m_e/m_p, than a typical fine structure energy-shift. If we convert the above energy into a wavelength then we obtain

$$\lambda = 21.1\,\text{cm}. \tag{12.163}$$

This is the wavelength of the radiation emitted by a hydrogen atom which is collisionally excited from the singlet to the triplet state, and then decays back to the lower energy singlet state. The 21 cm line is famous in radio astronomy because it was used to map out the spiral structure of our galaxy in the 1950's.

13 Time-Dependent Perturbation Theory

13.1 Introduction

Consider a system whose Hamiltonian can be written

$$H(t) = H_0 + H_1(t). \tag{13.1}$$

Here, H_0 is again a simple time-independent Hamiltonian whose eigenvalues and eigenstates are known exactly. However, H_1 now represents a small *time-dependent* external perturbation. Let the eigenstates of H_0 take the form

$$H_0\,\psi_m = E_m\,\psi_m. \tag{13.2}$$

We know (see Sect. 4.12) that if the system is in one of these eigenstates then, in the absence of an external perturbation, it remains in this state for ever. However, the presence of a small time-dependent perturbation can, in principle, give rise to a finite probability that if the system is initially in some eigenstate ψ_n of the unperturbed Hamiltonian then it is found in some other eigenstate at a subsequent time (since ψ_n is no longer an exact eigenstate of the total Hamiltonian). In other words, a time-dependent perturbation can cause the system to make transitions between its unperturbed energy eigenstates. Let us investigate this effect.

13.2 Preliminary Analysis

Suppose that at $t = 0$ the state of the system is represented by

$$\psi(0) = \sum_m c_m\,\psi_m, \tag{13.3}$$

where the c_m are complex numbers. Thus, the initial state is some linear superposition of the unperturbed energy eigenstates. In the absence of the time-dependent perturbation, the time evolution of the system is simply (see Sect. 4.12)

$$\psi(t) = \sum_m c_m \exp\left(-i\,E_m\,t/\hbar\right)\psi_m. \tag{13.4}$$

Now, the probability of finding the system in state n at time t is

$$P_n(t) = |\langle\psi_n|\psi\rangle|^2 = |c_n \exp\left(-i\,E_n\,t/\hbar\right)|^2 = |c_n|^2 = P_n(0), \tag{13.5}$$

since the unperturbed eigenstates are assummed to be orthonormal: *i.e.*,

$$\langle n|m\rangle = \delta_{nm}. \tag{13.6}$$

Clearly, with $H_1 = 0$, the probability of finding the system in state ψ_n at time t is exactly the same as the probability of finding the system in this state at the initial time, $t = 0$. However, with $H_1 \neq 0$, we expect P_n—and, hence, c_n—to vary with time. Thus, we can write

$$\psi(t) = \sum_m c_m(t)\, \exp\left(-i\, E_m\, t/\hbar\right) \psi_m, \tag{13.7}$$

where $P_n(t) = |c_n(t)|^2$. Here, we have carefully separated the fast phase oscillation of the eigenstates, which depends on the unperturbed Hamiltonian, from the slow variation of the amplitudes $c_n(t)$, which depends entirely on the perturbation (*i.e.*, c_n is constant in time if $H_1 = 0$). Note that in Eq. (13.7) the eigenstates ψ_m are *time-independent* (they are actually the eigenstates of H_0 evaluated at the initial time, $t = 0$).

The time-dependent Schrödinger equation [see Eq. (4.63)] yields

$$i\,\hbar\, \frac{\partial\psi(t)}{\partial t} = H(t)\, \psi(t) = [H_0 + H_1(t)]\, \psi(t). \tag{13.8}$$

Now, it follows from Eq. (13.7) that

$$(H_0 + H_1)\, \psi = \sum_m c_m\, \exp\left(-i\, E_m\, t/\hbar\right) (E_m + H_1)\, \psi_m. \tag{13.9}$$

We also have

$$i\,\hbar\, \frac{\partial\psi}{\partial t} = \sum_m \left(i\,\hbar\, \frac{dc_m}{dt} + c_m\, E_m\right) \exp\left(-i\, E_m\, t/\hbar\right) \psi_m, \tag{13.10}$$

since the ψ_m are time-independent. According to Eq. (13.8), we can equate the right-hand sides of the previous two equations to obtain

$$\sum_m i\,\hbar\, \frac{dc_m}{dt}\, \exp\left(-i\, E_m\, t/\hbar\right) \psi_m = \sum_m c_m\, exp\left(-i\, E_m\, t/\hbar\right) H_1\, \psi_m. \tag{13.11}$$

Projecting out the component of the above equation which is proportional to ψ_n, using Eq. (13.6), we obtain

$$i\,\hbar\, \frac{dc_n(t)}{dt} = \sum_m H_{nm}(t)\, \exp\left(i\, \omega_{nm}\, t\right) c_m(t), \tag{13.12}$$

where

$$H_{nm}(t) = \langle n|H_1(t)|m\rangle, \tag{13.13}$$

and

$$\omega_{nm} = \frac{E_n - E_m}{\hbar}. \tag{13.14}$$

Suppose that there are N linearly independent eigenstates of the unperturbed Hamiltonian. According to Eqs. (13.12), the time-dependence of the set of N coefficients c_n, which specify the probabilities of finding the system in these eigenstates at time t, is determined by N coupled first-order differential equations. Note that Eqs. (13.12) are exact—we have made no approximations at this stage. Unfortunately, we cannot generally find exact solutions to these equations. Instead, we have to obtain approximate solutions via suitable expansions in small quantities. However, for the particuilarly simple case of a two-state system (*i.e.*, $N = 2$), it is actually possible to solve Eqs. (13.12) without approximation. This solution is of great practical importance.

13.3 Two-State System

Consider a system in which the time-independent Hamiltonian possesses two eigenstates, denoted

$$H_0\,\psi_1 = E_1\,\psi_1, \tag{13.15}$$

$$H_0\,\psi_2 = E_2\,\psi_2. \tag{13.16}$$

Suppose, for the sake of simplicity, that the diagonal elements of the interaction Hamiltonian, H_1, are zero: *i.e.*,

$$\langle 1|H_1|1\rangle = \langle 2|H_1|2\rangle = 0. \tag{13.17}$$

The off-diagonal elements are assumed to oscillate sinusoidally at some frequency ω: *i.e.*,

$$\langle 1|H_1|2\rangle = \langle 2|H_1|1\rangle^{*} = \gamma\,\hbar\,\exp(i\,\omega\,t), \tag{13.18}$$

where γ and ω are real. Note that it is only the off-diagonal matrix elements which give rise to the effect which we are interested in—namely, transitions between states 1 and 2.

For a two-state system, Eq. (13.12) reduces to

$$i\,\frac{dc_1}{dt} = \gamma\,\exp\left[+i\,(\omega-\omega_{21})\,t\right]\,c_2, \tag{13.19}$$

$$i\,\frac{dc_2}{dt} = \gamma\,\exp\left[-i\,(\omega-\omega_{21})\,t\right]\,c_1, \tag{13.20}$$

where $\omega_{21} = (E_2 - E_1)/\hbar$. The above two equations can be combined to give a second-order differential equation for the time-variation of the amplitude c_2: *i.e.*,

$$\frac{d^2c_2}{dt^2} + i\,(\omega-\omega_{21})\,\frac{dc_2}{dt} + \gamma^2\,c_2 = 0. \tag{13.21}$$

Once we have solved for c_2, we can use Eq. (13.20) to obtain the amplitude c_1. Let us search for a solution in which the system is certain to be in state 1 (and, thus, has no chance of being in state 2) at time $t = 0$. Thus, our initial conditions are $c_1(0) = 1$ and $c_2(0) = 0$. It is easily demonstrated that the appropriate solutions to (13.21) and (13.20) are

$$c_2(t) = \left(\frac{-i\,\gamma}{\Omega}\right)\exp\left[\frac{-i\,(\omega-\omega_{21})\,t}{2}\right]\sin(\Omega\,t), \tag{13.22}$$

$$\begin{aligned}
c_1(t) &= \exp\left[\frac{i\,(\omega-\omega_{21})\,t}{2}\right]\cos(\Omega\,t) \\
&\quad -\left[\frac{i\,(\omega-\omega_{21})}{2\,\Omega}\right]\exp\left[\frac{i\,(\omega-\omega_{21})\,t}{2}\right]\sin(\Omega\,t),
\end{aligned} \tag{13.23}$$

where

$$\Omega=\sqrt{\gamma^2+(\omega-\omega_{21})^2/4}. \tag{13.24}$$

Now, the probability of finding the system in state 1 at time t is simply $P_1(t)=|c_1(t)|^2$. Likewise, the probability of finding the system in state 2 at time t is $P_2(t)=|c_2(t)|^2$. It follows that

$$P_1(t) = 1-P_2(t), \tag{13.25}$$

$$P_2(t) = \left[\frac{\gamma^2}{\gamma^2+(\omega-\omega_{21})^2/4}\right]\sin^2(\Omega\,t). \tag{13.26}$$

This result is known as *Rabi's formula*.

Equation (13.26) exhibits all the features of a classic *resonance*. At resonance, when the oscillation frequency of the perturbation, ω, matches the frequency ω_{21}, we find that

$$P_1(t) = \cos^2(\gamma\,t), \tag{13.27}$$

$$P_2(t) = \sin^2(\gamma\,t). \tag{13.28}$$

According to the above result, the system starts off in state 1 at $t=0$. After a time interval $\pi/(2\,\gamma)$ it is certain to be in state 2. After a further time interval $\pi/(2\,\gamma)$ it is certain to be in state 1 again, and so on. Thus, the system periodically flip-flops between states 1 and 2 under the influence of the time-dependent perturbation. This implies that the system alternatively absorbs and emits energy from the source of the perturbation.

The absorption-emission cycle also takes place away from the resonance, when $\omega\neq\omega_{21}$. However, the amplitude of the oscillation in the coefficient c_2 is reduced. This means that the maximum value of $P_2(t)$ is no longer unity, nor is the minimum of $P_1(t)$ zero. In fact, if we plot the maximum value of $P_2(t)$ as a function of the applied frequency, ω, we obtain a resonance curve whose maximum (unity) lies at the resonance, and whose full-width half-maximum (in frequency) is $4\,\gamma$. Thus, if the applied frequency differs from the resonant frequency by substantially more than $2\,\gamma$ then the probability of the system jumping from state 1 to state 2 is always very small. In other words, the time-dependent perturbation is only effective at causing transitions between states 1 and 2 if its frequency of oscillation lies in the approximate range $\omega_{21}\pm 2\,\gamma$. Clearly, the weaker the perturbation (*i.e.*, the smaller γ becomes), the narrower the resonance.

13.4 Spin Magnetic Resonance

Consider a system consisting of a spin one-half particle with no orbital angular momentum (*e.g.*, a bound electron) placed in a uniform z-directed magnetic field, and then subject to a small time-dependent magnetic field rotating in the x-y plane at the angular frequency ω. Thus,

$$\mathbf{B} = B_0\,\mathbf{e}_z + B_1\,[\cos(\omega\,t)\,\mathbf{e}_x + \sin(\omega\,t)\,\mathbf{e}_y]\,, \tag{13.29}$$

where B_0 and B_1 are constants, with $B_1 \ll B_0$. The rotating magnetic field usually represents the magnetic component of an electromagnetic wave propagating along the z-axis. In this system, the electric component of the wave has no effect. The Hamiltonian is written

$$H = -\boldsymbol{\mu}\cdot\mathbf{B} = H_0 + H_1, \tag{13.30}$$

where

$$H_0 = -\frac{g\,e\,B_0}{2\,m}\,S_z, \tag{13.31}$$

and

$$H_1 = -\frac{g\,e\,B_1}{2\,m}\,[\cos(\omega\,t)\,S_x + \sin(\omega\,t)\,S_y]\,. \tag{13.32}$$

Here, g and m are the gyromagnetic ratio [see Eq. (12.151)] and mass of the particle in question, respectively.

The eigenstates of the unperturbed Hamiltonian are the "spin up" and "spin down" states, denoted χ_+ and χ_-, respectively. Of course, these states are the eigenstates of S_z corresponding to the eigenvalues $+\hbar/2$ and $-\hbar/2$ respectively (see Sect. 10). Thus, we have

$$H_0\,\chi_\pm = \mp\frac{g\,e\,\hbar\,B_0}{4\,m}\,\chi_\pm. \tag{13.33}$$

The time-dependent Hamiltonian can be written

$$H_1 = -\frac{g\,e\,B_1}{4\,m}\,[\exp(\,i\,\omega\,t)\,S_- + \exp(-i\,\omega\,t)\,S_+]\,, \tag{13.34}$$

where S_+ and S_- are the conventional raising and lowering operators for spin angular momentum (see Sect. 10). It follows that

$$\langle +|H_1|+\rangle = \langle -|H_1|-\rangle = 0, \tag{13.35}$$

and

$$\langle -|H_1|+\rangle = \langle +|H_1|-\rangle^* = -\frac{g\,e\,B_1}{4\,m}\,\exp(\,i\,\omega\,t). \tag{13.36}$$

It can be seen that this system is exactly the same as the two-state system discussed in the previous subsection, provided that the make the following indentifications:

$$\psi_1 \;\rightarrow\; \chi_+, \tag{13.37}$$

$$\psi_2 \;\rightarrow\; \chi_-, \tag{13.38}$$

$$\omega_{21} \;\rightarrow\; \frac{g\,e\,B_0}{2\,m}, \tag{13.39}$$

$$\gamma \;\rightarrow\; -\frac{g\,e\,B_1}{4\,m}. \tag{13.40}$$

The resonant frequency, ω_{21}, is simply the spin precession frequency in a uniform magnetic field of strength B_0 (see Sect. 10.6). In the absence of the perturbation, the expectation values of S_x and S_y oscillate because of the spin precession, but the expectation value of S_z remains invariant. If we now apply a magnetic perturbation rotating at the resonant frequency then, according to the analysis of the previous subsection, the system undergoes a succession of spin flips, $\chi_+ \leftrightarrow \chi_-$, in addition to the spin precession. We also know that if the oscillation frequency of the applied field is very different from the resonant frequency then there is virtually zero probability of the field triggering a spin flip. The width of the resonance (in frequency) is determined by the strength of the oscillating magnetic perturbation. Experimentalists are able to measure the gyromagnetic ratios of spin one-half particles to a high degree of accuracy by placing the particles in a uniform magnetic field of known strength, and then subjecting them to an oscillating magnetic field whose frequency is gradually scanned. By determining the resonant frequency (*i.e.*, the frequency at which the particles absorb energy from the oscillating field), it is possible to determine the gyromagnetic ratio (assuming that the mass is known).

13.5 Perturbation Expansion

Let us recall the analysis of Sect. 13.2. The ψ_n are the stationary orthonormal eigenstates of the time-independent unperturbed Hamiltonian, H_0. Thus, $H_0\,\psi_n = E_n\,\psi_n$, where the E_n are the unperturbed energy levels, and $\langle n|m\rangle = \delta_{nm}$. Now, in the presence of a small time-dependent perturbation to the Hamiltonian, $H_1(t)$, the wavefunction of the system takes the form

$$\psi(t) = \sum_n c_n(t)\,\exp(-i\,\omega_n\,t)\,\psi_n, \tag{13.41}$$

where $\omega_n = E_n/\hbar$. The amplitudes $c_n(t)$ satisfy

$$i\,\hbar\,\frac{dc_n}{dt} = \sum_m H_{nm}\,\exp(\,i\,\omega_{nm}\,t)\,c_m, \tag{13.42}$$

where $H_{nm}(t) = \langle n|H_1(t)|m\rangle$ and $\omega_{nm} = (E_n - E_m)/\hbar$. Finally, the probability of finding the system in the nth eigenstate at time t is simply

$$P_n(t) = |c_n(t)|^2 \tag{13.43}$$

(assuing that, initially, $\sum_n |c_n|^2 = 1$).

