MONASH UNIVERSITY

SCHOOL OF PHYSICS

QUANTUM MECHANICS

PHS3031

David Paganin, 2012

David Paganin
Room 109, Building 26
Email: David.Paganin@monash.edu
Ph: 9905 3633

CHAPTER 1
INTRODUCTION

§1. Introduction

Suppose one were to plant a lemon tree in one's garden, subsequently noting that the tree grew a distance of one metre in one year. Given this observation, it would evidently be unwise to conclude that the lemon tree will grow a million metres in a million years.

The above unwise conclusion—that the lemon tree will grow a million metres in a million years—is an *unwarranted extrapolation* of the initial observation that the lemon tree grew one metre in one year. Stated differently, one cannot assume that a given trend, over a given period of time (in this case, one year) will continue over a period of time that is many orders of magnitude (in this case, six) longer.

The history of physics teaches us to be wary of unwarranted extrapolations. Four examples:

- Observation: To the naked eye of a human standing on the ground, the part of the earth that we can see appears to be flat. Unwarranted extrapolation: "Therefore the whole earth is flat."
- Observation: To the naked eye, matter (such as water, metal, wood *etc*.) appears to be a continuum. Unwarranted extrapolation: "Therefore matter is continuous down to arbitrarily small length scales."[1]
- Observation: If I apply a force to an object, that object accelerates at a rate proportional to the net applied force. Unwarranted extrapolation: "If the force is applied for an arbitrarily long time, the object will move arbitrarily quickly."[2]
- Observation: The laws of classical mechanics are applicable to mass scales, time scales and length scales that are within the ken of the everyday experience of humans unaided by instruments. These laws work beautifully in describing the behaviour of billiard balls, springs and bridges. Unwarranted extrapolation: "The laws of classical mechanics are applicable to phenomena such as atoms (which are many orders of magnitude lighter and many orders of magnitude smaller than anything within the ken of everyday human experience)".

Let us expand a little on the last dot-point listed above: The smallest mass that our senses can discern is on the order of milligram or so; the smallest time-scale that our senses can be sensitive to is on the order of a thousandth of a second; the smallest distance that the naked eye can see is on the order of a micron. Such mass-, time- and length-scales are many orders of magnitude longer that the corresponding characteristic scales in the atomic and sub-atomic domain (e.g. the electron mass is approximately 10^{-30} kg, strong interactions have characteristic time-scales on the order of 10^{-23} s, and atomic dimensions are on the order of 10^{-10} m).

There is no reason to make the extrapolation that Newtonian classical mechanics—whose domain of validity is well established for a vast array of phenomena over mass-, length- and time-scales within our everyday experience—should also be applicable at the level of atoms. Indeed, *classical mechanics breaks down at the atomic level*. To better describe such systems, quantum mechanics is required.

§2. Some watershed moments in the early years of quantum mechanics

In the following three subsections, we very briefly outline several examples in which classical mechanics breaks down at the atomic level. The successful application of the ideas of quantum mechanics, to such systems, constituted some early triumphs of the emerging quantum theory.

§2.1 Photoelectric effect

Following on from some experiments by Hertz in the late nineteenth century, Lenard investigated the so-called photoelectric effect using the apparatus shown in Figure 1. Here, we see an electric circuit containing a battery with potential difference V across its electrodes, a resistor R and an evacuated glass tube T. Within this glass tube are polished metal cathode and anode plates, respectively denoted C and A. Ultraviolet light passes from a source S, through a quartz window W, before striking the cathode C.

---

[1] The extrapolation is incorrect because atoms exist!

[2] According to Special Relativity, no object can move arbitrarily quickly, as the speed of light *in vacuo* forms a "natural speed limit".

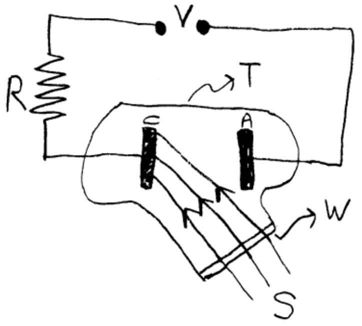


Figure 1: Photoelectric effect

In the absence of any illumination, no current was seen to flow through the circuit. In the presence of illumination by the ultraviolet light source, a current was seen to flow through the circuit. This indicates that electrons are emitted from C and travel to A. Such electrons are known as “photo-electrons”.

The following additional observations were made:

- The magnitude of the current is proportional to the intensity of the light striking the metal;
- The speed of the ejected electrons is independent of the intensity of the light striking the metal;
- The speed of the ejected electrons is dependent on the frequency of the incident ultraviolet light;
- The rate, at which electrons are emitted, is proportional to the intensity of the incident ultraviolet light;
- Ejected electrons are observed as soon as the ultraviolet lamp is turned on, even when the intensity of the lamp is very low.

In a very famous paper from 1905, Einstein developed a simple and powerful theory for describing the photoelectric effect. He considered the energy of the incident light to be “chopped up” into discrete packets or quanta of energy, which are now known as photons:

> “... the energy of a light ray spreading out from a point source is not continuously distributed over an increasing space but consists of a finite number of energy quanta which are localized at points in space, which move without dividing, and which can only be produced and absorbed as complete units”[3].

Each of these photons was considered to have an energy E given by:

(1) $E = h\nu$,

where h is Planck’s constant and ν is the frequency of the radiation. Assuming that an energy W (“work function”) is required to dislodge an electron from the cathode, conservation of energy then allowed Einstein to write[4]:

(2) $\frac{1}{2} m_e v^2 = h\nu - W$.

Here, m_e is the mass of the electron and v is its speed.

Equation (2) was seen to be in accord with the findings of experiment.

► *Exercise #1. Please explain how the four bullet points, listed earlier in this section, harmonise with Einstein’s photon concept.*

We see, here, our first example of *quantisation*. In this case, experiment is interpreted as implying a certain “granularity” in the electromagnetic field, with its energy being “chopped up” into discrete quanta (photons), rather than being continuously smeared through space in accord with classical electrodynamics.

Further examples, of quantisation, are given in the following two sub-sections.

§2.2 Compton effect

Another early triumph of Einstein’s photon concept lay in its successful application to the scattering of X-rays from stationary electrons[5]. This scenario, now known as “Compton scattering”, is sketched in Fig. 2.

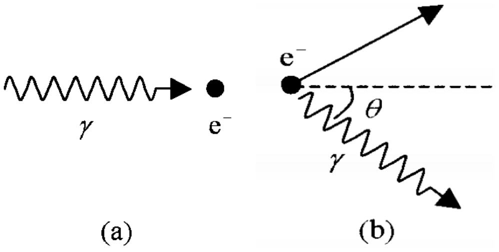


Figure 2: Compton scattering of an X-ray photon by a free stationary electron[6]

---

[3] A. Einstein, Ann. Phys. **17**, 132–148 (1905). The English translation, cited in the text, is given in A.B. Arons and M.B. Peppard (trans.), *Einstein's proposal of the photon concept—a translation of the Annalen der Physik paper of 1905*, American Journal of Physics **33**, 367–374 (1965).

[4] Note that (2) only holds if the right side is greater than or equal to zero.

[5] Prior analyses, using classical electromagnetic theory, were ultimately unsuccessful.

[6] Image taken from D.M. Paganin, *Coherent X-Ray Optics*, Oxford University Press, New York (2006), p. 116.

In part (a) of this figure, we see an X-ray photon γ, which is incident from the left upon a single stationary free electron e^-. After the interaction ("collision"), we have the scenario shown in part (b), with the electron recoiling and the photon being deflected through an angle θ. Since the electron has gained (kinetic) energy as a result of the collision, the photon must have lost energy. Since the energy of the scattered photon has been reduced, its wavelength must have increased.

By invoking both conservation of energy and conservation of momentum, and employing Einstein's photon concept, Compton was able to derive the following formula ("Compton formula"):

$$(3) \quad \lambda' - \lambda = \frac{2h}{m_e c}\sin^2\left(\frac{\theta}{2}\right).$$

Here, λ denotes the wavelength of the photon prior to the collision, λ' denotes the wavelength of the photon after the collision, and c denotes the speed of light.

Note, also, that the energy E and wavelength λ of the photon are related through the formula:

$$(4) \quad E = \frac{hc}{\lambda}.$$

► *Exercise #2. Derive equations (3) and (4). Briefly interpret equation (3). Note that this calculation won Compton a Nobel prize!*

§2.3 Atomic spectra

Consider the scenario shown in Figure 3. Here we see a sealed container A, whose walls enclose a hot elemental gas B. This gas is sufficiently hot to be glowing—i.e. to be emitting electromagnetic radiation in the form of visible light. Upon passing through a window C in the side of the container, the emitted light D is refracted by a prism E, with the refracted light F falling upon a film G.

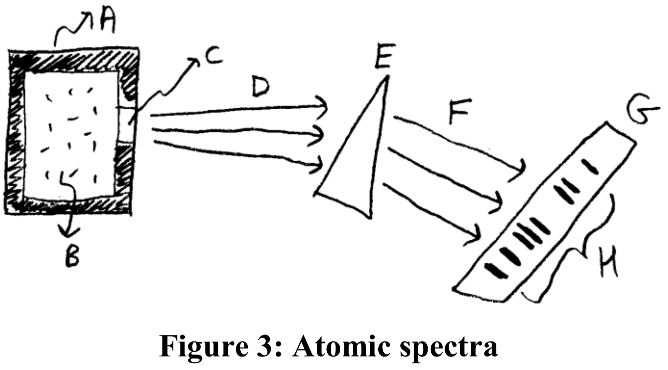


Figure 3: Atomic spectra

A rainbow-like continuum of radiation is *not* observed over the film. Rather, a series of *discrete lines* H, known as "spectral lines", are observed. Different sets of spectral lines are observed when different elements are used in the incandescent gas.

The presence of discrete spectral lines, each of which correspond to a given energy, is suggestive of the *quantisation of energy levels of the atoms in the incandescent gas.* Indeed, if the allowed energies of the atoms are quantised—that is, they are constrained to take on only the discrete values E_m, where m is an integer—then conservation of energy implies the energy E_γ of the emitted photons to also be quantised according to:

$$(5) \quad E_m - E_n = E_\gamma.$$

Here, m and n are integers, and $E_m \geq E_n$.