Suppose that at $t = 0$ the system is in some initial energy eigenstate labeled i. Equation (13.42) is, thus, subject to the initial condition

$$c_n(0) = \delta_{ni}. \tag{13.44}$$

Let us attempt a perturbative solution of Eq. (13.42) using the ratio of H_1 to H_0 (or H_{nm} to $\hbar\,\omega_{nm}$, to be more exact) as our expansion parameter. Now, according to (13.42), the c_n

are constant in time in the absence of the perturbation. Hence, the zeroth-order solution is simply

$$c_n^{(0)}(t) = \delta_{ni}. \tag{13.45}$$

The first-order solution is obtained, via iteration, by substituting the zeroth-order solution into the right-hand side of Eq. (13.42). Thus, we obtain

$$i\,\hbar\,\frac{dc_n^{(1)}}{dt} = \sum_m H_{nm}\,\exp(\,i\,\omega_{nm}\,t)\,c_m^{(0)} = H_{ni}\,\exp(\,i\,\omega_{ni}\,t), \tag{13.46}$$

subject to the boundary condition $c_n^{(1)}(0) = 0$. The solution to the above equation is

$$c_n^{(1)} = -\frac{i}{\hbar}\int_0^t H_{ni}(t')\,\exp(\,i\,\omega_{ni}\,t')\,dt'. \tag{13.47}$$

It follows that, up to first-order in our perturbation expansion,

$$c_n(t) = \delta_{ni} - \frac{i}{\hbar}\int_0^t H_{ni}(t')\,\exp(\,i\,\omega_{ni}\,t')\,dt'. \tag{13.48}$$

Hence, the probability of finding the system in some final energy eigenstate labeled f at time t, given that it is definitely in a different initial energy eigenstate labeled i at time $t = 0$, is

$$P_{i\to f}(t) = |c_f(t)|^2 = \left|-\frac{i}{\hbar}\int_0^t H_{fi}(t')\,\exp(\,i\,\omega_{fi}\,t')\,dt'\right|^2. \tag{13.49}$$

Note, finally, that our perturbative solution is clearly only valid provided

$$P_{i\to f}(t) \ll 1. \tag{13.50}$$

13.6 Harmonic Perturbations

Consider a (Hermitian) perturbation which oscillates sinusoidally in time. This is usually termed a *harmonic perturbation*. Such a perturbation takes the form

$$H_1(t) = V\,\exp(\,i\,\omega\,t) + V^\dagger\,\exp(-i\,\omega\,t), \tag{13.51}$$

where V is, in general, a function of position, momentum, and spin operators.

It follows from Eqs. (13.48) and (13.51) that, to first-order,

$$c_f(t) = -\frac{i}{\hbar}\int_0^t \left[V_{fi}\,\exp(\,i\,\omega\,t') + V_{fi}^\dagger\,\exp(-i\,\omega\,t')\right]\exp(\,i\,\omega_{fi}\,t')\,dt', \tag{13.52}$$

where

$$V_{fi} \;=\; \langle f|V|i\rangle, \tag{13.53}$$

$$V_{fi}^\dagger \;=\; \langle f|V^\dagger|i\rangle = \langle i|V|f\rangle^*. \tag{13.54}$$

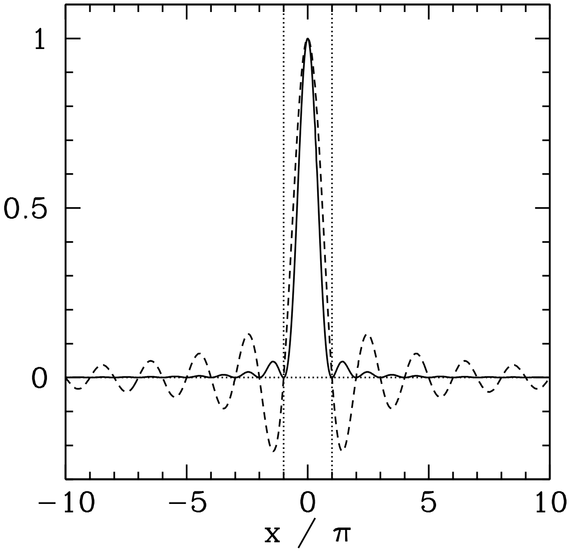

Figure 13.1: *The functions* sinc(x) *(dashed curve) and* sinc^2(x) *(solid curve). The vertical dotted lines denote the region* $|x| \leq \pi$.

Integration with respect to t' yields

$$\begin{aligned} c_f(t) &= -\frac{i\,t}{\hbar}\left(V_{fi}\,\exp\left[i\,(\omega+\omega_{fi})\,t/2\right]\text{sinc}\left[(\omega+\omega_{fi})\,t/2\right]\right. \\ &\quad \left.+V_{fi}^{\dagger}\,\exp\left[-i\,(\omega-\omega_{fi})\,t/2\right]\text{sinc}\left[(\omega-\omega_{fi})\,t/2\right]\right), \end{aligned} \tag{13.55}$$

where

$$\text{sinc}\,x \equiv \frac{\sin\,x}{x}. \tag{13.56}$$

Now, the function sinc(x) takes its largest values when $|x| \lesssim \pi$, and is fairly negligible when $|x| \gg \pi$ (see Fig. 13.1). Thus, the first and second terms on the right-hand side of Eq. (13.55) are only non-negligible when

$$|\omega+\omega_{fi}| \lesssim \frac{2\pi}{t}, \tag{13.57}$$

and

$$|\omega-\omega_{fi}| \lesssim \frac{2\pi}{t}, \tag{13.58}$$

respectively. Clearly, as t increases, the ranges in ω over which these two terms are non-negligible gradually shrink in size. Eventually, when $t \gg 2\pi/|\omega_{fi}|$, these two ranges become strongly non-overlapping. Hence, in this limit, $P_{i\rightarrow f} = |c_f|^2$ yields

$$P_{i\rightarrow f}(t) = \frac{t^2}{\hbar^2}\left\{|V_{fi}|^2\,\text{sinc}^2\left[(\omega+\omega_{fi})\,t/2\right] + |V_{fi}^{\dagger}|^2\,\text{sinc}^2\left[(\omega-\omega_{fi})\,t/2\right]\right\}. \tag{13.59}$$

Now, the function $\text{sinc}^2(x)$ is very strongly peaked at $x = 0$, and is completely negligible for $|x| \gtrsim \pi$ (see Fig. 13.1). It follows that the above expression exhibits a *resonant response* to the applied perturbation at the frequencies $\omega = \pm\omega_{fi}$. Moreover, the widths of these resonances decease linearly as time increases. At each of the resonances (*i.e.*, at $\omega = \pm\omega_{fi}$), the transition probability $P_{i\to f}(t)$ varies as t^2 [since $\sinh(0) = 1$]. This behaviour is entirely consistent with our earlier result (13.28), for the two-state system, in the limit $\gamma\, t \ll 1$ (recall that our perturbative solution is only valid as long as $P_{i\to f} \ll 1$).

The resonance at $\omega = -\omega_{fi}$ corresponds to

$$E_f - E_i = -\hbar\,\omega. \tag{13.60}$$

This implies that the system *loses* energy $\hbar\,\omega$ to the perturbing field, whilst making a transition to a final state whose energy is *less* than the initial state by $\hbar\,\omega$. This process is known as *stimulated emission*. The resonance at $\omega = \omega_{fi}$ corresponds to

$$E_f - E_i = \hbar\,\omega. \tag{13.61}$$

This implies that the system *gains* energy $\hbar\,\omega$ from the perturbing field, whilst making a transition to a final state whose energy is *greater* than that of the initial state by $\hbar\,\omega$. This process is known as *absorption*.

Stimulated emission and absorption are mutually exclusive processes, since the first requires $\omega_{fi} < 0$, whereas the second requires $\omega_{fi} > 0$. Hence, we can write the transition probabilities for both processes separately. Thus, from (13.59), the transition probability for stimulated emission is

$$P^{stm}_{i\to f}(t) = \frac{t^2}{\hbar^2}\,|V^\dagger_{if}|^{\,2}\,\text{sinc}^2\left[(\omega - \omega_{if})\,t/2\right], \tag{13.62}$$

where we have made use of the facts that $\omega_{if} = -\omega_{fi} > 0$, and $|V_{fi}|^2 = |V^\dagger_{if}|^2$. Likewise, the transition probability for absorption is

$$P^{abs}_{i\to f}(t) = \frac{t^2}{\hbar^2}\,|V^\dagger_{fi}|^{\,2}\,\text{sinc}^2\left[(\omega - \omega_{fi})\,t/2\right]. \tag{13.63}$$

13.7 Electromagnetic Radiation

Let us use the above results to investigate the interaction of an atomic electron with classical (*i.e.*, non-quantized) electromagnetic radiation.

The unperturbed Hamiltonian of the system is

$$H_0 = \frac{p^2}{2\,m_e} + V_0(r). \tag{13.64}$$

Now, the standard classical prescription for obtaining the Hamiltonian of a particle of charge q in the presence of an electromagnetic field is

$$\mathbf{p} \;\rightarrow\; \mathbf{p} + q\,\mathbf{A}, \tag{13.65}$$

$$H \;\rightarrow\; H - q\,\phi, \tag{13.66}$$

where $\mathbf{A}(\mathbf{r})$ is the vector potential, and $\phi(\mathbf{r})$ the scalar potential. Note that

$$\mathbf{E} = -\nabla\phi - \frac{\partial\mathbf{A}}{\partial t}, \tag{13.67}$$

$$\mathbf{B} = \nabla\times\mathbf{A}. \tag{13.68}$$

This prescription also works in quantum mechanics. Thus, the Hamiltonian of an atomic electron placed in an electromagnetic field is

$$H = \frac{(\mathbf{p} - e\,\mathbf{A})^2}{2\,m_e} + e\,\phi + V_0(r), \tag{13.69}$$

where $\mathbf{A}$ and ϕ are functions of the position operators. The above equation can be written

$$H = \frac{\left(p^2 - e\,\mathbf{A}\cdot\mathbf{p} - e\,\mathbf{p}\cdot\mathbf{A} + e^2A^2\right)}{2\,m_e} + e\,\phi + V_0(r). \tag{13.70}$$

Now,

$$\mathbf{p}\cdot\mathbf{A} = \mathbf{A}\cdot\mathbf{p}, \tag{13.71}$$

provided that we adopt the gauge $\nabla\cdot\mathbf{A} = 0$. Hence,

$$H = \frac{p^2}{2\,m_e} - \frac{e\,\mathbf{A}\cdot\mathbf{p}}{m_e} + \frac{e^2A^2}{2\,m_e} + e\,\phi + V_0(r). \tag{13.72}$$

Suppose that the perturbation corresponds to a linearly polarized, monochromatic, plane-wave. In this case,

$$\phi = 0, \tag{13.73}$$

$$\mathbf{A} = A_0\,\boldsymbol{\epsilon}\,\cos(\mathbf{k}\cdot\mathbf{r} - \omega t)\,, \tag{13.74}$$

where $\mathbf{k}$ is the wavevector (note that $\omega = k\,c$), and $\boldsymbol{\epsilon}$ a unit vector which specifies the direction of polarization (*i.e.*, the direction of $\mathbf{E}$). Note that $\boldsymbol{\epsilon}\cdot\mathbf{k} = 0$. The Hamiltonian becomes

$$H = H_0 + H_1(t), \tag{13.75}$$

with

$$H_0 = \frac{p^2}{2\,m_e} + V_0(r), \tag{13.76}$$

and

$$H_1 \simeq -\frac{e\,\mathbf{A}\cdot\mathbf{p}}{m_e}, \tag{13.77}$$

where the A^2 term, which is second order in A_0, has been neglected.

The perturbing Hamiltonian can be written

$$H_1 = -\frac{e\,A_0\,\boldsymbol{\epsilon}\cdot\mathbf{p}}{2\,m_e}\left[\exp(\,i\,\mathbf{k}\cdot\mathbf{r} - i\,\omega t) + \exp(-i\,\mathbf{k}\cdot\mathbf{r} + i\,\omega t)\right]. \tag{13.78}$$

This has the same form as Eq. (13.51), provided that

$$V^\dagger = -\frac{e\,A_0\,\boldsymbol{\epsilon}\cdot\mathbf{p}}{2\,m_e}\,\exp(\,i\,\mathbf{k}\cdot\mathbf{r}\,). \tag{13.79}$$

It follows from Eqs. (13.53), (13.63), and (13.79) that the transition probability for radiation induced absorption is

$$P^{abs}_{i\to f}(t) = \frac{t^2}{\hbar^2}\,\frac{e^2\,|A_0|^2}{4\,m_e^2}\,|\langle f|\boldsymbol{\epsilon}\cdot\mathbf{p}\,\exp(\,i\,\mathbf{k}\cdot\mathbf{r})|i\rangle|^2\,\mathrm{sinc}^2[(\omega-\omega_{fi})\,t/2]. \tag{13.80}$$

Now, the mean energy density of an electromagnetic wave is

$$u = \frac{1}{2}\left(\frac{\epsilon_0\,|E_0|^2}{2} + \frac{|B_0|^2}{2\,\mu_0}\right) = \frac{1}{2}\,\epsilon_0\,|E_0|^2, \tag{13.81}$$

where $E_0 = A_0\,\omega$ and $B_0 = E_0/c$ are the peak electric and magnetic field-strengths, respectively. It thus follows that

$$P^{abs}_{i\to f}(t) = \frac{t^2\,e^2}{2\,\epsilon_0\,\hbar^2\,m_e^2\,\omega^2}\,|\langle f|\boldsymbol{\epsilon}\cdot\mathbf{p}\,\exp(\,i\,\mathbf{k}\cdot\mathbf{r})|i\rangle|^2\,u\,\mathrm{sinc}^2[(\omega-\omega_{fi})\,t/2]. \tag{13.82}$$

Thus, not surprisingly, the transition probability for radiation induced absorption (or stimulated emission) is directly proportional to the *energy density* of the incident radiation.

Suppose that the incident radiation is not monochromatic, but instead extends over a range of frequencies. We can write

$$u = \int_{-\infty}^{\infty}\rho(\omega)\,d\omega, \tag{13.83}$$

where $\rho(\omega)\,d\omega$ is the energy density of radiation whose frequencies lie between ω and $\omega + d\omega$. Equation (13.82) generalizes to

$$P^{abs}_{i\to f}(t) = \int_{-\infty}^{\infty}\frac{t^2\,e^2}{2\,\epsilon_0\,\hbar^2\,m_e^2\,\omega^2}\,|\langle f|\boldsymbol{\epsilon}\cdot\mathbf{p}\,\exp(\,i\,\mathbf{k}\cdot\mathbf{r})|i\rangle|^2\,\rho(\omega)\,\mathrm{sinc}^2[(\omega-\omega_{fi})\,t/2]\,d\omega. \tag{13.84}$$

Note, however, that the above expression is only valid provided the radiation in question is *incoherent*: *i.e.*, there are no phase correlations between waves of different frequencies. This follows because it is permissible to add the *intensities* of incoherent radiation, whereas we must always add the *amplitudes* of coherent radiation. Given that the function $\mathrm{sinc}^2[(\omega-\omega_{fi})\,t/2]$ is very strongly peaked (see Fig. 13.1) about $\omega = \omega_{fi}$ (assuming that $t \gg 2\pi/\omega_{fi}$), and

$$\int_{-\infty}^{\infty}\mathrm{sinc}^2(x)\,dx = \pi, \tag{13.85}$$

the above equation reduces to

$$P^{abs}_{i\to f}(t) = \frac{\pi\,e^2\,\rho(\omega_{fi})}{\epsilon_0\,\hbar^2\,m_e^2\,\omega_{fi}^2}\,|\langle f|\boldsymbol{\epsilon}\cdot\mathbf{p}\,\exp(\,i\,\mathbf{k}\cdot\mathbf{r})|i\rangle|^2\,t. \tag{13.86}$$

Note that in integrating over the frequencies of the incoherent radiation we have transformed a transition probability which is basically proportional to t^2 [see Eq. (13.82)] to one which is proportional to t. As has already been explained, the above expression is only valid when $P^{abs}_{i\to f} \ll 1$. However, the result that

$$w^{abs}_{i\to f} \equiv \frac{dP^{abs}_{i\to f}}{dt} = \frac{\pi\, e^2\, \rho(\omega_{fi})}{\epsilon_0\, \hbar^2\, m_e^2\, \omega_{fi}^2}\, |\langle f|\boldsymbol{\epsilon}\cdot\mathbf{p}\, \exp(\,i\,\mathbf{k}\cdot\mathbf{r})|i\rangle|^2 \tag{13.87}$$

is *constant* in time is universally valid. Here, $w^{abs}_{i\to f}$ is the transition probability per unit time interval, otherwise known as the *transition rate*. Given that the transition rate is constant, we can write (see Cha. 2)

$$P^{abs}_{i\to f}(t+dt) - P^{abs}_{i\to f}(t) = \left[1 - P^{abs}_{i\to f}(t)\right]\, w^{abs}_{i\to f}\, dt: \tag{13.88}$$

i.e., the probability that the system makes a transition from state i to state f between times t and $t + dt$ is equivalent to the probability that the system does not make a transition between times 0 and t *and* then makes a transition in a time interval dt—the probabilities of these two events are $1 - P^{abs}_{i\to f}(t)$ and $w^{abs}_{i\to f}\, dt$, respectively. It follows that

$$\frac{dP^{abs}_{i\to f}}{dt} + w^{abs}_{i\to f}\, P^{abs}_{i\to f} = w^{abs}_{i\to f}, \tag{13.89}$$

with the initial condition $P^{abs}_{i\to f}(0) = 0$. The above equation can be solved to give

$$P^{abs}_{i\to f}(t) = 1 - \exp\left(-w^{abs}_{i\to f}\, t\right). \tag{13.90}$$

This result is consistent with Eq. (13.86) provided $w^{abs}_{i\to f}\, t \ll 1$: *i.e.*, provided that $P^{abs}_{i\to f} \ll 1$.