► *Exercise #3. Derive the central result of Bohr's model of the hydrogen atom*[7]*, namely:*

$$(6) \quad E_n = -\frac{m_e e^4}{32\pi^2\varepsilon_0^2\hbar^2 n^2}, \quad n = 1,2,3...$$

Here, E_n *denotes the* n*th quantised energy level of the hydrogen atom,* e *denotes the charge on an electron,* ε_0 *is the permittivity of free space, and*

$$(7) \quad \hbar \equiv \frac{h}{2\pi}$$

is known as the "reduced Planck constant".

► *Exercise #4. In 1885, nine hydrogen spectral lines were known. Balmer noted that the wavelengths* λ_p *of these nine lines could be empirically fitted to the following formula:*

$$(8) \quad \lambda_p = \frac{Cp^2}{p^2 - 4}, \quad p = \text{integer}.$$

Here, C *is a constant. Please show that the above relation is a direct consequence of Bohr's theory of*

---

[7] (a) Many of you will have met the Bohr theory in an earlier physics course, hence the brevity of this question! ☺
(b) Notwithstanding its early triumph in describing the energy levels of the hydrogen atom, the Bohr theory—which occupied an important place in the so-called "Old Quantum Theory"—was seen to be of limited utility in studying more complex quantum systems. In chapter 7 of the course, we will revisit our study of the hydrogen atom, from the more sophisticated perspective of the Schrödinger equation.

the hydrogen atom. Derive an expression for C, *and then obtain a suitable generalization of (8).*

§3. Brief outline of remainder of course

Having motivated the idea that a new mechanics—quantum mechanics—is needed for an analysis of the mechanics of very small systems such as atoms and photons, we are ready to give a brief outline of how this new mechanics will be treated in the following chapters.

Chapter 2 introduces the central concept of the wave-function, which is commonly denoted by the Greek letter Ψ ("psi"). The wavefunction is a function which completely defines the state of a given quantum-mechanical system. Typically, this is a function of both position and time, together with other variables that will be introduced in due course. Chapter 2 is also devoted to honing our physical interpretation of the wave function.

The third chapter introduces the famous Schrödinger equation, which governs the manner in which a system's wavefunction evolves through space and time. Two different forms of this equation will be written down: (a) a more general, time-dependent form, and (b) a time-independent form, which is less general than its time-dependent counterpart, and which may be applied to the study of mono-energetic quantum systems. We shall also briefly investigate the transition from quantum to classical mechanics—after all, the new theory (quantum mechanics) must reduce to the old theory (classical mechanics) for physical scenarios which lie within the domain of validity of the older, less general theory. Some further implications of the Schrödinger equation will also be explored in this chapter, such as conservation of particles, together with the role of operators in the quantum theory.

Chapter 4 discusses several one-dimensional quantum systems, including the free particle, the potential step and potential barrier, the square well and the linear harmonic oscillator. We will also introduce the notion of wave-packet revivals [in an assignment question, rather than in these notes!!], together with the very powerful notion of creation and destruction operators.

Chapter 5 develops a powerful and general mathematical framework which places the theory of quantum mechanics on a more formal footing.

Chapter 6 is devoted to angular momentum and spin, with Chapter 7 being devoted to the study of three-dimensional quantum systems. A highlight of this latter chapter will be a treatment of the hydrogen atom from the perspective of the Schrödinger equation.

Finally, in Chapter 8 we will give a brief survey of more advanced studies in quantum mechanics, together with an indication of how the material of this lecture course fits into the broader context of theoretical physics as a whole.

► *Exercise #5. The maximum energy of photo-electrons emitted from potassium is 2.1 eV when illuminated by light of wavelength 0.3 microns, and 0.5 eV when illuminated by light of wavelength 0.5 microns. Use these results to obtain an estimate for the numerical value of Planck's constant.*

► *Exercise #6. An X-ray photon of wavelength* 1.0×10^{-12} *m is incident on a stationary free electron. Calculate the wavelength of the scattered photon if it is detected at an angle of (i) 60, (ii) 90, and (iii) 120 degrees to the incident photon.*

CHAPTER 2

ENTER THE WAVEFUNCTION

The "wavefunction", commonly denoted by the Greek letter Ψ, is a concept of central importance in quantum mechanics. The wavefunction is a function of position **r** and time t that is postulated to completely define the state of a given quantum system.

Here, we outline some of the principles that underly the quantum-mechanical notion of a wavefunction. We begin our discussions with an analysis of one of the simplest wavefunctions, namely the "plane wave". We then make use of the Fourier integral to construct superpositions of these plane waves. Lastly, we include some interpretive comments regarding the meaning of the wavefunction.

§4. The plane wave

Our first example of a wavefunction is the so-called "plane wave"—see Figure 4. In the present context, we shall assume this wavefunction to describe a single electron, with Figure 4 corresponding to a "snapshot" of this wavefunction at a given instant of time.

As mentioned above, the wavefunction is postulated to completely define the state of a given quantum-mechanical system. The aim of the present discussion is to introduce us to the wavefunction concept, in the context of the example of the plane-wave wavefunction for a single electron.

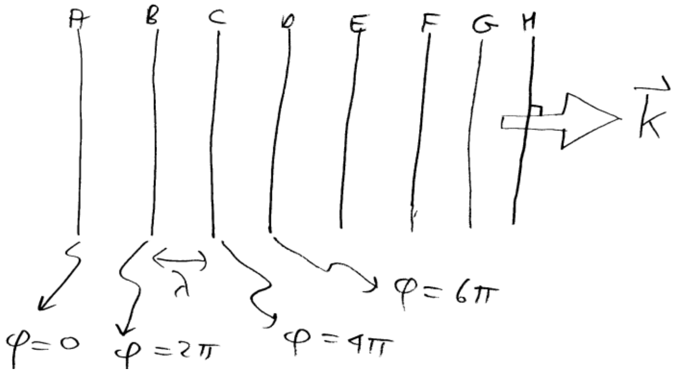


Figure 4: The plane wave

As shown in Figure 4, the plane wave has a series of parallel planar "wave-crests", labelled A, B, C, ..., with the distance between successive wave-crests being equal to the de-Broglie wavelength λ of the electron. The direction of propagation of the wave is given by the arrow in the figure, known as the "wave vector" **k.**[8]

As we have just pointed out, the *direction*, in which the wave-vector **k** points, is indicative of the direction in which the plane wave travels. The *magnitude* of the wave-vector **k**, denoted by $k = |\mathbf{k}|$, also has a physical meaning: the magnitude of the wavevector is inversely proportional to the wave-length of the radiation, and therefore inversely proportional to the spacing between successive wave-crests:

$$(9)\quad k = |\mathbf{k}| = \frac{2\pi}{\lambda}.$$

The wave-function, associated with our plane wave, is:

$$(10)\quad \Psi(\mathbf{r},t) = \exp[i(\mathbf{k}\bullet\mathbf{r} - \omega t)].$$

Here,

$$(11)\quad \omega = 2\pi\nu$$

is the angular frequency[9] corresponding to the frequency ν, $\mathbf{r} \equiv (x,y,z)$ is a position vector in three-dimensional space which corresponds to Cartesian coordinates (x,y,z), and t is time.

Note that the wave-crests, at a given instant of time, are associated with *surfaces of constant phase* at that given instant of time. (In the previous sentence, "phase" refers to the phase of the complex wavefunction.) The surfaces A, B, C, ..., in Figure 4, are examples of such surfaces of constant phase. The wavefront A corresponds to a phase of 0 radians—that is, the phase of the wavefunction is equal to zero radians at every point along the planar surface A. The wavefront B corresponds to a phase of 2π radians, C corresponds to a phase of 4π radians and so forth.

► *Exercise #7. The phase* $\varphi(\mathbf{r},t)$, *of the plane-wave wavefunction in (10), is given by:*

$$(12)\quad \varphi(\mathbf{r},t) = \mathbf{k}\bullet\mathbf{r} - \omega t.$$

(i) At any given instant of time, show that any surface of constant phase is a plane normal to the wave-vector **k**. *(ii) For a given fixed instant of time, show that the surfaces of constant phase w and w+2π, are separated by a distance of one de-Broglie wavelength, where w is any real number. (iii) Show that the wavefronts propagate in the direction of the wave-vector.*

Note: The above exercise serves to justify the constructions shown in Figure 4.

We close this section by noting a physical meaning of the plane-wave wavefunction in equation (10): it corresponds to a completely delocalised electron, with momentum:

$$(13)\quad \mathbf{p} = \hbar\mathbf{k}.$$

► *Exercise #8. Prove equation (13).*

§5. From plane waves to wave packets

In theoretical physics, it is often useful to express a given complicated "thing" (e.g. a number, a function, a sound wave, a light field, or a quantum-mechanical wavefunction) as a linear combination of "simpler things".

For example, in the decimal number system, we may express any real number as a linear combination/superposition of powers of 10, each coefficient being an integer. Thus the real number π, defined as the ratio of the circumference of any circle to its diameter, can be represented as:

---

[8] This vector is also termed the "propagation vector".

[9] Note that, via a very simple modification of equation (1)—obtained by dividing h by 2π and multiplying ν by 2π—we see that the angular frequency ω of the particle is related to its energy E via $E = \hbar\omega$.

$$\pi = (3\times10^{0}) + (1\times10^{-1}) + (4\times10^{-2}) + (1\times10^{-3}) + (5\times10^{-4}) + \cdots \quad (14)$$

The set $\{\cdots,10^{2},10^{1},10^{0},10^{-1},10^{-2},\cdots\}$ is said to form a "basis" for the real numbers, because *any real number may be expressed as a linear combination/superposition of the elements of the basis*, with integer coefficients. This basis is said to be "complete", meaning that *any real number can be represented with arbitrarily-high precision*, as a linear combination of different powers of ten (with integer coefficients), provided that enough powers of ten are included in the expansion.