Using similar arguments to the above, the transition probability for stimulated emission can be shown to take the form

$$P^{stm}_{i\to f}(t) = 1 - \exp\left(-w^{stm}_{i\to f}\, t\right), \tag{13.91}$$

where the corresponding transition rate is written

$$w^{stm}_{i\to f} = \frac{\pi\, e^2\, \rho(\omega_{if})}{\epsilon_0\, \hbar^2\, m_e^2\, \omega_{if}^2}\, |\langle i|\boldsymbol{\epsilon}\cdot\mathbf{p}\, \exp(\,i\,\mathbf{k}\cdot\mathbf{r})|f\rangle|^2. \tag{13.92}$$

13.8 Electric Dipole Approximation

In general, the wavelength of the type of electromagnetic radiation which induces, or is emitted during, transitions between different atomic energy levels is much larger than the typical size of an atom. Thus,

$$\exp(\,i\,\mathbf{k}\cdot\mathbf{r}) = 1 + i\,\mathbf{k}\cdot\mathbf{r} + \cdots, \tag{13.93}$$

can be approximated by its first term, unity. This approach is known as the *electric dipole approximation*. It follows that

$$\langle f|\boldsymbol{\epsilon}\cdot\mathbf{p}\,\exp(\,i\,\mathbf{k}\cdot\mathbf{r})|i\rangle \simeq \boldsymbol{\epsilon}\cdot\langle f|\mathbf{p}|i\rangle. \tag{13.94}$$

Now, it is readily demonstrated that

$$[\mathbf{r}, H_0] = \frac{i\,\hbar\,\mathbf{p}}{m_e}, \tag{13.95}$$

so

$$\langle f|\mathbf{p}|i\rangle = -i\,\frac{m_e}{\hbar}\,\langle f|[\mathbf{r}, H_0]|i\rangle = i\,m_e\,\omega_{fi}\,\langle f|\mathbf{r}|i\rangle. \tag{13.96}$$

Thus, our previous expressions for the transition rates for radiation induced absorption and stimulated emission reduce to

$$w^{abs}_{i\rightarrow f} = \frac{\pi}{\epsilon_0\,\hbar^2}\,|\boldsymbol{\epsilon}\cdot\mathbf{d}_{if}|^2\,\rho(\omega_{fi}), \tag{13.97}$$

$$w^{stm}_{i\rightarrow f} = \frac{\pi}{\epsilon_0\,\hbar^2}\,|\boldsymbol{\epsilon}\cdot\mathbf{d}_{if}|^2\,\rho(\omega_{if}), \tag{13.98}$$

respectively. Here,

$$\mathbf{d}_{if} = \langle f|e\,\mathbf{r}|i\rangle \tag{13.99}$$

is the effective electric dipole moment of the atom when making a transition from state i to state f.

Equations (13.97) and (13.98) give the transition rates for absorption and stimulated emission, respectively, induced by a linearly polarized plane-wave. Actually, we are more interested in the transition rates induced by *unpolarized isotropic* radiation. To obtain these we must average Eqs. (13.97) and (13.98) over all possible polarizations and propagation directions of the wave. To facilitate this process, we can define a set of Cartesian coordinates such that the wavevector $\mathbf{k}$, which specifies the direction of wave propagation, points along the z-axis, and the vector $\mathbf{d}_{if}$, which specifies the direction of the atomic dipole moment, lies in the x-z plane. It follows that the vector $\boldsymbol{\epsilon}$, which specifies the direction of wave polarization, must lie in the x-y plane, since it has to be orthogonal to $\mathbf{k}$. Thus, we can write

$$\mathbf{k} = (0,\,0,\,k), \tag{13.100}$$

$$\mathbf{d}_{if} = (d_{if}\,\sin\theta,\,0,\,d_{if}\,\cos\theta), \tag{13.101}$$

$$\boldsymbol{\epsilon} = (\cos\phi,\,\sin\phi,\,0), \tag{13.102}$$

which implies that

$$|\boldsymbol{\epsilon}\cdot\mathbf{d}_{if}|^2 = d_{if}^2\,\sin^2\theta\,\cos^2\phi. \tag{13.103}$$

We must now average the above quantity over all possible values of θ and ϕ. Thus,

$$\left\langle|\boldsymbol{\epsilon}\cdot\mathbf{d}_{if}|^2\right\rangle_{av} = d_{if}^2\,\frac{\int\!\int\sin^2\theta\,\cos^2\phi\,d\Omega}{4\pi}, \tag{13.104}$$

where $d\Omega = \sin\theta\,d\theta\,d\phi$, and the integral is taken over all solid angle. It is easily demonstrated that

$$\left\langle |\boldsymbol{\epsilon}\cdot\mathbf{d}_{if}|^2\right\rangle_{av} = \frac{d_{if}^2}{3}. \tag{13.105}$$

Here, d_{if}^2 stands for

$$d_{if}^2 = |\langle f|e\,x|i\rangle|^2 + |\langle f|e\,y|i\rangle|^2 + |\langle f|e\,z|i\rangle|^2. \tag{13.106}$$

Hence, the transition rates for absorption and stimulated emission induced by unpolarized isotropic radiation are

$$w_{i\to f}^{abs} = \frac{\pi}{3\,\epsilon_0\,\hbar^2}\,d_{if}^2\,\rho(\omega_{fi}), \tag{13.107}$$

$$w_{i\to f}^{stm} = \frac{\pi}{3\,\epsilon_0\,\hbar^2}\,d_{if}^2\,\rho(\omega_{if}), \tag{13.108}$$

respectively.

13.9 Spontaneous Emission

So far, we have calculated the rates of *radiation induced* transitions between two atomic states. This process is known as *absorption* when the energy of the final state exceeds that of the initial state, and *stimulated emission* when the energy of the final state is less than that of the initial state. Now, in the absence of any external radiation, we would not expect an atom in a given state to spontaneously jump into an state with a higher energy. On the other hand, it should be possible for such an atom to spontaneously jump into an state with a lower energy via the emission of a photon whose energy is equal to the difference between the energies of the initial and final states. This process is known as *spontaneous emission*.

It is possible to derive the rate of spontaneous emission between two atomic states from a knowledge of the corresponding absorption and stimulated emission rates using a famous thermodynamic argument due to Einstein. Consider a very large ensemble of similar atoms placed inside a closed cavity whose walls (which are assumed to be perfect emitters and absorbers of radiation) are held at the constant temperature T. Let the system have attained thermal equilibrium. According to statistical thermodynamics, the cavity is filled with so-called "black-body" electromagnetic radiation whose energy spectrum is

$$\rho(\omega) = \frac{\hbar}{\pi^2 c^3}\,\frac{\omega^3}{\exp(\hbar\,\omega/k_B\,T) - 1}, \tag{13.109}$$

where k_B is the Boltzmann constant. This well-known result was first obtained by Max Planck in 1900.

Consider two atomic states, labeled i and f, with $E_i > E_f$. One of the tenants of statistical thermodynamics is that in thermal equilibrium we have so-called *detailed balance*. This means that, irrespective of any other atomic states, the rate at which atoms in the ensemble leave state i due to transitions to state f is exactly balanced by the rate at which atoms enter state i due to transitions from state f. The former rate (*i.e.*, number of transitions per unit time in the ensemble) is written

$$W_{i\rightarrow f} = N_i\,(w^{spn}_{i\rightarrow f} + w^{stm}_{i\rightarrow f}), \tag{13.110}$$

where $w^{spn}_{i\rightarrow f}$ is the rate of spontaneous emission (for a single atom) between states i and f, and N_i is the number of atoms in the ensemble in state i. Likewise, the latter rate takes the form

$$W_{f\rightarrow i} = N_f\,w^{abs}_{f\rightarrow i}, \tag{13.111}$$

where N_f is the number of atoms in the ensemble in state f. The above expressions describe how atoms in the ensemble make transitions from state i to state f due to a combination of spontaneous and stimulated emission, and make the opposite transition as a consequence of absorption. In thermal equilibrium, we have $W_{i\rightarrow f} = W_{f\rightarrow i}$, which gives

$$w^{spn}_{i\rightarrow f} = \frac{N_f}{N_i}\,w^{abs}_{f\rightarrow i} - w^{stm}_{i\rightarrow f}. \tag{13.112}$$

According to Eqs. (13.107) and (13.108), we can also write

$$w^{spn}_{i\rightarrow f} = \left(\frac{N_f}{N_i} - 1\right)\frac{\pi}{3\,\epsilon_0\,\hbar^2}\,d_{if}^{\,2}\,\rho(\omega_{if}). \tag{13.113}$$

Now, another famous result in statistical thermodynamics is that in thermal equilibrium the number of atoms in an ensemble occupying a state of energy E is proportional to $\exp(-E/k_B\,T)$. This implies that

$$\frac{N_f}{N_i} = \frac{\exp(-E_f/k_B\,T)}{\exp(-E_i/k_B\,T)} = \exp(\,\hbar\,\omega_{if}/k_B\,T). \tag{13.114}$$

Thus, it follows from Eq. (13.109), (13.113), and (13.114) that the rate of spontaneous emission between states i and f takes the form

$$w^{spn}_{i\rightarrow f} = \frac{\omega_{if}^{3}\,d_{if}^{\,2}}{3\pi\,\epsilon_0\,\hbar\,c^3}. \tag{13.115}$$

Note, that, although the above result has been derived for an atom in a radiation-filled cavity, it remains correct even in the absence of radiation. Finally, the corresponding absorption and stimulated emission rates for an atom in a radiation-filled cavity are

$$w^{abs}_{i\rightarrow f} = \frac{\omega_{fi}^{3}\,d_{if}^{\,2}}{3\pi\,\epsilon_0\,\hbar\,c^3}\,\frac{1}{\exp(\hbar\,\omega_{fi}/k_B\,T) - 1}, \tag{13.116}$$

$$w^{stm}_{i\rightarrow f} = \frac{\omega_{if}^{3}\,d_{if}^{\,2}}{3\pi\,\epsilon_0\,\hbar\,c^3}\,\frac{1}{\exp(\hbar\,\omega_{if}/k_B\,T) - 1}, \tag{13.117}$$

respectively.

Let us estimate the typical value of the spontaneous emission rate for a hydrogen atom. We expect the dipole moment d_{if} to be of order $e\,a_0$, where a_0 is the Bohr radius [see Eq. (9.58)]. We also expect ω_{if} to be of order $|E_0|/\hbar$, where E_0 is the energy of the ground-state [see Eq. (9.57)]. It thus follows from Eq. (13.115) that

$$w_{i\to f}^{spn} \sim \alpha^3\,\omega_{if}, \tag{13.118}$$

where $\alpha = e^2/(4\pi\,\epsilon_0\,\hbar\,c) \simeq 1/137$ is the fine-structure constant. This is an important result, since our perturbation expansion is based on the assumption that the transition rate between different energy eigenstates is much slower than the frequency of phase oscillation of these states: *i.e.*, that $w_{i\to f}^{spn} \ll \omega_{if}$ (see Sect. 13.2). This is indeed the case.

13.10 Radiation from a Harmonic Oscillator

Consider an electron in a one-dimensional harmonic oscillator potential aligned along the x-axis. According to Sect. 5.8, the unperturbed energy eigenvalues of the system are

$$E_n = (n+1/2)\,\hbar\,\omega_0, \tag{13.119}$$

where ω_0 is the frequency of the corresponding classical oscillator. Here, the quantum number n takes the values $0, 1, 2, \cdots$. Let the $\psi_n(x)$ be the (real) properly normalized unperturbed eigenstates of the system.

Suppose that the electron is initially in an excited state: *i.e.*, $n > 0$. In principle, the electron can decay to a lower energy state via the spontaneous emission of a photon of the appropriate frequency. Let us investigate this effect. Now, according to Eq. (13.115), the system can only make a spontaneous transition from an energy state corresponding to the quantum number n to one corresponding to the quantum number n' if the associated electric dipole moment

$$(d_x)_{n,n'} = \langle n|e\,x|n'\rangle = e\int_{-\infty}^{\infty}\psi_n(x)\,x\,\psi_{n'}(x)\,dx \tag{13.120}$$

is non-zero [since $d_{if} \equiv (d_x)_{n,n'}^{\,2}$ for the case in hand]. However, according to Eq. (5.117),

$$\int_{-\infty}^{\infty}\psi_n\,x\,\psi_{n'}\,dx = \sqrt{\frac{\hbar}{2\,m_e\,\omega_0}}\left(\sqrt{n}\,\delta_{n,n'+1} + \sqrt{n'}\,\delta_{n,n'-1}\right). \tag{13.121}$$

Since we are dealing with emission, we must have $n > n'$. Hence, we obtain

$$(d_x)_{n,n'} = e\sqrt{\frac{\hbar\,n}{2\,m_e\,\omega_0}}\,\delta_{n,n'+1}. \tag{13.122}$$

It is clear that (in the electric dipole approximation) we can only have spontaneous emission between states whose quantum numbers differ by *unity*. Thus, the frequency of the photon emitted when the nth excited state decays is

$$\omega_{n,n-1} = \frac{E_n - E_{n-1}}{\hbar} = \omega_0. \tag{13.123}$$

Hence, we conclude that, no matter which state decays, the emitted photon always has the same frequency as the classical oscillator.

According to Eq. (13.115), the decay rate of the nth excited state is given by

$$w_n = \frac{\omega_{n,n-1}^3\,(d_x)_{n,n-1}^2}{3\pi\,\epsilon_0\,\hbar\,c^3}. \tag{13.124}$$

It follows that

$$w_n = \frac{n\,e^2\,\omega_0^2}{6\pi\,\epsilon_0\,m_e\,c^3}. \tag{13.125}$$

The mean radiated power is simply

$$P_n = \hbar\,\omega_0\,w_n = \frac{e^2\,\omega_0^2}{6\pi\,\epsilon_0\,m_e\,c^3}\,[E_n - (1/2)\,\hbar\,\omega_0]. \tag{13.126}$$

Classically, an electron in a one-dimensional oscillator potential radiates at the oscillation frequency ω_0 with the mean power

$$P = \frac{e^2\,\omega_0^2}{6\pi\,\epsilon_0\,m_e\,c^3}\,E, \tag{13.127}$$

where E is the oscillator energy. It can be seen that a quantum oscillator radiates in an almost exactly analogous manner to the equivalent classical oscillator. The only difference is the factor $(1/2)\,\hbar\,\omega_0$ in Eq. (13.126)—this is needed to ensure that the ground-state of the quantum oscillator does not radiate.

13.11 Selection Rules

Let us now consider spontaneous transitions between the different energy levels of a hydrogen atom. Since the perturbing Hamiltonian (13.77) does not contain any spin operators, we can neglect electron spin in our analysis. Thus, according to Sect. 9.4, the various energy eigenstates of the hydrogen atom are labeled by the familiar quantum numbers n, l, and m.

According to Eqs. (13.106) and (13.115), a hydrogen atom can only make a spontaneous transition from an energy state corresponding to the quantum numbers n, l, m to one corresponding to the quantum numbers n', l', m' if the modulus squared of the associated electric dipole moment

$$d^2 = |\langle n, l, m|e\,x|n', l', m'\rangle|^2 + |\langle n, l, m|e\,y|n', l', m'\rangle|^2 + |\langle n, l, m|e\,z|n', l', m'\rangle|^2 \tag{13.128}$$

is non-zero. Now, we have already seen, in Sect. 12.5, that the matrix element $\langle n, l, m|z|n', l',$ is only non-zero provided that $m' = m$ and $l' = l \pm 1$. It turns out that the proof that this matrix element is zero unless $l' = l \pm 1$ can, via a trivial modification, also be used to demonstrate that $\langle n, l, m|x|n', l', m'\rangle$ and $\langle n, l, m|y|n', l', m'\rangle$ are also zero unless $l' = l \pm 1$. Consider

$$x_\pm = x + i\,y. \tag{13.129}$$

It is easily demonstrated that

$$[L_z, x_\pm] = \pm\,\hbar\,x_\pm. \tag{13.130}$$

Hence,

$$\langle n, l, m|[L_z, x_+] - \hbar\,x_+|n', l', m'\rangle = \hbar\,(m - m' - 1)\,\langle n, l, m|x_+|n', l', m'\rangle = 0, \tag{13.131}$$

and

$$\langle n, l, m|[L_z, x_-] + \hbar\,x_-|n', l', m'\rangle = \hbar\,(m - m' + 1)\,\langle n, l, m|x_-|n', l', m'\rangle = 0. \tag{13.132}$$

Clearly, $\langle n, l, m|x_+|n', l', m'\rangle$ is zero unless $m' = m - 1$, and $\langle n, l, m|x_-|n', l', m'\rangle$ is zero unless $m' = m + 1$. Now, $\langle n, l, m|x|n', l', m'\rangle$ and $\langle n, l, m|y|n', l', m'\rangle$ are obviously both zero if $\langle n, l, m|x_+|n', l', m'\rangle$ and $\langle n, l, m|x_-|n', l', m'\rangle$ are both zero. Hence, we conclude that $\langle n, l, m|x|n', l', m'\rangle$ and $\langle n, l, m|y|n', l', m'\rangle$ are only non-zero if $m' = m \pm 1$.

The above arguments demonstrate that spontaneous transitions between different energy levels of a hydrogen atom are only possible provided

$$l' = l \pm 1, \tag{13.133}$$

$$m' = m,\ m \pm 1. \tag{13.134}$$

These are termed the *selection rules* for *electric dipole transitions* (*i.e.*, transitions calculated using the electric dipole approximation). Note, finally, that since the perturbing Hamiltonian does not contain any spin operators, the spin quantum number m_s cannot change during a transition. Hence, we have the additional selection rule that $m_s' = m_s$.

13.12 $2P \rightarrow 1S$ Transitions in Hydrogen

Let us calculate the rate of spontaneous emission between the first excited state (*i.e.*, $n = 2$) and the ground-state (*i.e.*, $n' = 1$) of a hydrogen atom. Now the ground-state is characterized by $l' = m' = 0$. Hence, in order to satisfy the selection rules (13.133) and (13.134), the excited state must have the quantum numbers $l = 1$ and $m = 0, \pm 1$. Thus, we are dealing with a spontaneous transition from a 2P to a 1S state. Note, incidentally, that a spontaneous transition from a 2S to a 1S state is forbidden by our selection rules.

According to Sect. 9.4, the wavefunction of a hydrogen atom takes the form

$$\psi_{n,l,m}(r, \theta, \phi) = R_{n,l}(r)\,Y_{l,m}(\theta, \phi), \tag{13.135}$$

where the radial functions $R_{n,l}$ are given in Sect. 9.4, and the spherical harmonics $Y_{l,m}$ are given in Sect. 8.7. Some straight-forward, but tedious, integration reveals that

$$\langle 1,0,0|x|2,1,\pm 1\rangle = \pm\frac{2^7}{3^5}\,a_0, \tag{13.136}$$

$$\langle 1,0,0|y|2,1,\pm 1\rangle = i\,\frac{2^7}{3^5}\,a_0, \tag{13.137}$$

$$\langle 1,0,0|z|2,1,0\rangle = \sqrt{2}\,\frac{2^7}{3^5}\,a_0, \tag{13.138}$$

where a_0 is the Bohr radius specified in Eq. (9.58). All of the other possible 2P → 1S matrix elements are zero because of the selection rules. If follows from Eq. (13.128) that the modulus squared of the dipole moment for the 2P → 1S transition takes the same value

$$d^2 = \frac{2^{15}}{3^{10}}\,(e\,a_0)^2 \tag{13.139}$$

for $m = 0$, 1, or -1. Clearly, the transition rate is independent of the quantum number m. It turns out that this is a general result.

Now, the energy of the eigenstate of the hydrogen atom characterized by the quantum numbers n, l, m is $E = E_0/n^2$, where the ground-state energy E_0 is specified in Eq. (9.57). Hence, the energy of the photon emitted during a 2P → 1S transition is

$$\hbar\,\omega = E_0/4 - E_0 = -\frac{3}{4}\,E_0 = 10.2\,\text{eV}. \tag{13.140}$$

This corresponds to a wavelength of 1.215×10^{-7} m.

Finally, according to Eq. (13.115), the 2P → 1S transition rate is written

$$w_{2P\to 1S} = \frac{\omega^3\,d^2}{3\pi\,\epsilon_0\,\hbar\,c^3}, \tag{13.141}$$

which reduces to

$$w_{2P\to 1S} = \left(\frac{2}{3}\right)^8 \alpha^5\,\frac{m_e\,c^2}{\hbar} = 6.27\times 10^8\,\text{s}^{-1} \tag{13.142}$$

with the aid of Eqs. (13.139) and (13.140). Here, $\alpha = 1/137$ is the fine-structure constant. Hence, the mean life-time of a hydrogen 2P state is

$$\tau_{2P} = (w_{2P\to 1S})^{-1} = 1.6\,\text{ns}. \tag{13.143}$$

Incidentally, since the 2P state only has a finite life-time, it follows from the energy-time uncertainty relation that the energy of this state is uncertain by an amount

$$\Delta E_{2P} \sim \frac{\hbar}{\tau_{2P}} \sim 4\times 10^{-7}\,\text{eV}. \tag{13.144}$$

This uncertainty gives rise to a *finite width* of the spectral line associated with the 2P → 1S transition. This natural line-width is of order

$$\frac{\Delta\lambda}{\lambda} \sim \frac{\Delta E_{2P}}{\hbar\,\omega} \sim 4\times 10^{-8}. \tag{13.145}$$

13.13 Intensity Rules

Now, we know, from Sect. 12.8, that when we take electron spin and spin-orbit coupling into account the degeneracy of the six 2P states of the hydrogen atom is broken. In fact, these states are divided into two groups with slightly different energies. There are four states characterized by the overall angular momentum quantum number $j = 3/2$—these are called the $2P_{3/2}$ states. The remaining two states are characterized by $j = 1/2$, and are thus called the $2P_{1/2}$ states. The energy of the $2P_{3/2}$ states is slightly higher than that of the $2P_{1/2}$ states. In fact, the energy difference is

$$\Delta E = -\frac{\alpha^2}{16}\,E_0 = 4.53\times 10^{-5}\,\text{eV}. \tag{13.146}$$

Thus, the wavelength of the spectral line associated with the 2P → 1S transition in hydrogen is split by a relative amount

$$\frac{\Delta\lambda}{\lambda} = \frac{\Delta E}{\hbar\,\omega} = 4.4\times 10^{-6}. \tag{13.147}$$

Note that this splitting is much greater than the natural line-width estimated in Eq. (13.145), so there really are two spectral lines. How does all of this affect the rate of the 2P → 1S transition?

Well, we have seen that the transition rate is independent of spin, and hence of the spin quantum number m_s, and is also independent of the quantum number m. It follows that the transition rate is independent of the z-component of total angular momentum quantum number $m_j = m + m_s$. However, if this is the case, then the transition rate is plainly also independent of the total angular momentum quantum number j. Hence, we expect the $2P_{3/2} \rightarrow 1S$ and $2P_{1/2} \rightarrow 1S$ transition rates to be the *same*. However, there are *four* $2P_{3/2}$ states and only *two* $2P_{1/2}$ states. If these states are equally populated—which we would certainly expect to be the case in thermal equilibrium, since they have almost the same energies—and since they decay to the 1S state at the same rate, it stands to reason that the spectral line associated with the $2P_{3/2} \rightarrow 1S$ transition is *twice* as bright as that associated with the $2P_{1/2} \rightarrow 1S$ transition.

13.14 Forbidden Transitions

Atomic transitions which are forbidden by the electric dipole selection rules (13.133) and (13.134) are unsurprisingly known as *forbidden transitions*. It is clear from the analysis in Sect. 13.8 that a forbidden transition is one for which the matrix element $\langle f|\epsilon\cdot\mathbf{p}|i\rangle$ is zero. However, this matrix element is only an approximation to the true matrix element for radiative transitions, which takes the form $\langle f|\epsilon\cdot\mathbf{p}\,\exp(\,i\,\mathbf{k}\cdot\mathbf{r})|i\rangle$. Expanding $\exp(\,i\,\mathbf{k}\cdot\mathbf{r})$, and keeping the first two terms, the matrix element for a forbidden transition becomes

$$\langle f|\epsilon\cdot\mathbf{p}\,\exp(\,i\,\mathbf{k}\cdot\mathbf{r})|i\rangle \simeq i\,\langle f|(\epsilon\cdot\mathbf{p})\,(\mathbf{k}\cdot\mathbf{r})|i\rangle. \tag{13.148}$$

Hence, if the residual matrix element on the right-hand side of the above expression is non-zero then a "forbidden" transition can take place, allbeit at a much reduced rate. In fact, in Sect. 13.9, we calculated that the typical rate of an electric dipole transition is

$$w_{i\rightarrow f} \sim \alpha^3\,\omega_{if}. \tag{13.149}$$

Since the transition rate is proportional to the square of the radiative matrix element, it is clear that the transition rate for a forbidden transition enabled by the residual matrix element (13.148) is smaller than that of an electric dipole transition by a factor $(k\,r)^2$. Estimating r as the Bohr radius, and k as the wavenumber of a typical spectral line of hydrogen, it is easily demonstrated that

$$w_{i\rightarrow f} \sim \alpha^5\,\omega_{if} \tag{13.150}$$

for such a transition. Of course, there are some transitions (in particular, the $2S \rightarrow 1S$ transition) for which the true radiative matrix element $\langle f|\boldsymbol{\epsilon}\cdot\mathbf{p}\,\exp(\,i\,\mathbf{k}\cdot\mathbf{r})|i\rangle$ is zero. Such transitions are absolutely forbidden.

Finally, it is fairly obvious that excited states which decay via forbidden transitions have much longer life-times than those which decay via electric dipole transitions. Since the natural width of a spectral line is inversely proportional to the life-time of the associated decaying state, it follows that spectral lines associated with forbidden transitions are generally *much sharper* than those associated with electric dipole transitions.

14 Variational Methods

14.1 Introduction

We have seen, in Sect. 9.4, that we can solve Schrödinger's equation *exactly* to find the stationary eigenstates of a hydrogen atom. Unfortunately, it is not possible to find exact solutions of Schrödinger's equation for atoms more complicated than hydrogen, or for molecules. In such systems, the best that we can do is to find *approximate* solutions. Most of the methods which have been developed for finding such solutions employ the so-called *variational principle* discussed below.

14.2 Variational Principle

Suppose that we wish to solve the time-independent Schrödinger equation

$$H\,\psi = E\,\psi, \tag{14.1}$$

where H is a known (presumably complicated) time-independent Hamiltonian. Let ψ be a *normalized* trial solution to the above equation. The variational principle states, quite simply, that the ground-state energy, E_0, is always less than or equal to the expectation value of H calculated with the trial wavefunction: *i.e.*,

$$E_0 \leq \langle\psi|H|\psi\rangle. \tag{14.2}$$

Thus, by varying ψ until the expectation value of H is *minimized,* we can obtain an approximation to the wavefunction and energy of the ground-state.

Let us prove the variational principle. Suppose that the ψ_n and the E_n are the true eigenstates and eigenvalues of H: *i.e.*,

$$H\,\psi_n = E_n\,\psi_n. \tag{14.3}$$

Furthermore, let

$$E_0 < E_1 < E_2 < \cdots, \tag{14.4}$$

so that ψ_0 is the ground-state, ψ_1 the first excited state, *etc*. The ψ_n are assumed to be orthonormal: *i.e.*,

$$\langle\psi_n|\psi_m\rangle = \delta_{nm}. \tag{14.5}$$

If our trial wavefunction ψ is properly normalized then we can write

$$\psi = \sum_n c_n\,\psi_n, \tag{14.6}$$

where

$$\sum_n |c_n|^2 = 1. \tag{14.7}$$

Now, the expectation value of H, calculated with ψ, takes the form

$$\begin{aligned}\langle\psi|H|\psi\rangle &= \left\langle \sum_n c_n\,\psi_n \middle| H \middle| \sum_m c_m\,\psi_m \right\rangle = \sum_{n,m} c_n^\ast\, c_m\,\langle\psi_n|H|\psi_m\rangle \\ &= \sum_n c_n^\ast\, c_m\, E_m\,\langle\psi_n|\psi_m\rangle = \sum_n E_n\,|c_n|^2, \end{aligned} \tag{14.8}$$

where use has been made of Eqs. (14.3) and (14.5). So, we can write

$$\langle\psi|H|\psi\rangle = |c_0|^2\,E_0 + \sum_{n>0} |c_n|^2\,E_n. \tag{14.9}$$

However, Eq. (14.7) can be rearranged to give

$$|c_0|^2 = 1 - \sum_{n>0} |c_n|^2. \tag{14.10}$$

Combining the previous two equations, we obtain

$$\langle\psi|H|\psi\rangle = E_0 + \sum_{n>0} |c_n|^2\,(E_n - E_0). \tag{14.11}$$

Now, the second term on the right-hand side of the above expression is *positive definite*, since $E_n - E_0 > 0$ for all $n > 0$ [see (14.4)]. Hence, we obtain the desired result

$$\langle\psi|H|\psi\rangle \geq E_0. \tag{14.12}$$

Suppose that we have found a good approximation, $\tilde{\psi}_0$, to the ground-state wavefunction. If ψ is a normalized trial wavefunction which is orthogonal to $\tilde{\psi}_0$ (*i.e.*, $\langle\psi|\tilde{\psi}_0\rangle = 0$) then, by repeating the above analysis, we can easily demonstrate that

$$\langle\psi|H|\psi\rangle \geq E_1. \tag{14.13}$$

Thus, by varying ψ until the expectation value of H is *minimized*, we can obtain an approximation to the wavefunction and energy of the first excited state. Obviously, we can continue this process until we have approximations to all of the stationary eigenstates. Note, however, that the errors are clearly cumulative in this method, so that any approximations to highly excited states are unlikely to be very accurate. For this reason, the variational method is generally only used to calculate the ground-state and first few excited states of complicated quantum systems.

14.3 Helium Atom

A helium atom consists of a nucleus of charge $+2\,e$ surrounded by two electrons. Let us attempt to calculate its ground-state energy.