The above ideas have a direct analogue in quantum physics. Rather than expressing real numbers as a linear superposition of elements of the set $\{\cdots,10^{2},10^{1},10^{0},10^{-1},10^{-2},\cdots\}$, we wish to express *arbitrary wavefunctions* $\Psi(\mathbf{r},t)$ as a *linear superposition of the set of all plane waves* (10)[10]. This superposition is effected via the Fourier integral:

$$\Psi(\mathbf{r},t) = \frac{1}{(2\pi)^{3/2}}\iiint \tilde{\Psi}(\mathbf{k})e^{i(\mathbf{k}\bullet\mathbf{r}-\omega t)}d\mathbf{k}. \quad (15)$$

Here, $\tilde{\Psi}$ denotes the Fourier transform of Ψ with respect to x, y, and z. The multiplicative factor, outside the front of the integral, is there because of the particular convention for Fourier transforms that is used in this course.

Before proceeding with more details regarding Fourier analysis in the context of quantum physics, it is imperative that we have a sound physical understanding of the meaning of (15)! This should be interpreted in very physical terms as a *linear combination/superposition of plane waves* $\exp[i(\mathbf{k}\bullet\mathbf{r}-\omega t)]$, of all possible different directions of propagation, and all possible different de-Broglie wavelengths. Stated differently: equation (15) represents an arbitrary wavefunction $\Psi(\mathbf{r},t)$ as *a linear superposition of plane waves, of all possible propagation vectors* **k**. In this superposition, a specified plane wave $\exp[i(\mathbf{k}\bullet\mathbf{r}-\omega t)]$ is weighted by the infinitesimally small weighting factor $(2\pi)^{-3/2}\tilde{\Psi}(\mathbf{k})d\mathbf{k}$, with all such re-weighted plane waves then being superposed (added together) using the triple integral. [Again, it may be useful to think once again of the analogy between this process, and the example of the real numbers given earlier in this section.]

The function $\tilde{\Psi}$ is known as the "momentum space wavefunction", being the Fourier transform of Ψ with respect to x, y, and z. Equation (15) may be viewed as a mathematical recipe for converting $\tilde{\Psi}$ into Ψ. The corresponding formula for "going the other way", namely a mathematical recipe for converting Ψ into $\tilde{\Psi}$, is given by:

$$\tilde{\Psi}(\mathbf{k}) = \frac{1}{(2\pi)^{3/2}}\iiint \Psi(\mathbf{r},t)e^{-i(\mathbf{k}\bullet\mathbf{r}-\omega t)}d\mathbf{r}. \quad (16)$$

Together, equations (15) and (16) form a "Fourier transform pair".

In view of the fact that each can be translated into the other, Ψ and $\tilde{\Psi}$ may be viewed as *different ways of representing the same physical entity.*

► *Exercise #9. Derive (16) from (15). In obtaining this result, you may like to make use of the integral representation of the three-dimensional Dirac delta, namely:*

$$\delta(\mathbf{r}) \equiv \delta(x)\delta(y)\delta(z) = \frac{1}{(2\pi)^{3}}\iiint e^{i\mathbf{k}\bullet\mathbf{r}}d\mathbf{k}. \quad (17)$$

Note that the Dirac delta has the property that its integral is unity:

$$\iiint \delta(\mathbf{r})d\mathbf{r} = \int\delta(x)dx\int\delta(y)dy\int\delta(z)dz = 1, \quad (18)$$

together with the "sifting property":

$$\iiint f(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}_0)d\mathbf{r} = f(\mathbf{r}_0) \quad (19)$$

for any function f *of position* $\mathbf{r}$*, where* $\mathbf{r}_0$ *is a given fixed position vector. Loosely speaking,* $\delta(\mathbf{r})$ *may be viewed as a "function" which is infinite at the origin, and zero elsewhere.*

Please spend an hour or so in studying any mathematics or physics text which deals with the basics of Fourier transforms, if the above mathematics is unfamiliar to you.

We close this section by introducing the notion of a "wave-packet". The plane wave, in equation (10), evidently fills all of space in the sense that its magnitude is equal to unity everywhere in space. However, the linear superposition of infinitely-many such plane waves, as given by integral in (15), may be localised in space in the sense that its magnitude is peaked over a given finite volume of space, with the magnitude being very small outside this volume. Such localised wavefunctions are known as "wavepackets". They may be viewed as "fuzzy blobs" of wavefunction which have a degree of

---

[10] The set of plane waves is a "basis" for expanding wavefunctions. This basis is "complete" in the sense that any wavefunction can be expanded in terms of this basis, using equation (15), to arbitrarily good accuracy.

localisation (see Figure 5), a state of affairs which may be contrasted with the infinitely-delocalised plane-wave wavefunction.

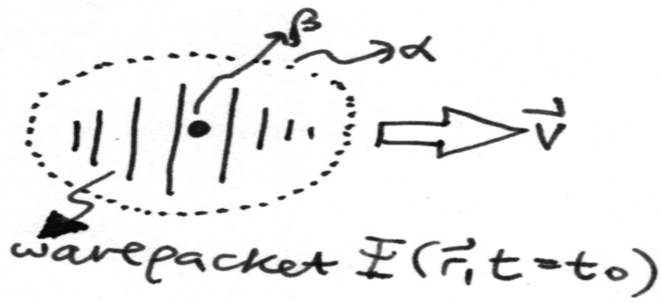


Figure 5: The wave-packet $\Psi(\mathbf{r}, t=t_0)$ is localised in the sense that its magnitude is non-negligible only inside the finite volume bounded by the dotted line α. The point β denotes the "centre of mass" of the wave-packet, about which more will be said later. The point β has velocity $\mathbf{v}$.

§6. Interpretation of the wave function

The discussions of the previous section have gone some way towards interpreting the meaning of the quantum-mechanical notion of a wavefunction. The interpretation of the wavefunction is further investigated in the present section.

§6.1 Born interpretation

Max Born is responsible for developing what is now called the "Born interpretation" of the wavefunction. On this view, the *modulus squared of the wavefunction is interpreted as being proportional to the probability of finding a particle at the said point.* Introducing the "position probability density" ("probability density" for short) $\rho(\mathbf{r},t)$ via:

(20) $\rho(\mathbf{r},t) = |\Psi(\mathbf{r},t)|^2$,

a more precise statement of the Born interpretation considers

(21) $\rho(\mathbf{r},t)d\mathbf{r} = |\Psi(\mathbf{r},t)|^2 \, d\mathbf{r}$

to be equal to the probability of detecting a particle in the volume of space (with volume d**r**) located at the position vector **r**, at time t.

Since the probability of finding a particle somewhere must be equal to unity, the integral of (21) over all space is unity, so that:

(22) $\iiint |\Psi(\mathbf{r},t)|^2 \, d\mathbf{r} = 1$.

The above condition expresses the demand that the wavefunction be "normalised". Note that if a wavefunction is not normalised, and the integral on the left of (22) is both finite and nonzero, the wavefunction can be normalised by making the replacement[11]:

(23) $\Psi(\mathbf{r},t) \to \dfrac{\Psi(\mathbf{r},t)}{\sqrt{\iiint |\Psi(\mathbf{r},t)|^2 \, d\mathbf{r}}}$.

Evidently, the Born interpretation of the wavefunction is of a statistical nature.

§6.2 Superposition principle

Forget about quantum mechanics for the moment, and examine Figure 6.

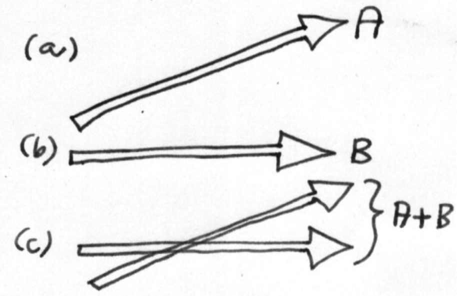


Figure 6: Superposition principle for light beams A and B

Suppose Figure 6(a) to represent a light beam A propagating through free space. Suppose Figure 6(b) to represent a different light beam B, propagating through the same space. Now suppose the two light beams are *superposed*, as shown in Figure 6(c). These light beams "go straight through one another" without interacting. This allows us to add the wavefield A to the wavefield B, to obtain another valid wavefield A+B. Mathematically speaking, this is a consequence of the fact that the underlying differential equations (in this case, the Maxwell equations) are *linear*. The fact that A+B is a solution to a given linear equation (or linear equations), whenever A and B are solutions, is known as the *superposition principle.*

Let us return to quantum mechanics. The underlying equation (the Schrödinger equation, see next chapter) governing wave functions is assumed to be linear, and therefore to obey the superposition principle. Thus if

---

[11] We have assumed that the factor, inside the square root in the denominator of (23), is not infinite. If this is the case, the wavefunction is said to be "square integrable" (because its squared modulus can be integrated over all space to yield a finite number). An important exception is the plane wave, which is not square integrable. It cannot be normalized using the procedures described above. Please see me if you would like to find out more about this point.

$\Psi_A(\mathbf{r},t)$ is an allowed wavefunction that obeys the Schrödinger equation, and $\Psi_B(\mathbf{r},t)$ is another allowed wavefunction that obeys the Schrödinger equation, then $\Psi_A(\mathbf{r},t)+\Psi_B(\mathbf{r},t)$ will also be an allowed wavefunction.

Denote the position probability density of Ψ_A by $\rho_A = |\Psi_A|^2$, with $\rho_B = |\Psi_B|^2$ and $\rho_{A+B} = |\Psi_A + \Psi_B|^2$ similarly defined. Now:

$$\begin{aligned} \rho_{A+B} &= |\Psi_A + \Psi_B|^2 \\ &= |\Psi_A|^2 + |\Psi_B|^2 + 2\,\mathrm{Re}(\Psi_A \Psi_B^*) \\ &= \rho_A + \rho_B + 2\,\mathrm{Re}(\Psi_A \Psi_B^*). \end{aligned} \tag{24}$$

The probability density of $\Psi_A + \Psi_B$ is not equal to the sum of the probability densities of Ψ_A and Ψ_B. Rather, there is an additional term—namely the third term on the last line of equation (24)—which is indicative of *interference* between the individual wavefunctions Ψ_A and Ψ_B.[12]

§6.3 Correspondence principle

As briefly mentioned earlier in this course, quantum mechanics must reduce to classical mechanics for physical scenarios which lie within the domain of validity of the older, less general theory. This demand is known as the *correspondence principle.*

We shall have much more to say about the correspondence principle, particularly in our later discussions on the Ehrenfest theorem. For the moment, we wish to give a brief indication of how the notion of wave-packets (see §5) can shed some light on the transition from quantum mechanics to classical mechanics.