Let the nucleus lie at the origin of our coordinate system, and let the position vectors of the two electrons be $\mathbf{r}_1$ and $\mathbf{r}_2$, respectively. The Hamiltonian of the system thus takes the form

$$H = -\frac{\hbar^2}{2\,m_e}\left(\nabla_1^{\,2} + \nabla_2^{\,2}\right) - \frac{e^2}{4\pi\,\epsilon_0}\left(\frac{2}{r_1} + \frac{2}{r_2} - \frac{1}{|\mathbf{r}_2 - \mathbf{r}_1|}\right), \tag{14.14}$$

where we have neglected any reduced mass effects. The terms in the above expression represent the kinetic energy of the first electron, the kinetic energy of the second electron, the electrostatic attraction between the nucleus and the first electron, the electrostatic attraction between the nucleus and the second electron, and the electrostatic repulsion between the two electrons, respectively. It is the final term which causes all of the difficulties. Indeed, if this term is neglected then we can write

$$H = H_1 + H_2, \tag{14.15}$$

where

$$H_{1,2} = -\frac{\hbar^2}{2\,m_e}\,\nabla_{1,2}^{\,2} - \frac{2\,e^2}{4\pi\,\epsilon_0\,r_{1,2}}. \tag{14.16}$$

In other words, the Hamiltonian just becomes the sum of separate Hamiltonians for each electron. In this case, we would expect the wavefunction to be separable: *i.e.*,

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_1(\mathbf{r}_1)\,\psi_2(\mathbf{r}_2). \tag{14.17}$$

Hence, Schrödinger's equation

$$H\,\psi = E\,\psi \tag{14.18}$$

reduces to

$$H_{1,2}\,\psi_{1,2} = E_{1,2}\,\psi_{1,2}, \tag{14.19}$$

where

$$E = E_1 + E_2. \tag{14.20}$$

Of course, Eq. (14.19) is the Schrödinger equation of a hydrogen atom whose nuclear charge is $+2\,e$, instead of $+e$. It follows, from Sect. 9.4 (making the substitution $e^2 \rightarrow 2\,e^2$), that if both electrons are in their lowest energy states then

$$\psi_1(\mathbf{r}_1) \;=\; \psi_0(\mathbf{r}_1), \tag{14.21}$$

$$\psi_2(\mathbf{r}_2) \;=\; \psi_0(\mathbf{r}_2), \tag{14.22}$$

where

$$\psi_0(\mathbf{r}) = \frac{4}{\sqrt{2\,\pi}\,a_0^{3/2}}\,\exp\left(-\frac{2\,r}{a_0}\right). \tag{14.23}$$

Here, a_0 is the Bohr radius [see Eq. (9.58)]. Note that ψ_0 is properly normalized. Furthermore,

$$E_1 = E_2 = 4\,E_0, \tag{14.24}$$

where $E_0 = -13.6\,\text{eV}$ is the hydrogen ground-state energy [see Eq. (9.57)]. Thus, our crude estimate for the ground-state energy of helium becomes

$$E = 4\,E_0 + 4\,E_0 = 8\,E_0 = -108.8\,\text{eV}. \tag{14.25}$$

Unfortunately, this estimate is significantly different from the experimentally determined value, which is $-78.98\,\text{eV}$. This fact demonstrates that the neglected electron-electron repulsion term makes a large contribution to the helium ground-state energy. Fortunately, however, we can use the variational principle to estimate this contribution.

Let us employ the separable wavefunction discussed above as our trial solution. Thus,

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_0(\mathbf{r}_1)\,\psi_0(\mathbf{r}_2) = \frac{8}{\pi\,a_0^3}\,\exp\left(-\frac{2\,[r_1 + r_2]}{a_0}\right). \tag{14.26}$$

The expectation value of the Hamiltonian (14.14) thus becomes

$$\langle H\rangle = 8\,E_0 + \langle V_{ee}\rangle, \tag{14.27}$$

where

$$\langle V_{ee}\rangle = \left\langle \psi\left|\frac{e^2}{4\pi\,\epsilon_0\,|\mathbf{r}_2 - \mathbf{r}_1|}\right|\psi\right\rangle = \frac{e^2}{4\pi\,\epsilon_0}\int\frac{|\psi(\mathbf{r}_1, \mathbf{r}_2)|^2}{|\mathbf{r}_2 - \mathbf{r}_1|}\,d^3\mathbf{r}_1\,d^3\mathbf{r}_2. \tag{14.28}$$

The variation principle only guarantees that (14.27) yields an *upper bound* on the ground-state energy. In reality, we hope that it will give a reasonably accurate estimate of this energy.

It follows from Eqs. (9.57), (14.26) and (14.28) that

$$\langle V_{ee}\rangle = -\frac{4\,E_0}{\pi^2}\int\frac{e^{-2\,(\hat{r}_1 + \hat{r}_2)}}{|\hat{\mathbf{r}}_1 - \hat{\mathbf{r}}_2|}\,d^3\hat{\mathbf{r}}_1\,d^3\hat{\mathbf{r}}_2, \tag{14.29}$$

where $\hat{\mathbf{r}}_{1,2} = 2\,\mathbf{r}_{1,2}/a_0$. Neglecting the hats, for the sake of clarity, the above expression can also be written

$$\langle V_{ee}\rangle = -\frac{4\,E_0}{\pi^2}\int\frac{e^{-2\,(r_1 + r_2)}}{\sqrt{r_1^2 + r_2^2 - 2\,r_1\,r_2\,\cos\theta}}\,d^3\mathbf{r}_1\,d^3\mathbf{r}_2, \tag{14.30}$$

where θ is the angle subtended between vectors $\mathbf{r}_1$ and $\mathbf{r}_2$. If we perform the integral in $\mathbf{r}_1$ space before that in $\mathbf{r}_2$ space then

$$\langle V_{ee}\rangle = -\frac{4\,E_0}{\pi^2}\int e^{-2\,r_2}\,I(\mathbf{r}_2)\,d^3\mathbf{r}_2, \tag{14.31}$$

where

$$I(\mathbf{r}_2) = \int \frac{e^{-2\,r_1}}{\sqrt{r_1^2 + r_2^2 - 2\,r_1\,r_2\,\cos\theta}}\,d^3\mathbf{r}_1. \tag{14.32}$$

Our first task is to evaluate the function $I(\mathbf{r}_2)$. Let $(r_1,\,\theta_1,\,\phi_1)$ be a set of spherical polar coordinates in $\mathbf{r}_1$ space whose axis of symmetry runs in the direction of $\mathbf{r}_2$. It follows that $\theta = \theta_1$. Hence,

$$I(\mathbf{r}_2) = \int_0^\infty\int_0^\pi\int_0^{2\pi} \frac{e^{-2\,r_1}}{\sqrt{r_1^2 + r_2^2 - 2\,r_1\,r_2\,\cos\theta_1}}\,r_1^2\,dr_1\,\sin\theta_1\,d\theta_1\,d\phi_1, \tag{14.33}$$

which trivially reduces to

$$I(\mathbf{r}_2) = 2\pi\int_0^\infty\int_0^\pi \frac{e^{-2\,r_1}}{\sqrt{r_1^2 + r_2^2 - 2\,r_1\,r_2\,\cos\theta_1}}\,r_1^2\,dr_1\,\sin\theta_1\,d\theta_1. \tag{14.34}$$

Making the substitution $\mu = \cos\theta_1$, we can see that

$$\int_0^\pi \frac{1}{\sqrt{r_1^2 + r_2^2 - 2\,r_1\,r_2\,\cos\theta_1}}\,\sin\theta_1\,d\theta_1 = \int_{-1}^{1}\frac{d\mu}{\sqrt{r_1^2 + r_2^2 - 2\,r_1\,r_2\,\mu}}. \tag{14.35}$$

Now,

$$\begin{aligned}\int_{-1}^{1}\frac{d\mu}{\sqrt{r_1^2 + r_2^2 - 2\,r_1\,r_2\,\mu}} &= \left[\frac{\sqrt{r_1^2 + r_2^2 - 2\,r_1\,r_2\,\mu}}{r_1\,r_2}\right]_{+1}^{-1} \\ &= \frac{(r_1 + r_2) - |r_1 - r_2|}{r_1\,r_2} \\ &= \left\{\begin{array}{ll} 2/r_1 & \text{for } r_1 > r_2 \\ 2/r_2 & \text{for } r_1 < r_2 \end{array}\right., \end{aligned} \tag{14.36}$$

giving

$$I(\mathbf{r}_2) = 4\pi\left(\frac{1}{r_2}\int_0^{r_2} e^{-2\,r_1}\,r_1^2\,dr_1 + \int_{r_2}^\infty e^{-2\,r_1}\,r_1\,dr_1\right). \tag{14.37}$$

But,

$$\int e^{-\beta\,x}\,x\,dx = -\frac{e^{-\beta\,x}}{\beta^2}\,(1 + \beta\,x), \tag{14.38}$$

$$\int e^{-\beta\,x}\,x^2\,dx = -\frac{e^{-\beta\,x}}{\beta^3}\,(2 + 2\,\beta\,x + \beta^2\,x^2), \tag{14.39}$$

yielding

$$I(\mathbf{r}_2) = \frac{\pi}{r_2}\left[1 - e^{-2\,r_2}\,(1 + r_2)\right]. \tag{14.40}$$

Since the function $I(\mathbf{r}_2)$ only depends on the magnitude of $\mathbf{r}_2$, the integral (14.31) reduces to

$$\langle V_{ee}\rangle = -\frac{16\,E_0}{\pi}\int_0^\infty e^{-2\,r_2}\,I(r_2)\,r_2^2\,dr_2, \tag{14.41}$$

which yields

$$\langle V_{ee}\rangle = -16\,E_0\int_0^\infty e^{-2\,r_2}\left[1-e^{-2\,r_2}\,(1+r_2)\right]r_2\,dr_2 = -\frac{5}{2}\,E_0. \tag{14.42}$$

Hence, from (14.27), our estimate for the ground-state energy of helium is

$$\langle H\rangle = 8\,E_0 - \frac{5}{2}\,E_0 = \frac{11}{2}\,E_0 = -74.8\,\text{eV}. \tag{14.43}$$

This is remarkably close to the correct result.

We can actually refine our estimate further. The trial wavefunction (14.26) essentially treats the two electrons as *non-interacting* particles. In reality, we would expect one electron to partially shield the nuclear charge from the other, and *vice versa*. Hence, a better trial wavefunction might be

$$\psi(\mathbf{r}_1,\mathbf{r}_2) = \frac{Z^3}{\pi\,a_0^3}\,\exp\left(-\frac{Z\,[r_1+r_2]}{a_0}\right), \tag{14.44}$$

where $Z < 2$ is effective nuclear charge number seen by each electron. Let us recalculate the ground-state energy of helium as a function of Z, using the above trial wavefunction, and then *minimize* the result with respect to Z. According to the variational principle, this should give us an even better estimate for the ground-state energy.

We can rewrite the expression (14.14) for the Hamiltonian of the helium atom in the form

$$H = H_1(Z) + H_2(Z) + V_{ee} + U(Z), \tag{14.45}$$

where

$$H_{1,2}(Z) = -\frac{\hbar^2}{2\,m_e}\,\nabla_{1,2}^2 - \frac{Z\,e^2}{4\pi\,\epsilon_0\,r_{1,2}} \tag{14.46}$$

is the Hamiltonian of a hydrogen atom with nuclear charge $+Z\,e$,

$$V_{ee} = \frac{e^2}{4\pi\,\epsilon_0}\,\frac{1}{|\mathbf{r}_2-\mathbf{r}_1|} \tag{14.47}$$

is the electron-electron repulsion term, and

$$U(Z) = \frac{e^2}{4\pi\,\epsilon_0}\left(\frac{[Z-2]}{r_1} + \frac{[Z-2]}{r_2}\right). \tag{14.48}$$

It follows that

$$\langle H\rangle(Z) = 2\,E_0(Z) + \langle V_{ee}\rangle(Z) + \langle U\rangle(Z), \tag{14.49}$$

where $E_0(Z) = Z^2\,E_0$ is the ground-state energy of a hydrogen atom with nuclear charge $+Z\,e$, $\langle V_{ee}\rangle(Z) = -(5\,Z/4)\,E_0$ is the value of the electron-electron repulsion term when recalculated with the wavefunction (14.44) [actually, all we need to do is to make the substitution $a_0 \rightarrow (2/Z)\,a_0$], and

$$\langle U\rangle(Z) = 2\,(Z-2)\left(\frac{e^2}{4\pi\,\epsilon_0}\right)\left\langle\frac{1}{r}\right\rangle. \tag{14.50}$$

Here, $\langle 1/r\rangle$ is the expectation value of $1/r$ calculated for a hydrogen atom with nuclear charge $+Z\,e$. It follows from Eq. (9.74) [with $n = 1$, and making the substitution $a_0 \rightarrow a_0/Z$] that

$$\left\langle\frac{1}{r}\right\rangle = \frac{Z}{a_0}. \tag{14.51}$$

Hence,

$$\langle U\rangle(Z) = -4\,Z\,(Z-2)\,E_0, \tag{14.52}$$

since $E_0 = -e^2/(8\pi\,\epsilon_0\,a_0)$. Collecting the various terms, our new expression for the expectation value of the Hamiltonian becomes

$$\langle H\rangle(Z) = \left[2\,Z^2 - \frac{5}{4}\,Z - 4\,Z\,(Z-2)\right]E_0 = \left[-2\,Z^2 + \frac{27}{4}\,Z\right]E_0. \tag{14.53}$$

The value of Z which minimizes this expression is the root of

$$\frac{d\langle H\rangle}{dZ} = \left[-4\,Z + \frac{27}{4}\right]E_0 = 0. \tag{14.54}$$

It follows that

$$Z = \frac{27}{16} = 1.69. \tag{14.55}$$

The fact that $Z < 2$ confirms our earlier conjecture that the electrons partially shield the nuclear charge from one another. Our new estimate for the ground-state energy of helium is

$$\langle H\rangle(1.69) = \frac{1}{2}\left(\frac{3}{2}\right)^6 E_0 = -77.5\,\text{eV}. \tag{14.56}$$

This is clearly an improvement on our previous estimate (14.43) [recall that the correct result is -78.98 eV].

Obviously, we could get even closer to the correct value of the helium ground-state energy by using a more complicated trial wavefunction with more adjustable parameters.

Note, finally, that since the two electrons in a helium atom are *indistinguishable fermions*, the overall wavefunction must be *anti-symmetric* with respect to exchange of particles (see Sect. 6). Now, the overall wavefunction is the product of the *spatial wavefunction* and the *spinor* representing the spin-state. Our spatial wavefunction (14.44) is obviously *symmetric* with respect to exchange of particles. This means that the spinor must be *anti-symmetric*.

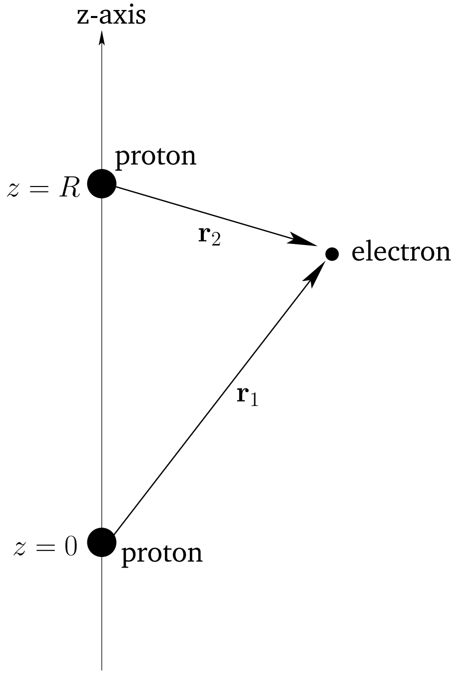

Figure 14.1: *The hydrogen molecule ion.*

It is clear, from Sect. 11.4, that if the spin-state of an $l = 0$ system consisting of two spin one-half particles (*i.e.*, two electrons) is *anti-symmetric* with respect to interchange of particles then the system is in the so-called *singlet* state with overall spin zero. Hence, the ground-state of helium has overall electron spin zero.

14.4 Hydrogen Molecule Ion

The hydrogen molecule ion consists of an electron orbiting about two protons, and is the simplest imaginable molecule. Let us investigate whether or not this molecule possesses a bound state: *i.e.*, whether or not it possesses a ground-state whose energy is less than that of a hydrogen atom and a free proton. According to the variation principle, we can deduce that the H_2^+ ion has a bound state if we can find *any* trial wavefunction for which the total Hamiltonian of the system has an expectation value less than that of a hydrogen atom and a free proton.

Suppose that the two protons are separated by a distance R. In fact, let them lie on the z-axis, with the first at the origin, and the second at $z = R$ (see Fig. 14.1). In the following, we shall treat the protons as essentially stationary. This is reasonable, since the electron moves far more rapidly than the protons.

Let us try

$$\psi(\mathbf{r})_\pm = A\,[\psi_0(\mathbf{r}_1) \pm \psi_0(\mathbf{r}_2)] \tag{14.57}$$

as our trial wavefunction, where

$$\psi_0(\mathbf{r}) = \frac{1}{\sqrt{\pi}\, a_0^{3/2}}\, e^{-r/a_0} \tag{14.58}$$

is a normalized hydrogen ground-state wavefunction centered on the origin, and $\mathbf{r}_{1,2}$ are the position vectors of the electron with respect to each of the protons (see Fig. 14.1). Obviously, this is a very simplistic wavefunction, since it is just a linear combination of hydrogen ground-state wavefunctions centered on each proton. Note, however, that the wavefunction respects the obvious symmetries in the problem.

Our first task is to normalize our trial wavefunction. We require that

$$\int |\psi_\pm|^2\, d^3\mathbf{r} = 1. \tag{14.59}$$

Hence, from (14.57), $A = I^{-1/2}$, where

$$I = \int \left[|\psi_0(\mathbf{r}_1)|^2 + |\psi_0(\mathbf{r}_2)|^2 \pm 2\, \psi_0(\mathbf{r}_1)\, \psi(\mathbf{r}_2)\right] d^3\mathbf{r}. \tag{14.60}$$

It follows that

$$I = 2\,(1 \pm J), \tag{14.61}$$

with

$$J = \int \psi_0(\mathbf{r}_1)\, \psi_0(\mathbf{r}_2)\, d^3\mathbf{r}. \tag{14.62}$$

Let us employ the standard spherical polar coordinates (r, θ, ϕ). Now, it is easily seen that $r_1 = r$ and $r_2 = (r^2 + R^2 - 2\, r\, R\, \cos\theta)^{1/2}$. Hence,

$$J = 2 \int_0^\infty \int_0^\pi \exp\left[-x - (x^2 + X^2 - 2\, x\, X\, \cos\theta)^{1/2}\right]\, x^2\, dx\, \sin\theta\, d\theta, \tag{14.63}$$

where $X = R/a_0$. Here, we have already performed the trivial ϕ integral. Let $y = (x^2 + X^2 - 2\, x\, X\, \cos\theta)^{1/2}$. It follows that $d(y^2) = 2\, y\, dy = 2\, x\, X\, \sin\theta\, d\theta$, giving

$$\begin{aligned} \int_0^\pi e^{\,(x^2 + X^2 - 2\, x\, X\, \cos\theta)^{1/2}}\, \sin\theta\, d\theta &= \frac{1}{x\, X} \int_{|x-X|}^{x+X} e^{-y}\, y\, dy \qquad (14.64) \\ &= -\frac{1}{x\, X}\left[e^{-(x+X)}\,(1 + x + X) - e^{-|x-X|}\,(1 + |x - X|)\right]. \end{aligned}$$

Thus,

$$\begin{aligned} J &= -\frac{2}{X}\, e^{-X} \int_0^X \left[e^{-2\,x}\,(1 + X + x) - (1 + X - x)\right] x\, dx \\ &\quad -\frac{2}{X} \int_X^\infty e^{-2\,x}\left[e^{-X}\,(1 + X + x) - e^{X}\,(1 - X + x)\right] x\, dx, \qquad (14.65) \end{aligned}$$

which evaluates to

$$J = e^{-X}\left(1 + X + \frac{X^3}{3}\right). \tag{14.66}$$

Now, the Hamiltonian of the electron is written

$$H = -\frac{\hbar^2}{2\,m_e}\,\nabla^2 - \frac{e^2}{4\pi\,\epsilon_0}\left(\frac{1}{r_1} + \frac{1}{r_2}\right). \tag{14.67}$$

Note, however, that

$$\left(-\frac{\hbar^2}{2\,m_e}\,\nabla^2 - \frac{e^2}{4\pi\,\epsilon_0\,r_{1,2}}\right)\psi_0(\mathbf{r}_{1,2}) = E_0\,\psi_0(\mathbf{r}_{1,2}), \tag{14.68}$$

since $\psi_0(\mathbf{r}_{1,2})$ are hydrogen ground-state wavefunctions. It follows that

$$\begin{aligned} H\,\psi_\pm &= A\left[-\frac{\hbar^2}{2\,m_e}\,\nabla^2 - \frac{e^2}{4\pi\,\epsilon_0}\left(\frac{1}{r_1} + \frac{1}{r_2}\right)\right][\psi_0(\mathbf{r}_1) \pm \psi_0(\mathbf{r}_2)] \\ &= E_0\,\psi - A\left(\frac{e^2}{4\pi\,\epsilon_0}\right)\left[\frac{\psi_0(\mathbf{r}_1)}{r_2} \pm \frac{\psi_0(\mathbf{r}_2)}{r_1}\right]. \end{aligned} \tag{14.69}$$

Hence,

$$\langle H\rangle = E_0 + 4\,A^2\,(D \pm E)\,E_0, \tag{14.70}$$

where

$$D = \left\langle \psi_0(\mathbf{r}_1)\left|\frac{a_0}{r_2}\right|\psi_0(\mathbf{r}_1)\right\rangle, \tag{14.71}$$

$$E = \left\langle \psi_0(\mathbf{r}_1)\left|\frac{a_0}{r_1}\right|\psi_0(\mathbf{r}_2)\right\rangle. \tag{14.72}$$

Now,

$$D = 2\int_0^\infty\int_0^\pi \frac{e^{-2\,x}}{(x^2 + X^2 - 2\,x\,X\,\cos\theta)^{1/2}}\,x^2\,dx\,\sin\theta\,d\theta, \tag{14.73}$$

which reduces to

$$D = \frac{4}{X}\int_0^X e^{-2\,x}\,x^2\,dx + 4\int_X^\infty e^{-2\,x}\,x\,dx, \tag{14.74}$$

giving

$$D = \frac{1}{X}\left(1 - [1 + X]\,e^{-2\,X}\right). \tag{14.75}$$

Furthermore,

$$E = 2\int_0^\infty\int_0^\pi \exp\left[-x - (x^2 + X^2 - 2\,x\,X\,\cos\theta)^{1/2}\right]\,x\,dx\,\sin\theta\,d\theta, \tag{14.76}$$

which reduces to

$$
\begin{aligned}
E &= -\frac{2}{X}\,e^{-X}\int_0^X\left[e^{-2x}\,(1+X+x)-(1+X-x)\right]dx\\
&\quad -\frac{2}{X}\int_X^\infty e^{-2x}\left[e^{-X}\,(1+X+x)-e^{X}\,(1-X+x)\right]dx, \qquad (14.77)
\end{aligned}
$$

yielding

$$E = (1+X)\,e^{-X}. \qquad (14.78)$$

Our expression for the expectation value of the electron Hamiltonian is

$$\langle H\rangle = \left[1+2\,\frac{(D\pm E)}{(1\pm J)}\right]E_0, \qquad (14.79)$$

where J, D, and E are specified as functions of $X = R/a_0$ in Eqs. (14.66), (14.75), and (14.78), respectively. In order to obtain the total energy of the molecule, we must add to this the potential energy of the two protons. Thus,

$$E_{\rm total} = \langle H\rangle + \frac{e^2}{4\pi\,\epsilon_0\,R} = \langle H\rangle - \frac{2}{X}\,E_0, \qquad (14.80)$$

since $E_0 = -e^2/(8\pi\,\epsilon_0\,a_0)$. Hence, we can write

$$E_{\rm total} = -F_\pm(R/a_0)\,E_0, \qquad (14.81)$$

where E_0 is the hydrogen ground-state energy, and

$$F_\pm(X) = -1 + \frac{2}{X}\left[\frac{(1+X)\,e^{-2X}\pm(1-2\,X^2/3)\,e^{-X}}{1\pm(1+X+X^2/3)\,e^{-X}}\right]. \qquad (14.82)$$

The functions $F_+(X)$ and $F_-(X)$ are both plotted in Fig. 14.2. Recall that in order for the H_2^+ ion to be in a bound state it must have a lower energy than a hydrogen atom and a free proton: *i.e.*, $E_{\rm total} < E_0$. It follows from Eq. (14.81) that a bound state corresponds to $F_\pm < -1$. Clearly, the *even* trial wavefunction ψ_+ possesses a bound state, whereas the *odd* trial wavefunction ψ_- does not [see Eq. (14.57)]. This is hardly surprising, since the even wavefunction *maximizes* the electron probability density between the two protons, thereby reducing their mutual electrostatic repulsion. On the other hand, the odd wavefunction does exactly the opposite. The *binding energy* of the H_2^+ ion is defined as the difference between its energy and that of a hydrogen atom and a free proton: *i.e.*,

$$E_{\rm bind} = E_{\rm total} - E_0 = -(F_+ + 1)\,E_0. \qquad (14.83)$$

According to the variational principle, the binding energy is less than or equal to the *minimum* binding energy which can be inferred from Fig. 14.2. This minimum occurs when $X \simeq 2.5$ and $F_+ \simeq -1.13$. Thus, our estimates for the separation between the two protons, and the binding energy, for the H_2^+ ion are $R = 2.5\,a_0 = 1.33\times10^{-10}\,\text{m}$ and $E_{\rm bind} = 0.13\,E_0 = -1.77$ eV, respectively. The experimentally determined values are $R = 1.06\times10^{-10}$ m, and $E_{\rm bind} = -2.8$ eV, respectively. Clearly, our estimates are not particularly accurate. However, our calculation does establish, beyond any doubt, the existence of a bound state of the H_2^+ ion, which is all that we set out to achieve.

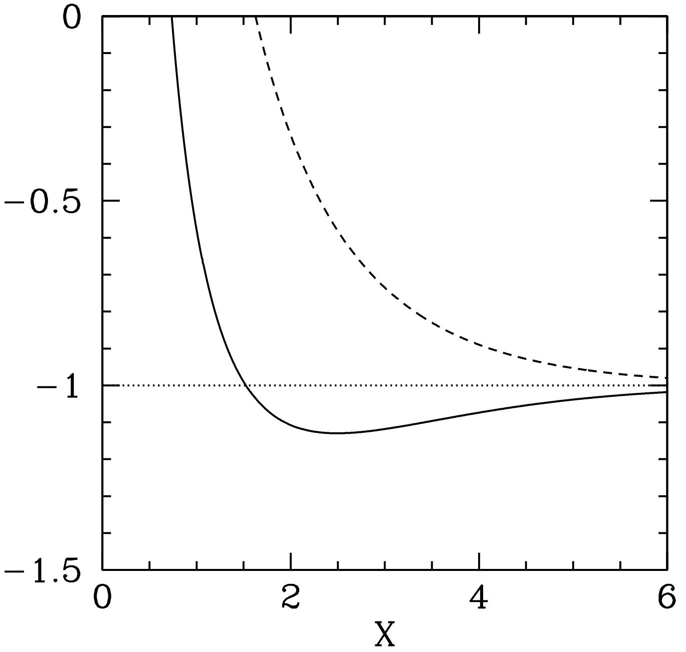

Figure 14.2: *The functions* $F_+(X)$ *(solid curve) and* $F_-(X)$ *(dashed curve)*.

15 Scattering Theory

15.1 Introduction

Historically, data regarding quantum phenomena has been obtained from two main sources. Firstly, from the study of spectroscopic lines, and, secondly, from scattering experiments. We have already developed theories which account for some aspects of the spectrum of hydrogen, and hydrogen-like, atoms. Let us now examine the quantum theory of scattering.

15.2 Fundamentals

Consider time-independent, energy conserving scattering in which the Hamiltonian of the system is written

$$H = H_0 + V(\mathbf{r}), \tag{15.1}$$

where

$$H_0 = \frac{p^2}{2\,m} \equiv -\frac{\hbar^2}{2\,m}\,\nabla^2 \tag{15.2}$$

is the Hamiltonian of a free particle of mass m, and $V(\mathbf{r})$ the scattering potential. This potential is assumed to only be non-zero in a fairly localized region close to the origin. Let

$$\psi_0(\mathbf{r}) = \sqrt{n}\,e^{\,i\,\mathbf{k}\cdot\mathbf{r}} \tag{15.3}$$

represent an incident beam of particles, of number density n, and velocity $\mathbf{v} = \hbar\,\mathbf{k}/m$. Of course,

$$H_0\,\psi_0 = E\,\psi_0, \tag{15.4}$$

where $E = \hbar^2\,k^2/2\,m$ is the particle energy. Schrödinger's equation for the scattering problem is

$$(H_0 + V)\,\psi = E\,\psi, \tag{15.5}$$

subject to the boundary condition $\psi \rightarrow \psi_0$ as $V \rightarrow 0$.

The above equation can be rearranged to give

$$(\nabla^2 + k^2)\,\psi = \frac{2\,m}{\hbar^2}\,V\,\psi. \tag{15.6}$$

Now,

$$(\nabla^2 + k^2)\,u(\mathbf{r}) = \rho(\mathbf{r}) \tag{15.7}$$

is known as the *Helmholtz equation*. The solution to this equation is well-known:[1]

$$u(\mathbf{r}) = u_0(\mathbf{r}) - \int \frac{e^{\,i\,k\,|\mathbf{r}-\mathbf{r}'|}}{4\pi\,|\mathbf{r}-\mathbf{r}'|}\,\rho(\mathbf{r}')\,d^3\mathbf{r}'. \tag{15.8}$$

---

[1] See Griffiths, Sect. 11.4.

Here, $u_0(\mathbf{r})$ is any solution of $(\nabla^2 + k^2)\,u_0 = 0$. Hence, Eq. (15.6) can be inverted, subject to the boundary condition $\psi \to \psi_0$ as $V \to 0$, to give

$$\psi(\mathbf{r}) = \psi_0(\mathbf{r}) - \frac{2\,m}{\hbar^2} \int \frac{e^{\,i\,k\,|\mathbf{r}-\mathbf{r}'|}}{4\pi\,|\mathbf{r}-\mathbf{r}'|}\,V(\mathbf{r}')\,\psi(\mathbf{r}')\,d^3\mathbf{r}'. \tag{15.9}$$

Let us calculate the value of the wavefunction $\psi(\mathbf{r})$ well outside the scattering region. Now, if $r \gg r'$ then

$$|\mathbf{r}-\mathbf{r}'| \simeq r - \hat{\mathbf{r}}\cdot\mathbf{r}' \tag{15.10}$$

to first-order in r'/r, where $\hat{\mathbf{r}}/r$ is a unit vector which points from the scattering region to the observation point. It is helpful to define $\mathbf{k}' = k\,\hat{\mathbf{r}}$. This is the wavevector for particles with the same energy as the incoming particles (*i.e.*, $k' = k$) which propagate from the scattering region to the observation point. Equation (15.9) reduces to

$$\psi(\mathbf{r}) \simeq \sqrt{n}\left[e^{\,i\mathbf{k}\cdot\mathbf{r}} + \frac{e^{\,i\,k\,r}}{r}\,f(\mathbf{k},\mathbf{k}')\right], \tag{15.11}$$

where

$$f(\mathbf{k},\mathbf{k}') = -\frac{m}{2\pi\,\sqrt{n}\,\hbar^2}\int e^{-i\mathbf{k}'\cdot\mathbf{r}'}\,V(\mathbf{r}')\,\psi(\mathbf{r}')\,d^3\mathbf{r}'. \tag{15.12}$$

The first term on the right-hand side of Eq. (15.11) represents the incident particle beam, whereas the second term represents an outgoing *spherical wave* of scattered particles.