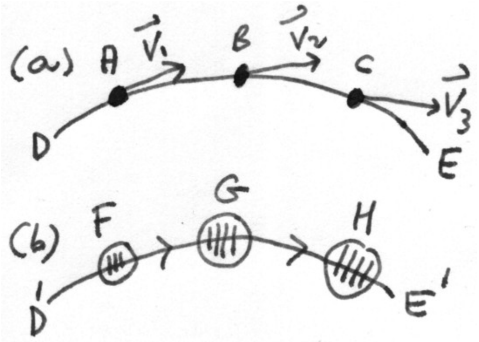


Figure 7: An electron moves along a curved trajectory on traversing a magnetic field

Consider Figure 7(a). This represents the classical-physics scenario where an electron (considered as a point particle, namely a particle of infinitely small size) moves along a curved trajectory DE, as it passes through a magnetic field. At three succeeding instants of time, the electron respectively occupies the points A, B, and C, each of these three points lying on the trajectory DE. At each of the points A, B, and C, the respective velocity vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are easily defined.

A quantum-mechanical version is given in Figure 7(b). Rather than having a point particle of zero spatial extent, one instead has a wave-packet of non-zero extent (cf. Figure 5). The trajectory D'E' may be defined as being traced out by the "centre of mass" of the wave-packet[13]. The labels F, G, and H, respectively denote the wave-packets of the (reasonably well localised) electron at three succeeding times. In the classical limit, the non-zero extent of the wave-packets is ignored. In this same limit, the centre-of-mass of the wave-packet should follow the trajectory predicted by classical mechanics.

§6.4 Bound states and unbound states

The plane wave is an example of a so-called "unbound state" or "unbound wavefunction", this terminology being used to indicate that the said wavefunction is not localised to a given finite volume of space.

Another example of an unbound wavefunction is shown in Figure 8. Here, we have sketched the distorted spherical surfaces of constant phase associated with an irregular source S of matter waves (say, the tip of an electron gun in an electron microscope). Again, the wavefunction is said to be "unbound" on account of the fact that the said wavefunction is not localised to a given finite volume.

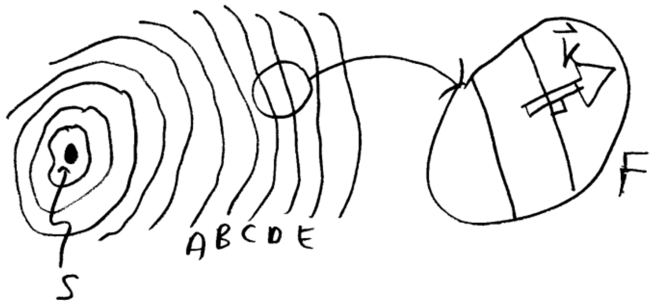


Figure 8: Distorted spherical wave-function emerging from a source S of matter waves[14].

---

[12] A classic example, of such interference, is given by the Young double-slit experiment using matter waves.

[13] Note that it will be a lecture or two before we render more precise the notion of the "centre of mass" of a wave-packet.

[14] In contrast to the case of a plane wave given in Figure 4, the wave-vector **k** is now a function of position. This observation is closely related to the notion of *probability current density*, to be introduced in the next chapter.

A bound wavefunction is localised to a given volume of space. One example of a bound wavefunction is the wavefunction of the electron in a hydrogen atom, which is in a state of sufficiently low energy that the electron is unable to "escape" from the Coulomb attraction of the nucleus (see chapter 7). Other examples include a particle confined to an infinite square well, and the harmonic oscillator (both of these latter examples will be treated in chapter 4).

► *Exercise #10. Using the superposition principle, study the interference fringes formed by superposing two plane waves, each of which have the same angular frequency, but which point in different directions. Show that the total probability density is time-independent.*

CHAPTER 3

THE SCHRÖDINGER EQUATION

The Schrödinger equation, which governs wavefunction dynamics, is introduced and studied in the present chapter.

In §7, we use the elementary plane wave (10) to introduce the so-called "Schrödinger correspondence rules". These rules assign a mathematical *operator* to both the energy and momentum of a material particle[15].

Using these correspondence rules, we construct the time-dependent form of the Schrödinger equation, in §8. A corresponding time-independent form of this equation, valid for quantum systems of well-defined energy E, is obtained in §9.

Certain boundary conditions, which restrict the class of physically admissible wavefunctions, are introduced in §10.

In §11, we use the Schrödinger equation to obtain an expression for local conservation of probability density, known as the continuity equation[16]. This study will allow us to obtain an expression for the probability current density **j**(**r**,t) associated with a given quantum-mechanical wavefunction.

Section 12 extends the operator concept, which was first introduced in §7. We introduce the notion of non-commuting operators, commutators and expectation values. We are then able, for example, to give a precise meaning to the notion of the "centre of mass" and the "momentum" of a wavepacket.

In §13, we return to the question of how classical mechanics emerges as a limiting case of quantum mechanics. In particular, we derive the Ehrenfest theorem, which obtains Newton's equations for a point particle in a variable scalar potential, as a limiting case of the Schrödinger equation, for the case of wavepackets that are sufficiently small for their spatial extent to be neglected.

§7. Schrödinger correspondence rules

Return consideration to the expression for the wavefunction of an elementary plane-wave, which was given in equation (10):

$$(10)\quad \Psi(\mathbf{r},t)=\exp[i(\mathbf{k}\bullet\mathbf{r}-\omega t)].$$

Recall that **k** is the wave-vector of the plane wave, with $|\mathbf{k}| = 2\pi/\lambda$, where λ is the de-Broglie wavelength of the plane wave (see equation (9) and Figure 4). Also, $\mathbf{r}$ = (x,y,z) denotes the position vector in three-dimensional space, t denotes time, and ω is the angular frequency of the wave (see equation (11)).

From equation (13), we have:

$$(13)\quad \mathbf{p}=\hbar\mathbf{k},$$

which states that the momentum **p** and wave-vector **k** are proportional to one another. We may therefore write:

$$(25)\quad \mathbf{k}=\mathbf{p}/\hbar.$$

Also, note that the relation $E=\hbar\omega$ (see footnote 9) implies that the energy E and angular frequency ω are related by:

$$(26)\quad \omega=E/\hbar.$$

Substitute (25) and (26) into (10), in order to obtain an expression for the plane wave in terms of its energy E and momentum **p**:

$$(27)\quad \Psi(\mathbf{r},t)=\exp\left[i\left(\frac{\mathbf{p}\bullet\mathbf{r}}{\hbar}-\frac{Et}{\hbar}\right)\right].$$

We leave it as an exercise to show that the above equation implies the following pair of expressions:

$$(28)\quad i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r},t)=E\Psi(\mathbf{r},t),\ \Psi=\text{plane wave},$$

$$(29)\quad -i\hbar\nabla\Psi(\mathbf{r},t)=\mathbf{p}\Psi(\mathbf{r},t),\ \Psi=\text{plane wave}.$$

---

[15] Note that more will be said about operators in §12.

[16] Note that continuity equations arise in a very broad variety of physical theories, including fluid mechanics, classical electrodynamics, classical optics, quantum field theory, general relativity, *etc*.

► *Exercise #11. Prove (28) and (29).*

In words, equation (28) states that, "When the operator $i\hbar\partial/\partial t$ acts on the plane-wave wavefunction $\Psi(\mathbf{r},t)$, the result is equal to the energy E of the plane wave multiplied by the wave-function $\Psi(\mathbf{r},t)$". This motivates the *first Schrödinger correspondence rule:*

$$(30)\quad E \to i\hbar\frac{\partial}{\partial t}.$$

Thus the energy E is replaced by the *operator* $i\hbar\partial/\partial t$.

In words, equation (29) states that, "When the operator $-i\hbar\nabla$ acts on the plane-wave wavefunction $\Psi(\mathbf{r},t)$, the result is equal to the momentum **p** of the plane wave multiplied by the wave-function $\Psi(\mathbf{r},t)$". This motivates the *second Schrödinger correspondence rule:*

$$(31)\quad \mathbf{p} \to -i\hbar\nabla.$$

Thus the momentum **p** is replaced by the *operator* $-i\hbar\nabla$.

We shall have a more extensive discussion, on the use of mathematical operators in quantum mechanics, in §12. For the moment, the single salient point is that the two Schrödinger correspondence rules, (30) and (31), allow us to "extract" the physical quantities E and **p** from a plane-wave wavefunction, according to equations (28) and (29)[17].

As we shall see in the next section, the Schrödinger correspondence rules will prove useful in obtaining the Schrödinger equation.

§8. Time-dependent Schrödinger equation

Here, we introduce the famous Schrödinger equation—as mentioned earlier, this equation governs the dynamics of the wavefunction.

We emphasize that the Schrödinger equation is *not* obtained via a logical process of deductive reasoning. Rather, what follows is in essence a *plausibility argument*, albeit one that was deeply inspired by some of the observations of the era of the old quantum theory, with the ultimate justification for the Schrödinger equation being its success in correctly predicting the results of a wide variety of experiments (up to a given accuracy).

Before using the Schrödinger correspondence rules to motivate the introduction of the Schrödinger equation, we list three key requirements that we have for this equation.

- *Requirement #1: The Schrödinger equation must be first order with respect to time.* Recall that the wavefunction is postulated to completely define the dynamical state of a given quantum system[18]. Since it completely defines the state of a system, knowledge of the wavefunction at a given time t_0, should give sufficient information for one to determine the wavefunction at any future time. This will be the case if the Schrödinger equation is first order in time[19], that is, if it contains only first-order derivatives with respect to time.
- *Requirement #2: The Schrödinger equation must be a linear equation.* We discussed the superposition principle in §6.2: if $\Psi_A(\mathbf{r},t)$ is a wavefunction that obeys the Schrödinger equation, and $\Psi_B(\mathbf{r},t)$ is another wavefunction that obeys the Schrödinger equation, then $\Psi_A(\mathbf{r},t)+\Psi_B(\mathbf{r},t)$ will also obey the Schrödinger equation. The superposition principle, which is obeyed by linear equations and disobeyed (in general) for non-linear equations, implies that the Schrödinger equation must be linear in Ψ.
- *Requirement #3*: *The Schrödinger equation must preserve the normalisation of the wavefunction.* Equation (22) states that the probability, of finding a particle somewhere, is equal to unity at all times[20]. The Schrödinger equation must be consistent with this demand. In particular, if the left side of (22) is equal to unity at a given time, it must equal unity at all times.