The *differential scattering cross-section* $d\sigma/d\Omega$ is defined as the number of particles per unit time scattered into an element of solid angle $d\Omega$, divided by the incident particle flux. From Sect. 7.2, the probability flux (*i.e.*, the particle flux) associated with a wavefunction ψ is

$$\mathbf{j} = \frac{\hbar}{m}\,\mathrm{Im}(\psi^*\,\nabla\psi). \tag{15.13}$$

Thus, the particle flux associated with the incident wavefunction ψ_0 is

$$\mathbf{j} = n\,\mathbf{v}, \tag{15.14}$$

where $\mathbf{v} = \hbar\,\mathbf{k}/m$ is the velocity of the incident particles. Likewise, the particle flux associated with the scattered wavefunction $\psi - \psi_0$ is

$$\mathbf{j}' = n\,\frac{|f(\mathbf{k},\mathbf{k}')|^2}{r^2}\,\mathbf{v}', \tag{15.15}$$

where $\mathbf{v}' = \hbar\,\mathbf{k}'/m$ is the velocity of the scattered particles. Now,

$$\frac{d\sigma}{d\Omega}\,d\Omega = \frac{r^2\,d\Omega\,|\mathbf{j}'|}{|\mathbf{j}|}, \tag{15.16}$$

which yields

$$\frac{d\sigma}{d\Omega} = |f(\mathbf{k},\mathbf{k}')|^2. \tag{15.17}$$

Thus, $|f(\mathbf{k},\mathbf{k}')|^2$ gives the differential cross-section for particles with incident velocity $\mathbf{v} = \hbar\,\mathbf{k}/m$ to be scattered such that their final velocities are directed into a range of solid angles $d\Omega$ about $\mathbf{v}' = \hbar\,\mathbf{k}'/m$. Note that the scattering conserves energy, so that $|\mathbf{v}'| = |\mathbf{v}|$ and $|\mathbf{k}'| = |\mathbf{k}|$.

15.3 Born Approximation

Equation (15.17) is not particularly useful, as it stands, because the quantity $f(\mathbf{k},\mathbf{k}')$ depends on the, as yet, unknown wavefunction $\psi(\mathbf{r})$ [see Eq. (15.12)]. Suppose, however, that the scattering is not particularly strong. In this case, it is reasonable to suppose that the total wavefunction, $\psi(\mathbf{r})$, does not differ substantially from the incident wavefunction, $\psi_0(\mathbf{r})$. Thus, we can obtain an expression for $f(\mathbf{k},\mathbf{k}')$ by making the substitution $\psi(\mathbf{r}) \to \psi_0(\mathbf{r}) = \sqrt{n}\,\exp(\,i\,\mathbf{k}\cdot\mathbf{r})$ in Eq. (15.12). This procedure is called the *Born approximation*.

The Born approximation yields

$$f(\mathbf{k},\mathbf{k}') \simeq \frac{m}{2\pi\,\hbar^2}\int e^{\,i\,(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}'}\,V(\mathbf{r}')\,d^3\mathbf{r}'. \tag{15.18}$$

Thus, $f(\mathbf{k},\mathbf{k}')$ is proportional to the *Fourier transform* of the scattering potential $V(\mathbf{r})$ with respect to the wavevector $\mathbf{q} = \mathbf{k} - \mathbf{k}'$.

For a spherically symmetric potential,

$$f(\mathbf{k}',\mathbf{k}) \simeq -\frac{m}{2\pi\,\hbar^2}\iiint \exp(\,i\,q\,r'\cos\theta')\,V(r')\,r'^2\,dr'\,\sin\theta'\,d\theta'\,d\phi', \tag{15.19}$$

giving

$$f(\mathbf{k}',\mathbf{k}) \simeq -\frac{2\,m}{\hbar^2\,q}\int_0^\infty r'\,V(r')\sin(q\,r')\,dr'. \tag{15.20}$$

Note that $f(\mathbf{k}',\mathbf{k})$ is just a function of q for a spherically symmetric potential. It is easily demonstrated that

$$q \equiv |\mathbf{k} - \mathbf{k}'| = 2\,k\,\sin(\theta/2), \tag{15.21}$$

where θ is the angle subtended between the vectors $\mathbf{k}$ and $\mathbf{k}'$. In other words, θ is the scattering angle. Recall that the vectors $\mathbf{k}$ and $\mathbf{k}'$ have the same length, via energy conservation.

Consider scattering by a Yukawa potential

$$V(r) = \frac{V_0\,\exp(-\mu\,r)}{\mu\,r}, \tag{15.22}$$

where V_0 is a constant, and $1/\mu$ measures the "range" of the potential. It follows from Eq. (15.20) that

$$f(\theta) = -\frac{2\,m\,V_0}{\hbar^2\,\mu}\frac{1}{q^2+\mu^2}, \tag{15.23}$$

since

$$\int_0^\infty \exp(-\mu\, r')\, \sin(q\, r')\, dr' = \frac{q}{q^2+\mu^2}. \tag{15.24}$$

Thus, in the Born approximation, the differential cross-section for scattering by a Yukawa potential is

$$\frac{d\sigma}{d\Omega} \simeq \left(\frac{2\, m\, V_0}{\hbar^2\, \mu}\right)^2 \frac{1}{[2\, k^2\, (1-\cos\theta) + \mu^2]^2}, \tag{15.25}$$

given that

$$q^2 = 4\, k^2\, \sin^2(\theta/2) = 2\, k^2\, (1-\cos\theta). \tag{15.26}$$

The Yukawa potential reduces to the familiar Coulomb potential as $\mu \rightarrow 0$, provided that $V_0/\mu \rightarrow Z\, Z'\, e^2/4\pi\, \epsilon_0$. In this limit, the Born differential cross-section becomes

$$\frac{d\sigma}{d\Omega} \simeq \left(\frac{2\, m\, Z\, Z'\, e^2}{4\pi\, \epsilon_0\, \hbar^2}\right)^2 \frac{1}{16\, k^4\, \sin^4(\theta/2)}. \tag{15.27}$$

Recall that $\hbar\, k$ is equivalent to $|\mathbf{p}|$, so the above equation can be rewritten

$$\frac{d\sigma}{d\Omega} \simeq \left(\frac{Z\, Z'\, e^2}{16\pi\, \epsilon_0\, E}\right)^2 \frac{1}{\sin^4(\theta/2)}, \tag{15.28}$$

where $E = p^2/2\, m$ is the kinetic energy of the incident particles. Of course, Eq. (15.28) is the famous *Rutherford scattering cross-section* formula.

The Born approximation is valid provided that $\psi(\mathbf{r})$ is not too different from $\psi_0(\mathbf{r})$ in the scattering region. It follows, from Eq. (15.9), that the condition for $\psi(\mathbf{r}) \simeq \psi_0(\mathbf{r})$ in the vicinity of $\mathbf{r} = \mathbf{0}$ is

$$\left|\frac{m}{2\pi\, \hbar^2} \int \frac{\exp(\,i\, k\, r')}{r'}\, V(\mathbf{r}')\, d^3\mathbf{r}'\right| \ll 1. \tag{15.29}$$

Consider the special case of the Yukawa potential. At low energies, (*i.e.*, $k \ll \mu$) we can replace $\exp(\,i\, k\, r')$ by unity, giving

$$\frac{2\, m}{\hbar^2} \frac{|V_0|}{\mu^2} \ll 1 \tag{15.30}$$

as the condition for the validity of the Born approximation. The condition for the Yukawa potential to develop a bound state is

$$\frac{2\, m}{\hbar^2} \frac{|V_0|}{\mu^2} \geq 2.7, \tag{15.31}$$

where V_0 is negative. Thus, if the potential is strong enough to form a bound state then the Born approximation is likely to break down. In the high-k limit, Eq. (15.29) yields

$$\frac{2\, m}{\hbar^2} \frac{|V_0|}{\mu\, k} \ll 1. \tag{15.32}$$

This inequality becomes progressively easier to satisfy as k increases, implying that the Born approximation is more accurate at *high* incident particle energies.

15.4 Partial Waves

We can assume, without loss of generality, that the incident wavefunction is characterized by a wavevector $\mathbf{k}$ which is aligned parallel to the z-axis. The scattered wavefunction is characterized by a wavevector $\mathbf{k}'$ which has the same magnitude as $\mathbf{k}$, but, in general, points in a different direction. The direction of $\mathbf{k}'$ is specified by the polar angle θ (*i.e.*, the angle subtended between the two wavevectors), and an azimuthal angle ϕ about the z-axis. Equations (15.20) and (15.21) strongly suggest that for a spherically symmetric scattering potential [*i.e.*, $V(\mathbf{r}) = V(r)$] the scattering amplitude is a function of θ only: *i.e.*,

$$f(\theta, \phi) = f(\theta). \tag{15.33}$$

It follows that neither the incident wavefunction,

$$\psi_0(\mathbf{r}) = \sqrt{n}\,\exp(\,i\,k\,z) = \sqrt{n}\,\exp(\,i\,k\,r\cos\theta), \tag{15.34}$$

nor the large-r form of the total wavefunction,

$$\psi(\mathbf{r}) = \sqrt{n}\left[\exp(\,i\,k\,r\cos\theta) + \frac{\exp(\,i\,k\,r)\,f(\theta)}{r}\right], \tag{15.35}$$

depend on the azimuthal angle ϕ.

Outside the range of the scattering potential, both $\psi_0(\mathbf{r})$ and $\psi(\mathbf{r})$ satisfy the free space Schrödinger equation

$$(\nabla^2 + k^2)\,\psi = 0. \tag{15.36}$$

What is the most general solution to this equation in spherical polar coordinates which does not depend on the azimuthal angle ϕ? Separation of variables yields

$$\psi(r, \theta) = \sum_l R_l(r)\,P_l(\cos\theta), \tag{15.37}$$

since the *Legendre functions* $P_l(\cos\theta)$ form a complete set in θ-space. The Legendre functions are related to the *spherical harmonics*, introduced in Cha. 8, via

$$P_l(\cos\theta) = \sqrt{\frac{4\pi}{2\,l+1}}\,Y_{l,0}(\theta, \varphi). \tag{15.38}$$

Equations (15.36) and (15.37) can be combined to give

$$r^2\,\frac{d^2R_l}{dr^2} + 2\,r\,\frac{dR_l}{dr} + [k^2\,r^2 - l\,(l+1)]R_l = 0. \tag{15.39}$$

The two independent solutions to this equation are the *spherical Bessel functions*, $j_l(k\,r)$ and $y_l(k\,r)$, introduced in Sect. 9.3. Recall that

$$j_l(z) = z^l\left(-\frac{1}{z}\frac{d}{dz}\right)^l\left(\frac{\sin z}{z}\right), \tag{15.40}$$

$$y_l(z) = -z^l\left(-\frac{1}{z}\frac{d}{dz}\right)^l\left(\frac{\cos z}{z}\right). \tag{15.41}$$

Note that the $j_l(z)$ are well-behaved in the limit $z \to 0$, whereas the $y_l(z)$ become singular. The asymptotic behaviour of these functions in the limit $z \to \infty$ is

$$j_l(z) \rightarrow \frac{\sin(z - l\,\pi/2)}{z}, \tag{15.42}$$

$$y_l(z) \rightarrow -\frac{\cos(z - l\,\pi/2)}{z}. \tag{15.43}$$

We can write

$$\exp(\,i\,k\,r\cos\theta) = \sum_l a_l\,j_l(k\,r)\,P_l(\cos\theta), \tag{15.44}$$

where the a_l are constants. Note there are no $y_l(k\,r)$ functions in this expression, because they are not well-behaved as $r \to 0$. The Legendre functions are orthonormal,

$$\int_{-1}^{1} P_n(\mu)\,P_m(\mu)\,d\mu = \frac{\delta_{nm}}{n + 1/2}, \tag{15.45}$$

so we can invert the above expansion to give

$$a_l\,j_l(k\,r) = (l + 1/2)\int_{-1}^{1} \exp(\,i\,k\,r\,\mu)\,P_l(\mu)\,d\mu. \tag{15.46}$$

It is well-known that

$$j_l(y) = \frac{(-i)^l}{2}\int_{-1}^{1} \exp(\,i\,y\,\mu)\,P_l(\mu)\,d\mu, \tag{15.47}$$

where $l = 0, 1, 2, \cdots$ [see M. Abramowitz and I.A. Stegun, *Handbook of mathematical functions*, (Dover, New York NY, 1965), Eq. 10.1.14]. Thus,

$$a_l = i^{\,l}\,(2\,l + 1), \tag{15.48}$$

giving

$$\psi_0(\mathbf{r}) = \sqrt{n}\,\exp(\,i\,k\,r\cos\theta) = \sqrt{n}\,\sum_l i^{\,l}\,(2\,l + 1)\,j_l(k\,r)\,P_l(\cos\theta). \tag{15.49}$$

The above expression tells us how to decompose the incident plane-wave into a series of spherical waves. These waves are usually termed “partial waves”.

The most general expression for the total wavefunction outside the scattering region is

$$\psi(\mathbf{r}) = \sqrt{n}\sum_l\,[A_l\,j_l(k\,r) + B_l\,y_l(k\,r)]\,P_l(\cos\theta), \tag{15.50}$$

where the A_l and B_l are constants. Note that the $y_l(k\,r)$ functions are allowed to appear in this expansion, because its region of validity does not include the origin. In the large-r limit, the total wavefunction reduces to

$$\psi(\mathbf{r}) \simeq \sqrt{n}\sum_l\left[A_l\,\frac{\sin(k\,r - l\,\pi/2)}{k\,r} - B_l\,\frac{\cos(k\,r - l\,\pi/2)}{k\,r}\right]P_l(\cos\theta), \tag{15.51}$$

where use has been made of Eqs. (15.42) and (15.43). The above expression can also be written

$$\psi(\mathbf{r}) \simeq \sqrt{n} \sum_l C_l \frac{\sin(k\,r - l\,\pi/2 + \delta_l)}{k\,r} P_l(\cos\theta), \tag{15.52}$$

where the sine and cosine functions have been combined to give a sine function which is phase-shifted by δ_l. Note that $A_l = C_l \cos\delta_l$ and $B_l = -C_l \sin\delta_l$.

Equation (15.52) yields

$$\psi(\mathbf{r}) \simeq \sqrt{n} \sum_l C_l \left[\frac{e^{\,i\,(k\,r - l\,\pi/2 + \delta_l)} - e^{-i\,(k\,r - l\,\pi/2 + \delta_l)}}{2\,i\,k\,r}\right] P_l(\cos\theta), \tag{15.53}$$

which contains both incoming and outgoing spherical waves. What is the source of the incoming waves? Obviously, they must be part of the large-r asymptotic expansion of the incident wavefunction. In fact, it is easily seen from Eqs. (15.42) and (15.49) that

$$\psi_0(\mathbf{r}) \simeq \sqrt{n} \sum_l i^{\,l}\,(2l+1) \left[\frac{e^{\,i\,(k\,r - l\,\pi/2)} - e^{-i\,(k\,r - l\,\pi/2)}}{2\,i\,k\,r}\right] P_l(\cos\theta) \tag{15.54}$$

in the large-r limit. Now, Eqs. (15.34) and (15.35) give

$$\frac{\psi(\mathbf{r}) - \psi_0(\mathbf{r})}{\sqrt{n}} = \frac{\exp(\,i\,k\,r)}{r}\, f(\theta). \tag{15.55}$$

Note that the right-hand side consists of an *outgoing* spherical wave only. This implies that the coefficients of the incoming spherical waves in the large-r expansions of $\psi(\mathbf{r})$ and $\psi_0(\mathbf{r})$ must be the same. It follows from Eqs. (15.53) and (15.54) that

$$C_l = (2\,l+1)\,\exp[\,i\,(\delta_l + l\,\pi/2)]. \tag{15.56}$$

Thus, Eqs. (15.53)–(15.55) yield

$$f(\theta) = \sum_{l=0}^{\infty} (2\,l+1)\,\frac{\exp(\,i\,\delta_l)}{k}\,\sin\delta_l\,P_l(\cos\theta). \tag{15.57}$$

Clearly, determining the scattering amplitude $f(\theta)$ via a decomposition into partial waves (*i.e.*, spherical waves) is equivalent to determining the phase-shifts δ_l.

Now, the differential scattering cross-section $d\sigma/d\Omega$ is simply the modulus squared of the scattering amplitude $f(\theta)$ [see Eq. (15.17)]. The total cross-section is thus given by

$$\begin{aligned}
\sigma_{\rm total} &= \int |f(\theta)|^2\,d\Omega \\
&= \frac{1}{k^2}\oint d\phi \int_{-1}^{1} d\mu \sum_l \sum_{l'} (2\,l+1)\,(2\,l'+1)\,\exp[\,i\,(\delta_l - \delta_{l'})] \\
&\qquad \times \sin\delta_l\,\sin\delta_{l'}\,P_l(\mu)\,P_{l'}(\mu),
\end{aligned} \tag{15.58}$$

where $\mu = \cos\theta$. It follows that

$$\sigma_{\text{total}} = \frac{4\pi}{k^2} \sum_l (2\,l+1)\,\sin^2\delta_l, \tag{15.59}$$

where use has been made of Eq. (15.45).

15.5 Determination of Phase-Shifts

Let us now consider how the phase-shifts δ_l in Eq. (15.57) can be evaluated. Consider a spherically symmetric potential $V(r)$ which vanishes for $r > a$, where a is termed the *range* of the potential. In the region $r > a$, the wavefunction $\psi(\mathbf{r})$ satisfies the free-space Schrödinger equation (15.36). The most general solution which is consistent with no incoming spherical-waves is

$$\psi(\mathbf{r}) = \sqrt{n} \sum_{l=0}^{\infty} i^{\,l}\,(2\,l+1)\,\mathcal{R}_l(r)\,P_l(\cos\theta), \tag{15.60}$$

where

$$\mathcal{R}_l(r) = \exp(\,i\,\delta_l)\;[\cos\delta_l\,j_l(k\,r) - \sin\delta_l\,y_l(k\,r)]\,. \tag{15.61}$$

Note that $y_l(k\,r)$ functions are allowed to appear in the above expression, because its region of validity does not include the origin (where $V \neq 0$). The logarithmic derivative of the lth radial wavefunction, $\mathcal{R}_l(r)$, just outside the range of the potential is given by

$$\beta_{l+} = k\,a\left[\frac{\cos\delta_l\,j_l'(k\,a) - \sin\delta_l\,y_l'(k\,a)}{\cos\delta_l\,j_l(k\,a) - \sin\delta_l\,y_l(k\,a)}\right], \tag{15.62}$$

where $j_l'(x)$ denotes $dj_l(x)/dx$, *etc.* The above equation can be inverted to give

$$\tan\delta_l = \frac{k\,a\,j_l'(k\,a) - \beta_{l+}\,j_l(k\,a)}{k\,a\,y_l'(k\,a) - \beta_{l+}\,y_l(k\,a)}. \tag{15.63}$$

Thus, the problem of determining the phase-shift δ_l is equivalent to that of obtaining β_{l+}.

The most general solution to Schrödinger's equation inside the range of the potential ($r < a$) which does not depend on the azimuthal angle ϕ is

$$\psi(\mathbf{r}) = \sqrt{n} \sum_{l=0}^{\infty} i^{\,l}\,(2\,l+1)\,\mathcal{R}_l(r)\,P_l(\cos\theta), \tag{15.64}$$

where

$$\mathcal{R}_l(r) = \frac{u_l(r)}{r}, \tag{15.65}$$

and

$$\frac{d^2u_l}{dr^2} + \left[k^2 - \frac{l\,(l+1)}{r^2} - \frac{2\,m}{\hbar^2}\,V\right]u_l = 0. \tag{15.66}$$

The boundary condition

$$u_l(0) = 0 \tag{15.67}$$

ensures that the radial wavefunction is well-behaved at the origin. We can launch a well-behaved solution of the above equation from $r = 0$, integrate out to $r = a$, and form the logarithmic derivative

$$\beta_{l-} = \frac{1}{(u_l/r)} \frac{d(u_l/r)}{dr}\bigg|_{r=a}. \tag{15.68}$$

Since $\psi(\mathbf{r})$ and its first derivatives are necessarily continuous for physically acceptible wavefunctions, it follows that

$$\beta_{l+} = \beta_{l-}. \tag{15.69}$$

The phase-shift δ_l is then obtainable from Eq. (15.63).

15.6 Hard Sphere Scattering

Let us test out this scheme using a particularly simple example. Consider scattering by a *hard sphere*, for which the potential is infinite for $r < a$, and zero for $r > a$. It follows that $\psi(\mathbf{r})$ is zero in the region $r < a$, which implies that $u_l = 0$ for all l. Thus,

$$\beta_{l-} = \beta_{l+} = \infty, \tag{15.70}$$

for all l. Equation (15.63) thus gives

$$\tan \delta_l = \frac{j_l(k\,a)}{y_l(k\,a)}. \tag{15.71}$$

Consider the $l = 0$ partial wave, which is usually referred to as the S-wave. Equation (15.71) yields

$$\tan \delta_0 = \frac{\sin(k\,a)/k\,a}{-\cos(k\,a)/ka} = -\tan(k\,a), \tag{15.72}$$

where use has been made of Eqs. (15.40) and (15.41). It follows that

$$\delta_0 = -k\,a. \tag{15.73}$$

The S-wave radial wave function is [see Eq. (15.61)]

$$\begin{aligned}\mathcal{R}_0(r) &= \exp(-i\,k\,a)\frac{[\cos(k\,a)\,\sin(k\,r) - \sin(k\,a)\,\cos(k\,r)]}{k\,r} \\ &= \exp(-i\,k\,a)\,\frac{\sin[k\,(r-a)]}{k\,r}.\end{aligned} \tag{15.74}$$

The corresponding radial wavefunction for the incident wave takes the form [see Eq. (15.49)]

$$\tilde{\mathcal{R}}_0(r) = \frac{\sin(k\,r)}{k\,r}. \tag{15.75}$$

Thus, the actual $l = 0$ radial wavefunction is similar to the incident $l = 0$ wavefunction, except that it is phase-shifted by $k\,a$.

Let us examine the low and high energy asymptotic limits of $\tan\delta_l$. Low energy implies that $k\,a \ll 1$. In this regime, the spherical Bessel functions reduce to:

$$j_l(k\,r) \simeq \frac{(k\,r)^l}{(2\,l+1)!!}, \tag{15.76}$$

$$y_l(k\,r) \simeq -\frac{(2\,l-1)!!}{(k\,r)^{l+1}}, \tag{15.77}$$

where $n!! = n\,(n-2)\,(n-4)\cdots 1$. It follows that

$$\tan\delta_l = \frac{-(k\,a)^{2\,l+1}}{(2\,l+1)\,[(2\,l-1)!!]^{\,2}}. \tag{15.78}$$

It is clear that we can neglect δ_l, with $l > 0$, with respect to δ_0. In other words, at low energy only S-wave scattering (*i.e.*, spherically symmetric scattering) is important. It follows from Eqs. (15.17), (15.57), and (15.73) that

$$\frac{d\sigma}{d\Omega} = \frac{\sin^2 k\,a}{k^2} \simeq a^2 \tag{15.79}$$

for $k\,a \ll 1$. Note that the total cross-section

$$\sigma_{\rm total} = \int \frac{d\sigma}{d\Omega}\,d\Omega = 4\pi\,a^2 \tag{15.80}$$

is *four times* the *geometric cross-section* $\pi\,a^2$ (*i.e.*, the cross-section for classical particles bouncing off a hard sphere of radius a). However, low energy scattering implies relatively long wavelengths, so we would not expect to obtain the classical result in this limit.

Consider the high energy limit $k\,a \gg 1$. At high energies, all partial waves up to $l_{\rm max} = k\,a$ contribute significantly to the scattering cross-section. It follows from Eq. (15.59) that

$$\sigma_{\rm total} \simeq \frac{4\pi}{k^2}\sum_{l=0}^{l_{\rm max}}(2\,l+1)\,\sin^2\delta_l. \tag{15.81}$$

With so many l values contributing, it is legitimate to replace $\sin^2\delta_l$ by its average value $1/2$. Thus,

$$\sigma_{\rm total} \simeq \sum_{l=0}^{k\,a}\frac{2\pi}{k^2}\,(2\,l+1) \simeq 2\pi\,a^2. \tag{15.82}$$

This is *twice* the classical result, which is somewhat surprizing, since we might expect to obtain the classical result in the short wavelength limit. For hard sphere scattering, incident waves with impact parameters less than a must be deflected. However, in order to produce a "shadow" behind the sphere, there must also be some scattering in the forward direction in order to produce destructive interference with the incident plane-wave. In fact, the interference is not completely destructive, and the shadow has a bright spot (the so-called "Poisson spot") in the forward direction. The effective cross-section associated with this bright spot is πa^2 which, when combined with the cross-section for classical reflection, πa^2, gives the actual cross-section of $2\pi a^2$.

15.7 Low Energy Scattering

In general, at low energies (*i.e.*, when $1/k$ is much larger than the range of the potential) partial waves with $l > 0$ make a negligible contribution to the scattering cross-section. It follows that, at these energies, with a finite range potential, only S-wave scattering is important.

As a specific example, let us consider scattering by a finite potential well, characterized by $V = V_0$ for $r < a$, and $V = 0$ for $r \geq a$. Here, V_0 is a constant. The potential is repulsive for $V_0 > 0$, and attractive for $V_0 < 0$. The outside wavefunction is given by [see Eq. (15.61)]

$$\begin{aligned}\mathcal{R}_0(r) &= \exp(\,i\,\delta_0)\,[\cos\delta_0\,j_0(k\,r) - \sin\delta_0\,y_0(k\,r)]\\ &= \frac{\exp(\,i\,\delta_0)\,\sin(k\,r+\delta_0)}{k\,r},\end{aligned} \tag{15.83}$$

where use has been made of Eqs. (15.40) and (15.41). The inside wavefunction follows from Eq. (15.66). We obtain

$$\mathcal{R}_0(r) = B\,\frac{\sin(k'\,r)}{r}, \tag{15.84}$$

where use has been made of the boundary condition (15.67). Here, B is a constant, and

$$E - V_0 = \frac{\hbar^2\,k'^2}{2\,m}. \tag{15.85}$$

Note that Eq. (15.84) only applies when $E > V_0$. For $E < V_0$, we have

$$\mathcal{R}_0(r) = B\,\frac{\sinh(\kappa\,r)}{r}, \tag{15.86}$$

where

$$V_0 - E = \frac{\hbar^2\kappa^2}{2\,m}. \tag{15.87}$$

Matching $\mathcal{R}_0(r)$, and its radial derivative, at $r = a$ yields

$$\tan(k\,a+\delta_0) = \frac{k}{k'}\,\tan(k'\,a) \tag{15.88}$$

for $E > V_0$, and

$$\tan(k\,a+\delta_0) = \frac{k}{\kappa}\,\tanh(\kappa\,a) \tag{15.89}$$

for $E < V_0$.

Consider an attractive potential, for which $E > V_0$. Suppose that $|V_0| \gg E$ (*i.e.*, the depth of the potential well is much larger than the energy of the incident particles), so that $k' \gg k$. We can see from Eq. (15.88) that, unless $\tan(k'\,a)$ becomes extremely large, the right-hand side is much less that unity, so replacing the tangent of a small quantity with the quantity itself, we obtain

$$k\,a + \delta_0 \simeq \frac{k}{k'}\,\tan(k'\,a). \tag{15.90}$$

This yields

$$\delta_0 \simeq k\,a\left[\frac{\tan(k'\,a)}{k'\,a} - 1\right]. \tag{15.91}$$

According to Eq. (15.81), the scattering cross-section is given by

$$\sigma_{\rm total} \simeq \frac{4\pi}{k^2}\,\sin^2\delta_0 = 4\pi\,a^2\left[\frac{\tan(k'a)}{k'\,a} - 1\right]^2. \tag{15.92}$$

Now

$$k'\,a = \sqrt{k^2\,a^2 + \frac{2\,m\,|V_0|\,a^2}{\hbar^2}}, \tag{15.93}$$

so for sufficiently small values of $k\,a$,

$$k'\,a \simeq \sqrt{\frac{2\,m\,|V_0|\,a^2}{\hbar^2}}. \tag{15.94}$$

It follows that the total (S-wave) scattering cross-section is *independent* of the energy of the incident particles (provided that this energy is sufficiently small).

Note that there are values of $k'\,a$ (*e.g.*, $k'\,a \simeq 4.49$) at which $\delta_0 \to \pi$, and the scattering cross-section (15.92) vanishes, despite the very strong attraction of the potential. In reality, the cross-section is not exactly zero, because of contributions from $l > 0$ partial waves. But, at low incident energies, these contributions are small. It follows that there are certain values of V_0 and k which give rise to almost perfect transmission of the incident wave. This is called the *Ramsauer-Townsend effect*, and has been observed experimentally.

15.8 Resonances

There is a significant exception to the independence of the cross-section on energy mentioned above. Suppose that the quantity $\sqrt{2\,m\,|V_0|\,a^2/\hbar^2}$ is slightly less than $\pi/2$. As the incident energy increases, $k'\,a$, which is given by Eq. (15.93), can reach the value $\pi/2$. In this case, $\tan(k'\,a)$ becomes infinite, so we can no longer assume that the right-hand side of Eq. (15.88) is small. In fact, it follows from Eq. (15.88) that at the value of the incident energy when $k' a = \pi/2$ then we also have $k a + \delta_0 = \pi/2$, or $\delta_0 \simeq \pi/2$ (since we are assuming that $k a \ll 1$). This implies that

$$\sigma_{\rm total} = \frac{4\pi}{k^2} \sin^2 \delta_0 = 4\pi\, a^2 \left(\frac{1}{k^2 a^2}\right). \tag{15.95}$$

Note that the cross-section now depends on the energy. Furthermore, the magnitude of the cross-section is much larger than that given in Eq. (15.92) for $k' a \neq \pi/2$ (since $k a \ll 1$).

The origin of this rather strange behaviour is quite simple. The condition

$$\sqrt{\frac{2\, m\, |V_0|\, a^2}{\hbar^2}} = \frac{\pi}{2} \tag{15.96}$$

is equivalent to the condition that a spherical well of depth V_0 possesses a *bound state* at zero energy. Thus, for a potential well which satisfies the above equation, the energy of the scattering system is essentially the same as the energy of the bound state. In this situation, an incident particle would like to form a bound state in the potential well. However, the bound state is not stable, since the system has a small positive energy. Nevertheless, this sort of *resonance scattering* is best understood as the capture of an incident particle to form a metastable bound state, and the subsequent decay of the bound state and release of the particle. The cross-section for resonance scattering is generally *much larger* than that for non-resonance scattering.

We have seen that there is a resonant effect when the phase-shift of the S-wave takes the value $\pi/2$. There is nothing special about the $l = 0$ partial wave, so it is reasonable to assume that there is a similar resonance when the phase-shift of the lth partial wave is $\pi/2$. Suppose that δ_l attains the value $\pi/2$ at the incident energy E_0, so that

$$\delta_l(E_0) = \frac{\pi}{2}. \tag{15.97}$$

Let us expand $\cot\delta_l$ in the vicinity of the resonant energy:

$$\begin{aligned} \cot\delta_l(E) &= \cot\delta_l(E_0) + \left(\frac{d\cot\delta_l}{dE}\right)_{E=E_0} (E - E_0) + \cdots \\ &= -\left(\frac{1}{\sin^2\delta_l}\frac{d\delta_l}{dE}\right)_{E=E_0} (E - E_0) + \cdots. \end{aligned} \tag{15.98}$$

Defining

$$\left(\frac{d\delta_l(E)}{dE}\right)_{E=E_0} = \frac{2}{\Gamma}, \tag{15.99}$$

we obtain

$$\cot\delta_l(E) = -\frac{2}{\Gamma}\,(E - E_0) + \cdots. \tag{15.100}$$

Recall, from Eq. (15.59), that the contribution of the lth partial wave to the scattering cross-section is

$$\sigma_l = \frac{4\pi}{k^2}\,(2\,l+1)\,\sin^2\delta_l = \frac{4\pi}{k^2}\,(2\,l+1)\,\frac{1}{1+\cot^2\delta_l}. \tag{15.101}$$

Thus,

$$\sigma_l \simeq \frac{4\pi}{k^2}\,(2\,l+1)\,\frac{\Gamma^2/4}{(E-E_0)^2+\Gamma^2/4}. \tag{15.102}$$

This is the famous *Breit-Wigner formula*. The variation of the partial cross-section σ_l with the incident energy has the form of a classical *resonance curve*. The quantity Γ is the width of the resonance (in energy). We can interpret the Breit-Wigner formula as describing the absorption of an incident particle to form a metastable state, of energy E_0, and lifetime $\tau = \hbar/\Gamma$.