With a view to motivating the Schrödinger equation, let us write down the relation between the energy E and momentum **p**, of a classical point particle in a force-field derivable from a scalar potential $V(\mathbf{r},t)$:

$$(32)\quad E = \frac{\mathbf{p}\bullet\mathbf{p}}{2m} + V(\mathbf{r},t).$$

Making use of the Schrödinger correspondence rules (30) and (31), this becomes the operator equation:

---

[17] Indeed, both (28) and (29) have the form of eigenvalue equations—that is, equations in the form "An operator acts on a (wave- or eigen-)function to give a constant ('eigen-value') multiplied by the (wave- or eigen-) function". We shall have a great deal more to say about eigenvalues and eigenfunctions, later in this course.

[18] See opening paragraph of chapter 2.

[19] Please see me if you would like more detail on this point.

[20] We implicitly assume the wavefunction to be a single-particle wavefunction, throughout the entire course. Please bear in mind, however, that one may also write down wavefunctions for multi-particle systems. This is the realm of many-body theory, which is beyond the scope of these lectures.

$$i\hbar\frac{\partial}{\partial t} = \frac{(-i\hbar\nabla)\bullet(-i\hbar\nabla)}{2m} + V(\mathbf{r},t)$$

$$(34)\quad i\hbar\frac{\partial}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r},t)\,.$$

If we now post-multiply[21] both sides of this equation by the wavefunction, we obtain the famous *time-dependent Schrödinger equation for a particle in a potential* V(**r**,t)*:*

$$(35)\quad \boxed{i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r},t) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r},t)\right]\Psi(\mathbf{r},t)}.$$

[Special case: In the absence of a potential, i.e. when V = 0, we obtain the *time-dependent Schrödinger equation for a free particle:*

$$(36)\quad i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r},t) = -\frac{\hbar^2}{2m}\nabla^2\Psi(\mathbf{r},t)\,.]$$

Again, we emphasize that the above plausibility argument is ultimately justified by the success of the Schrödinger equation in predicting the results of experiment.

Let us return to the three requirements for the Schrödinger equation, which were enumerated on the previous page.

(a) Requirement #1: The Schrödinger equation must be first order with respect to time. From equation (35), this requirement is trivially satisfied.

(b) Requirement #2: The Schrödinger equation must be a linear equation. From equation (35), this requirement is trivially satisfied. Note that the superposition principle follows as a direct consequence of this linearity, as we now show. Suppose that $\Psi_A(\mathbf{r},t)$ is a wavefunction that obeys the Schrödinger equation, so that:

$$(37)\quad \begin{aligned} &i\hbar\frac{\partial}{\partial t}\Psi_A(\mathbf{r},t)\\ &\quad = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r},t)\right]\Psi_A(\mathbf{r},t). \end{aligned}$$

Suppose that $\Psi_B(\mathbf{r},t)$ is another wavefunction that obeys the Schrödinger equation, so that:

$$(38)\quad \begin{aligned} &i\hbar\frac{\partial}{\partial t}\Psi_B(\mathbf{r},t)\\ &\quad = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r},t)\right]\Psi_B(\mathbf{r},t). \end{aligned}$$

Add together equations (37) and (38):

$$(39)\quad \begin{aligned} &i\hbar\frac{\partial}{\partial t}\Psi_A(\mathbf{r},t) + i\hbar\frac{\partial}{\partial t}\Psi_B(\mathbf{r},t)\\ &\quad = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r},t)\right]\Psi_A(\mathbf{r},t)\\ &\qquad + \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r},t)\right]\Psi_B(\mathbf{r},t). \end{aligned}$$

This may be written as:

$$(40)\quad \begin{aligned} &i\hbar\frac{\partial}{\partial t}\{\Psi_A(\mathbf{r},t) + \Psi_B(\mathbf{r},t)\}\\ &\quad = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r},t)\right]\{\Psi_A(\mathbf{r},t) + \Psi_B(\mathbf{r},t)\}. \end{aligned}$$

Thus, if $\Psi_A(\mathbf{r},t)$ is a solution to the Schrödinger equation (35), and $\Psi_B(\mathbf{r},t)$ is another solution to the Schrödinger equation, then $\Psi_A(\mathbf{r},t) + \Psi_B(\mathbf{r},t)$ is also a solution.

(c) Requirement #3: *The Schrödinger equation must preserve the normalisation of the wavefunction.* I have left it to you, in the following exercise, to show that the Schrödinger equation preserves the normalisation of the wavefunction—that is, if the wavefunction is normalised at a given time (i.e. if the left side of (22) is equal to unity at a given time), then the Schrödinger equation implies it to be normalised at all times.

► *Exercise #12. Prove the above statement.*

We close this section by noting that the total energy of a quantum system is often termed its “Hamiltonian”[22]. Since the left side of (34) was seen to be equal to the energy operator, we may replace this side by the Hamiltonian operator $H(\mathbf{r},t)$, to arrive at the following expression for the Hamiltonian operator:

---

[21] That is, multiply the wavefunction on the right side of each term in the equation. Thus, for example, if we were to *post-multiply* the equation A = B + C by D, we would obtain the equation AD = BD + CD. If, on the other hand, we were to *pre-multiply* the equation A = B + C by D, we would obtain the equation DA = DB + DC.

[22] There is a deeper reason for this terminology, which arises from the Hamiltonian formulation of both quantum and classical mechanics. We shall not be exploring this point further in the present course, but I though it would be worth mentioning!

For more on the use of Hamiltonians in classical mechanics, see e.g. the book on classical mechanics by Herbert Goldstein. For more on the use of Hamiltonians in quantum mechanics, see e.g. the book on quantum mechanics by Albert Messiah.

$$(41)\quad H(\mathbf{r},t) = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r},t).$$

► *Exercise #13. Show that the Fourier representation (15), for an arbitrary time-dependent wavefunction, is indeed a solution to the time-dependent Schrödinger equation for a free particle (equation (36)).*

§9. Time-independent Schrödinger equation

Consider, once again, the Fourier integral in equation (15):

$$(15)\quad \Psi(\mathbf{r},t) = \frac{1}{(2\pi)^{3/2}}\iiint \widetilde{\Psi}(\mathbf{k})e^{i(\mathbf{k}\bullet\mathbf{r}-\omega t)}d\mathbf{k}.$$

Recall that this expresses an arbitrary wave-function $\Psi(\mathbf{r},t)$ as a superposition of plane waves, of all possible energies and all possible propagation directions.

Suppose, now, that rather than the general form (15), we were to consider a restricted form of (15), in which *only plane waves of a given energy* E are superposed. Since (26) states that $\omega = E/\hbar$, we are restricting the superposition (15) to contain only plane waves of a given angular frequency ω. Under this restriction, ω—which, in general, depends on the magnitude of **k** and therefore should be contained within the integral, which integrates over **k**—may be taken outside the integral in (15), leaving us with[23]:

$$(42)\quad \Psi_E(\mathbf{r},t) = \frac{e^{-i\omega t}}{(2\pi)^{3/2}}\iiint \widetilde{\Psi}(\mathbf{k})e^{i\mathbf{k}\bullet\mathbf{r}}d\mathbf{k}.$$

Now, equation (26) allows us to write $\exp(-i\omega t)$ as $\exp(-iEt/\hbar)$. Further, let us introduce the *spatial wavefunction* $\psi(\mathbf{r})$ via:

$$(43)\quad \psi(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}}\iiint \widetilde{\Psi}(\mathbf{k})e^{i\mathbf{k}\bullet\mathbf{r}}d\mathbf{k}.$$

Thus (42) becomes:

$$(44)\quad \boxed{\Psi_E(\mathbf{r},t) = \psi(\mathbf{r})\exp(-iEt/\hbar)}.$$

We see that a wavefunction, built up of plane waves that all have the same energy E, may be expressed as a product of the spatial wavefunction $\psi(\mathbf{r})$ (which depends on position but not time) and the harmonic factor $\exp(-iEt/\hbar)$ (which depends on time but not position). Solutions of the form (44), which evidently correspond to a well-defined energy E, are known as *stationary-state* solutions.

What equation is obeyed by the spatial wavefunction $\psi(\mathbf{r})$? To find out, substitute (44) into the time-dependent Schrödinger equation (35). This gives:

$$(45)\quad \begin{aligned} &i\hbar\frac{\partial}{\partial t}\{\psi(\mathbf{r})\exp(-iEt/\hbar)\} \\ &= \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\{\psi(\mathbf{r})\exp(-iEt/\hbar)\}. \end{aligned}$$

Note that we have replaced V(**r**,t) with V(**r**) in the above expression, which amounts to the *assumption that the potential does not depend on time.*

Perform the differentiation with respect to time on the left side of this equation, to see that:

$$(46)\quad \begin{aligned} &i\hbar(-iE/\hbar)\{\psi(\mathbf{r})\exp(-iEt/\hbar)\} \\ &= \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\{\psi(\mathbf{r})\exp(-iEt/\hbar)\}. \end{aligned}$$

Cancelling common factors, we arrive at the *time-independent Schrödinger equation for a particle in a time-independent potential* $V(\mathbf{r})$:

$$(47)\quad \boxed{\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\psi(\mathbf{r}) = E\psi(\mathbf{r})}.$$

The above time-independent equation might be solved to yield the spatial wavefunction $\psi(\mathbf{r})$. The full (time-dependent) wavefunction is then obtained using (44).

We shall spend much of our time, in subsequent lectures, solving the time-independent Schrödinger equation (47) for the spatial wavefunction. Indeed, this is probably the most-used result of the course.

We close this section by mentioning that (47) has the form of an *eigenvalue equation*[24]—that is, equation (47) has the form, "An operator acts on the eigenfunction to give the eigenfunction back again, multiplied by a *number* known as the eigenvalue". In the previous sentence in quotation marks, the eigenfunction is the spatial wavefunction $\psi(\mathbf{r})$, the eigenvalue is the energy E, and the operator is the operator in square brackets (of equation (47)).

§10. Boundary conditions for the wavefunction

The wave function $\Psi(\mathbf{r},t)$ is assumed to satisfy the following three conditions:

---

[23] Note that we have put an "E" subscript on the left side of (42), to indicate that this wavefunction is made up by summing all possible plane waves that have a fixed energy E.

[24] See footnote to §7.

- *Condition I:* The wavefunction $\Psi(\mathbf{r},t)$ must be a single-valued, continuous function of position and time;
- *Condition II:* (b) Bound wavefunctions should be square integrable[25];
- *Condition III:* At any given time, the spatial derivatives of $\Psi(\mathbf{r},t)$ (i.e. $\partial\Psi/\partial x,\ \partial\Psi/\partial y,\ \partial\Psi/\partial z$) must be continuous everywhere, except where there is an infinite discontinuity in the potential.

From time to time throughout this course, we shall make use of the above boundary conditions:

(a) For example, we shall see that Condition I's demand for a single-valued wavefunction—namely, the demand that the complex wavefunction have a single complex value at each point in space and at each instant of time—is very closely related to the quantisation of angular momentum in the hydrogen atom (see chapter 7).

(b) As another example, condition III will be of use in studying the one-dimensional barrier penetration problem (see chapter 4).

§11. Probability current density

Here, we develop the notion of *probability current density*, often abbreviated *probability current* or even just *current*, for a quantum-mechanical wavefunction obeying the time-dependent Schrödinger equation (35). Loosely speaking, the probability current density is analogous to the notion of "current at a point" which is employed in classical fluid mechanics.

With a view to obtaining an expression for the probability current density, pre-multiply the time-dependent Schrödinger equation (35) by the complex conjugate Ψ^* of the wavefunction, to give:

$$(48)\quad i\hbar\Psi^*\frac{\partial}{\partial t}\Psi=\Psi^*\left[-\frac{\hbar^2}{2m}\nabla^2+V\right]\Psi .$$

Note that we are suppressing explicit functional dependence of the wavefunction and potential on position and time, to make our formulae less cluttered.

Put equation (48) to one side, for the moment. Take the complex conjugate of the time-dependent Schrödinger equation:

$$(49)\quad -i\hbar\frac{\partial}{\partial t}\Psi^*=\left[-\frac{\hbar^2}{2m}\nabla^2+V^*\right]\Psi^* .$$

Use the fact that the potential is real, to replace V^* by V in the above equation, and then pre-multiply by the wavefunction Ψ, to see that:

$$(50)\quad -i\hbar\Psi\frac{\partial}{\partial t}\Psi^*=\Psi\left[-\frac{\hbar^2}{2m}\nabla^2+V\right]\Psi^* .$$

Subtract equation (50) from equation (48), and note that the potential drops out, leaving:

(51)

$$i\hbar\left(\Psi^*\frac{\partial}{\partial t}\Psi+\Psi\frac{\partial}{\partial t}\Psi^*\right)=\frac{\hbar^2}{2m}\left(\Psi\nabla^2\Psi^*-\Psi^*\nabla^2\Psi\right).$$

We leave it as an exercise to show that the above equation is equivalent to:

$$(52)\quad \frac{\partial}{\partial t}|\Psi|^2+\nabla\bullet\left[\frac{i\hbar}{2m}\left(\Psi\nabla\Psi^*-\Psi^*\nabla\Psi\right)\right]=0 .$$

► *Exercise #14. Show that (51) and (52) are equivalent. You may find it helpful to invoke the vector identity:*

$$(53)\quad \nabla\bullet(A\nabla B)=A\nabla^2B+\nabla A\bullet\nabla B,$$

where A and B are functions of position in three-dimensional space.

Making use of the fact that $|\Psi|^2$ is equal to the probability density ρ, and introducing the desired current density vector $\mathbf{j}$ via:

$$(54)\ \mathbf{j}(\mathbf{r},t)=\frac{i\hbar}{2m}\left[\Psi(\mathbf{r},t)\nabla\Psi^*(\mathbf{r},t)-\Psi^*(\mathbf{r},t)\nabla\Psi(\mathbf{r},t)\right]$$

we see that (52) assumes the form of a so-called *continuity equation*:

$$(55)\quad \frac{\partial}{\partial t}\rho(\mathbf{r},t)+\nabla\bullet\mathbf{j}(\mathbf{r},t)=0 .$$

This equation expresses the local conservation of probability density, as we now argue.

Consider a small volume of space Ω, the smooth closed surface of which is denoted by $\partial\Omega$ (see Figure 9). Integrating over this three-dimensional volume, (55) becomes:

$$(56)\quad \iiint_\Omega\frac{\partial}{\partial t}\rho(\mathbf{r},t)dxdydz+\iiint_\Omega\nabla\bullet\mathbf{j}(\mathbf{r},t)dxdydz=0.$$

---

[25] That is, $\iiint|\Psi(\mathbf{r},t)|^2\,d\mathbf{r}$ should be finite. Indeed, for normalized wavefunctions, the just-mentioned integral will be equal to unity. Cf. § 6.1.

In the first triple integral, take the partial derivative with respect to time outside the front of the integral. In the second integral, make use of the Gauss divergence theorem[26], thereby converting this integral into an integral over the surface $\partial\Omega$. We thereby arrive at:

$$(57)\quad \iint_{\partial\Omega} \mathbf{j}(\mathbf{r},t)\bullet\hat{n}\, d\sigma = -\frac{\partial}{\partial t}\iiint_{\Omega} \rho(\mathbf{r},t)dxdydz\,.$$

Here, $d\sigma$ denotes an element of area on the surface $\partial\Omega$, and $\hat{n}$ denotes an outward-pointing unit vector that is normal to the same surface (see Figure 9).

Figure 9: Construction for use in studying the continuity equation. A volume Ω is bounded by the surface $\partial\Omega$. At a given point on the surface, $\hat{n}$ denotes an outward-pointing unit vector that is normal to the surface, while $d\sigma$ denotes an infinitesimal element of area.

Equation (57) gives the promised expression for the local conservation of probability density—in words, this equation states that the flux of the current density **j**, through the surface $\partial\Omega$, is equal to the negative of the time rate of change of the probability of finding a particle within the said volume. If there is a net positive (i.e., outward) flux of particles, the left side will be positive, and therefore the probability, of finding a particle within the volume, will decrease with time. If there is a net negative (i.e., inward) flux of particles, the left side will be negative, and therefore the probability, of finding a particle within the volume, will increase with time.

► *Exercise #15. Without loss of generality, any complex wavefunction can be written in the form:*

$$(58)\quad \Psi(\mathbf{r},t) = \sqrt{\rho(\mathbf{r},t)}\,\exp[iS(\mathbf{r},t)]\,.$$

Here, the real function $\rho(\mathbf{r},t)$ *is the now-familiar probability density, and the real function* $S(\mathbf{r},t)$ *is the phase of the wavefunction. Substitute (58) into (54) to obtain an expression for the probability current density, in terms of both* $\rho(\mathbf{r},t)$ *and* $S(\mathbf{r},t)$*. Use this result to show that the local current vector is perpendicular to the surfaces of constant phase (namely the "wavefronts" of the wavefunction—see paragraph of text just before Exercise #7).*

► *Exercise #16. Show that, for monoenergetic wavefunctions of the form (44), the current density vector is independent of time. (Note that this result is somewhat analogous to the notion of "steady state flow" in classical fluid mechanics.)*

► *Exercise #17. Show that the current density vanishes everywhere, for any wavefunction that is a real function of position and time. (Note that this fact is sometimes used to support the argument that the wavefunction should be a complex function of position and time, rather than a real function of position and time.) [OPTIONAL]*

§12. Operators, expectation values, commutators and complementarity

We have already seen mention of the notion of *operators* in quantum mechanics, one example being the Schrödinger correspondence rules encountered in §7, with another being the Hamiltonian operator that was introduced at the end of §8. Here, we deepen our understanding of quantum-mechanical operators, in a discussion that will include the important concepts of *Hermitian operators*, *expectation values*, *commutators*, *uncertainty relations* and *complementary variables*.

§12.1 Operators and expectation values

In Figure 5, we introduced—in a completely qualitative way—the notion of the "centre of mass" of a wavepacket. In the present section, we show how to calculate the centre-of-mass of a wavepacket. We then generalize this finding, by developing the notion of the *expectation value* of a given operator.

Now, "centre of mass" may be thought of as the *expectation value* of the position of the wavepacket. For a normalised wavefunction—that is, a wavefunction that obeys equation (22)—we may write the expectation value of position[27] as follows:

$$(59)\quad \langle\mathbf{r}\rangle = \iiint \mathbf{r}\,\rho(\mathbf{r},t)d\mathbf{r} = \iiint \Psi^*(\mathbf{r},t)\,\mathbf{r}\,\Psi(\mathbf{r},t)d\mathbf{r}.$$

---

[26] Please spend about half an hour briefly revising the Gauss divergence theorem, if this concept is unfamiliar to you.

[27] Note that the expectation value of a given quantity is denoted by putting angular brackets around that quantity. Thus, for example, the expectation value of **r** is written as <**r**>.

Note that the above vector equation is actually three scalar equations, on account of the fact that **r** is a position vector in three-dimensional space. Explicitly, the x, y, and z components of the above equations are as follows:

$$(60)\quad \begin{cases} \langle x\rangle = \iiint \Psi^*(\mathbf{r},t)\, x\, \Psi(\mathbf{r},t)d\mathbf{r}, \\ \langle y\rangle = \iiint \Psi^*(\mathbf{r},t)\, y\, \Psi(\mathbf{r},t)d\mathbf{r}, \\ \langle z\rangle = \iiint \Psi^*(\mathbf{r},t)\, z\, \Psi(\mathbf{r},t)d\mathbf{r}. \end{cases}$$

As a special case of (59), suppose that at some instant of time t_0, a given wavepacket is completely localized to the point $\mathbf{r}_0$, allowing us to write its probability density at this time as:

$$(61)\quad \rho(r,t_0) = \delta(\mathbf{r}-\mathbf{r}_0),$$

where δ is the three-dimensional Dirac delta (see equations (17) and (18)). The expectation value of position is then:

$$(62)\quad \langle \mathbf{r}\rangle = \iiint \mathbf{r}\,\rho(\mathbf{r},t)d\mathbf{r} = \iiint \mathbf{r}\,\delta(\mathbf{r}-\mathbf{r}_0)d\mathbf{r} = \mathbf{r}_0,$$

as expected[28].

The lower part of equation (59), for the expectation value of **r**, can be generalized. To this end, replace **r** with an arbitrary function or operator A, to give the following expression for the expectation value of A:

$$(63)\quad \langle A(\mathbf{r},t)\rangle = \iiint \Psi^*(\mathbf{r},t)\, A(\mathbf{r},t)\, \Psi(\mathbf{r},t)d\mathbf{r}.$$

As a special case of the above formula, let us consider A to be the momentum operator **p**, so that:

$$(64)\quad \begin{aligned} \langle \mathbf{p}\rangle &= \iiint \Psi^*(\mathbf{r},t)\, \mathbf{p}\, \Psi(\mathbf{r},t)d\mathbf{r}, \\ &= \iiint \Psi^*(\mathbf{r},t)\, (-i\hbar\nabla)\, \Psi(\mathbf{r},t)d\mathbf{r}. \end{aligned}$$

Note that we have made use of the Schrödinger correspondence rule (31), in writing down the second line of the above equation.

Since ∇ is a vector (operator) in three-dimensional space, the above vector equation corresponds to the following three scalar equations:

$$(65)\quad \begin{cases} \langle p_x\rangle = \iiint \Psi^*(\mathbf{r},t)\left(-i\hbar\dfrac{\partial}{\partial x}\right)\Psi(\mathbf{r},t)d\mathbf{r}, \\ \langle p_y\rangle = \iiint \Psi^*(\mathbf{r},t)\left(-i\hbar\dfrac{\partial}{\partial y}\right)\Psi(\mathbf{r},t)d\mathbf{r}, \\ \langle p_z\rangle = \iiint \Psi^*(\mathbf{r},t)\left(-i\hbar\dfrac{\partial}{\partial z}\right)\Psi(\mathbf{r},t)d\mathbf{r}. \end{cases}$$

Here, p_x, p_y, p_z are respectively equal to the x, y and z components of the momentum **p.**

The expectation values of operators that are associated with physical quantities, which are known as *observables*, must be real. Thus if A is an observable, its expectation value

$$(66)\quad \langle A\rangle = \iiint \Psi^* A \Psi d\mathbf{r}$$

must be real, so that:

$$(67)\quad \langle A\rangle = \langle A\rangle^*.$$

Substituting (66) into (67), we see that the demand, for $< A >$ to be real, reduces to the requirement that:

$$(68)\quad \begin{aligned} \iiint \Psi^* A\Psi d\mathbf{r} &= \left[\iiint \Psi^* A\Psi d\mathbf{r}\right]^* \\ &= \iiint \left[\Psi^* A\Psi\right]^* d\mathbf{r} \\ &= \iiint \left[\Psi^*(A\Psi)\right]^* d\mathbf{r} \\ &= \iiint \left[\Psi^{**}(A\Psi)^*\right] d\mathbf{r} \\ &= \iiint \left[\Psi(A\Psi)^*\right] d\mathbf{r} \\ &= \iiint (A\Psi)^*\Psi d\mathbf{r}. \end{aligned}$$

Operators, that obey the above restriction, are known as *Hermitian* operators.

► *Exercise #18. Show that both the momentum operator and the energy operator, as given by the Schrödinger correspondence rules, are Hermitian. [Hint: Integration by parts.]*

§12.2 Commutators and complementarity

The order, in which two real numbers are multiplied, is immaterial in the sense that the result is the same for either order: for example, $2\times 3 = 3\times 2 = 6$.

The order, in which mathematical quantities are multiplied, is not always immaterial: for example, if A and B are square matrices of the same size, then it is not in general true that AB=BA. Stated differently, it is not in general true that AB-BA is equal to zero, when A and B are matrices.

In this context, it is useful to introduce the "commutator of A and B", denoted by "[A,B]", via:

$$(69)\quad [A,B] = AB - BA.$$

If A and B are numbers, their commutator vanishes ("multiplication of two real numbers is commutative").

---

[28] Note that we have made use of the sifting property of the Dirac delta, as given in equation (19), in writing down equation (62).

If A and B are square matrices of the same dimension, their commutator will not in general vanish ("matrix multiplication is not commutative").

If A and B are *operators*—and this is the case that we are most interested in for the purposes of this course—then their commutator [A,B] will in general be non-vanishing.

► *Exercise #19. Derive the following identities of commutator algebra, as a direct consequence of the definition in (69). Here, A, B,C denote arbitrary linear operators*[29].

(70a) $[A,B] = -[B,A]$,

(70b) $[A,B+C] = [A,B]+[A,C]$,

(70c) $[A,BC] = [A,B]C + B[A,C]$,

(70d) $[A,[B,C]]+[B,[C,A]]+[C,[A,B]] = 0$.

[OPTIONAL]

► *Exercise #20. Derive the following commutators.*

(71a) $[x,p_x] = [y,p_y] = [z,p_z] = i\hbar$,

(71b) $[E,t] = i\hbar$.

Note that the energy E, *and the momentum* **p** = (p$_x$,p$_y$,p$_z$)*, are to be interpreted as operators, as given by the Schrödinger correspondence rules. [Hints: (a) Always post-multiply by an arbitrary wavefunction when proving these identities. Thus, for example, when proving (71b), consider the function* $[E,t]\Psi$*, and then show that* $[E,t]\Psi = i\hbar\Psi$*, which is equivalent to (71b). (b) Use the product rule!]*

We shall now show that the non-vanishing, of quantum-mechanical operator commutators, is closely related to the existence of an associated Heisenberg uncertainty principle.

To this end, we may introduce the *uncertainty* ΔA, in a measurement of the expectation value of A, as the root-mean-square deviation in that quantity:

$$(72)\quad \begin{aligned}\Delta A &= \sqrt{\left\langle \left(A - \langle A\rangle\right)^2 \right\rangle}\\ &= \sqrt{\left\langle \left(A - \langle A\rangle\right)\left(A - \langle A\rangle\right) \right\rangle}\\ &= \sqrt{\left\langle A^2 - A\langle A\rangle - A\langle A\rangle + \langle A\rangle^2 \right\rangle}\\ &= \sqrt{\left\langle A^2 - 2A\langle A\rangle + \langle A\rangle^2 \right\rangle}\\ &= \sqrt{\left\langle A^2\right\rangle - 2\langle A\rangle\langle A\rangle + \langle A\rangle^2}\\ &= \sqrt{\left\langle A^2\right\rangle - \langle A\rangle^2}.\end{aligned}$$

Equipped with this definition, we state that the non-vanishing of a commutator [A,B] implies the following uncertainty relation between the uncertainties ΔA and ΔB that are respectively associated with A and B:

(73) $\Delta A\,\Delta B \ge \frac{1}{2}\,|\langle[A,B]\rangle|$

► *Exercise #21 (Difficult!). Prove (73). [OPTIONAL]*

As an example, equations (71a) and (71b) respectively imply the position-momentum uncertainty principle:

$$(74a)\quad \begin{cases}\Delta x\,\Delta p_x \ge \frac{1}{2}\hbar,\\ \Delta y\,\Delta p_y \ge \frac{1}{2}\hbar,\\ \Delta z\,\Delta p_z \ge \frac{1}{2}\hbar,\end{cases}$$

and the energy-time uncertainty principle:

(74b) $\Delta E\,\Delta t \ge \frac{1}{2}\hbar$.

Since you have met these so-called Heisenberg uncertainty relations in a previous course, our discussion of these will be limited to three salient points:

(a) The just-mentioned connection between non-vanishing operator commutators and their associated uncertainty relation;

(b) The notion of complementary pairs of variables, such as energy and time, for which the associated uncertainty relation implies that it is not possible to measure both quantities with arbitrary accuracy;

(c) The notion of *minimum-uncertainty wavepackets*, such as the Gaussian wavepacket studied in §19, for which the inequalities in (74) are replaced by equalities.

---

[29] By definition, A is a "linear operator" if $A(a\Psi_1 + b\Psi 2) = a(A\Psi_1) + b(A\Psi_2)$, for any two wavefunctions Ψ_1 and Ψ_2, and any complex numbers a and b.

§13. Ehrenfest's theorem

A key foundational concept, in the classical mechanics of point particles, is the notion of a *trajectory*. As we now argue, in the quantum domain the notion of a trajectory may be meaningless.

Consider, for example, a hydrogen atom in its ground state. The size of the ground-state hydrogen atom is on the order of 1 angstrom. If the notion of the trajectory of this electron were to have any meaning, this trajectory should be experimentally observable[30]. For example, one might scatter a photon from the electron at multiple points along the electron trajectory, thereby "seeing" the position of the electron at a number of subsequent times. The photon that is used should have a wavelength that is somewhat smaller than the size of the atom. Yet even if the photon has the maximum allowed wavelength of approximately 1 angstrom (i.e., roughly the size of the hydrogen atom), which corresponds to a minimum photon energy in the high-energy ("hard") X-ray regime, the photon will have sufficient energy to strip the electron from the hydrogen atom. Thus the act of observing even a single point, along the putative electron trajectory, serves to *irreversibly perturb the electron*, rendering all subsequent observations of the electron to be meaningless, insofar as ascertaining the original "electron trajectory" is concerned. Thus, in the context of the present example, the notion of the electron's "trajectory" may be dismissed as meaningless.

Rather than considering the electron to be always "somewhere" at a given time, which is the essence of the now-discarded notion of a trajectory, we instead have the weaker notion of a wavefunction whose probability density is in general smeared throughout space, in accord with the Born interpretation of the wavefunction (cf. Figs. 4, 5 and 8).

Notwithstanding the above considerations, we do know that quantum mechanics should reduce to classical mechanics, in regimes where the latter is applicable. Thus, for example, the classical notion of a trajectory, while not being meaningful (in general) in the quantum world, may nevertheless be considered to emerge from an appropriate *limiting case of a quantum-mechanical calculation*.

The *Ehrenfest theorem* gives a detailed examination of the transition from the quantum to the classical domains, by showing that two of Newton's equations, of classical mechanics, emerge as a limit case of the time-dependent Schrödinger equation (35).

To this end, consider the time-rate-of-change of the expectation value of the x-component of the momentum operator, as given by the first member of equations (65):

$$(75)\quad \begin{aligned}\frac{d}{dt}\langle p_x\rangle &= \frac{d}{dt}\iiint \Psi^*\left(-i\hbar\frac{\partial}{\partial x}\right)\Psi d\mathbf{r}\\ &= -i\hbar\iiint \frac{d}{dt}\left(\Psi^*\frac{\partial\Psi}{\partial x}\right)d\mathbf{r}.\end{aligned}$$

Making use of the product rule, we have:

$$(76)\quad \frac{d}{dt}\langle p_x\rangle = -i\hbar\iiint\left(\frac{\partial\Psi^*}{\partial t}\frac{\partial\Psi}{\partial x}+\Psi^*\frac{\partial}{\partial x}\frac{\partial\Psi}{\partial t}\right)d\mathbf{r}.$$

► *Exercise #22. Use the Schrödinger equation (35), to obtain an expression for* $\partial\Psi/\partial t$*; the complex conjugate of this result will give an expression for* $\partial\Psi^*/\partial t$*. Substitute your expressions, for* $\partial\Psi/\partial t$ *and* $\partial\Psi^*/\partial t$*, into (76). By making using of an argument based on the Gauss divergence theorem, or otherwise, hence show that (76) leads to:*

$$(77)\quad \frac{d}{dt}\langle p_x\rangle = -\left\langle\frac{\partial V}{\partial x}\right\rangle.$$

In working this exercise, please assume that: (i) the potential is real; (ii) the wavefunction vanishes at infinity. [Hint: Study our previous discussions on the probability current.]

Equation (77) is the x-component of the following relation (first Ehrenfest equation):

$$(78)\quad \frac{d}{dt}\langle\mathbf{p}\rangle = -\langle\nabla V\rangle.$$

Since, in classical mechanics, a conservative force **F** is equal to $-\nabla V$, (78) becomes:

$$(79)\quad \frac{d}{dt}\langle\mathbf{p}\rangle = \langle\mathbf{F}\rangle.$$

This result—that the expectation value of the force is equal to the time-rate-of-change of the expectation value of the momentum—evidently makes contact with the classical mechanics of a point particle in a conservative force field describable by the scalar potential V, for which the force is equal to the time-rate-of-change of momentum. If the spatial extent of the wavepacket can be ignored, the classical result holds.

---

[30] This is a key point in the argument, which assumes that in order for a physical construct to be meaningful, it must be able to be *measured in an experiment*. If this assumption is not made, then the argument, given in the main text, falls apart.

We close this section by noting that Ehrenfest's second equation is:

$$(80)\quad m\frac{d}{dt}\langle \mathbf{r}\rangle = \langle \mathbf{p}\rangle .$$

This, too, directly relates to the classical limit that "momentum equals mass times velocity". In particular, the classical result is exactly recovered if the spatial extent of the wavepacket may be ignored.

► *Exercise #23. Derive equation (80).*

► *Exercise #24. Show that the operator* $x^m p_x^n$ *is not Hermitian, but that the combination* $(x^m p_x^n + p_x^n x^m)/2$ *is Hermitian, where* m *and* n *are positive integers.*

► *Exercise #25. Prove that if a constant* V_0 *is added to the time-independent potential energy* $V(\mathbf{r})$*, (i) the energy eigenvalues of the time-independent Schrödinger equation are shifted from* E *to* E+V$_0$*; and (ii) the corresponding spatial wavefunctions* $\psi(\mathbf{r})$ *remain unchanged.*

► *Exercise #26. Utilizing the fact that any wavefunction can be considered as a superposition of plane waves, show that in the absence of an external force field, the wave-function* $\Psi(\mathbf{r}_2, t_2)$*, at the point* $\mathbf{r}_2$ *at the instant* t_2*, can be deduced from the values* $\Psi(\mathbf{r}_1, t_1)$ *taken by the wave at the instant* t_1*, by:*

$$(81)\quad \begin{aligned}&\Psi(\mathbf{r}_2, t_2)\\ &= \iiint K(\mathbf{r}_2 - \mathbf{r}_1; t_2 - t_1)\Psi(\mathbf{r}_1, t_1)d\mathbf{r}_1,\end{aligned}$$

where the "propagator" K *is defined by:*

$$(82)\quad K(\mathbf{r}, t) = \frac{1}{(2\pi)^3}\iiint \exp[i(\mathbf{k}\bullet\mathbf{r} - \omega t)]d\mathbf{k} .$$

[OPTIONAL]

CHAPTER 4

ONE-DIMENSIONAL QUANTUM SYSTEMS

Thus far, most of our discussions have been rather general. It is time to get more specific and set about solving the Schrödinger equation in several important contexts.

In the present chapter, we restrict ourselves to one-dimensional quantum systems, in which the wavefunction is considered to be a function of a single spatial dimension x, together with time t.

This idealisation, of one spatial dimension, is useful for at least two reasons: (a) the mathematics is simpler, allowing us to develop some useful techniques and insight prior to tacking more complex problems; (b) some systems are well approximated as one dimensional.

§14. Schrödinger equation in one dimension

We may transform the three-dimensional Schrödinger equation (35) into the one-dimensional Schrödinger equation by replacing **r** = (x,y,z) with x.

We thereby arrive at the *one-dimensional form of the time-dependent Schrödinger equation*:

$$(83)\quad i\hbar\frac{\partial}{\partial t}\Psi(x,t) = \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x,t)\right]\Psi(x,t).$$

The corresponding stationary-state solutions Ψ_E are obtained by replacing **r** = (x,y,z) with x in equations (44) and (47), so that:

$$(84)\quad \Psi_E(x,t) = \psi(x)\exp(-iEt/\hbar),$$

$$(85)\quad \left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right]\psi(x) = E\psi(x).$$

Equation (85) is the *one-dimensional form of the time-independent Schrödinger equation*[31].

► *Exercise #27. Write down an expression for the current density* j(x,t) *associated with a 1+1-dimensional*[32] *wavefunction* $\Psi(x,t)$*. Show that, for stationary-state wavefunctions* $\psi(x)$ *in the presence of a time-independent potential* V(x)*, the current density is independent of position. [Note that the current density is now a scalar rather than a vector, hence in this chapter the current density is no longer represented with a boldface* j.*]*

---

[31] Notes: (i) Equation (85) is known as a *Sturm-Liouville equation*; (ii) In passing from (83) to (85), the partial derivative with respect to x has been replaced with d/dx, since ψ(x) depends only on x.

[32] Note that "1+1-dimensional" means "one space dimension and one time dimension". Previously, we have been working with a 3+1-dimensional theory, i.e. a theory with 3 space dimensions and 1 time dimension.

§15. Free particle

For a free particle, V(x) is zero and so (85) becomes the following time-independent Schrödinger equation:

$$(86)\quad -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) = E\psi(x)\,.$$

► *Exercise #28. Show that the general solution to (86) is:*

$$(87)\quad \psi(x) = A\exp(ikx) + B\exp(-ikx)\,,\ \textit{where:}$$

$$(88)\quad k = \frac{\sqrt{2mE}}{\hbar}$$

is the wavenumber of the stationary state solution[33]*, and* A *and* B *are arbitrary complex numbers. Deduce the corresponding expression for the probability current density* j(x)*, investigate the limit cases as A tends to zero and as B tends to zero, and briefly interpret your results.*

Equation (87) can be interpreted as a pair of *counter-propagating one-dimensional plane waves.* These may be compared to the three-dimensional plane waves studied in §4.

§16. Barrier penetration (tunnelling)

Here, we consider the solution to the one-dimensional time-independent Schrödinger equation (85), for the case where the potential V(x) is a "rectangular potential barrier":

$$(89)\quad V(x) = \begin{cases} 0, & x < 0, \\ V_0, & 0 \le x \le a, \\ 0, & x > a. \end{cases}$$

Here, V_0 is a positive real number that specifies the "height" of the potential barrier. This barrier is sketched in Figure 10.

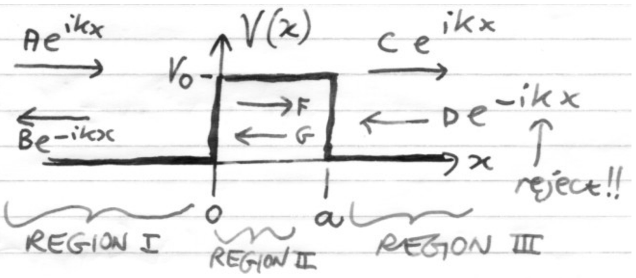


Figure 10: Potential barrier

In both "Region I" and "Region III" (see Fig. 10), the potential V is equal to zero and so (85) reduces to (86). In exercise #28, you showed that the general solution to (86) is given by the pair of counter-propagating plane waves in (87). Thus we may write:

$$(87)\quad \psi_I(x) = A\exp(ikx) + B\exp(-ikx)$$

for the solution in Region I, and:

$$(90)\quad \psi_{III}(x) = C\exp(ikx) + D\exp(-ikx)$$

for the solution in Region III, where A, B, C and D are complex numbers.

Making use of the discussion in §15, we note:

- A exp(ikx) is a plane wave incident upon the left side of the barrier;
- B exp(-ikx) is a plane wave reflected from the left side of the barrier;
- C exp(ikx) is a plane wave that exits the right side of the barrier, and travels to the right;
- D exp(-ikx) is a plane wave incident upon the right side of the barrier.

By assumption, we take D = 0.[34]

We then define the reflection coefficient R, and the transmission coefficient T, via:

$$(91a)\quad R = \frac{|B|^2}{|A|^2}\,,$$

$$(91b)\quad T = \frac{|C|^2}{|A|^2}\,.$$

To proceed further, we separately consider the following two special cases:

- Case #1, where $0 < E < V_0$;
- Case #2, where $E > V_0$.

<u>Case #1</u> $(0 < E < V_0)$

Solving the time-independent Schrödinger equation in Region II, one obtains the following expression for the wavefunction $\psi_{II}(x)$ in this region:

$$(92)\quad \psi_{II}(x) = F\exp(\kappa x) + G\exp(-\kappa x)\,,$$

where F and G are complex numbers, and:

---

[33] Note that in writing (88), we have implicitly assumed the energy to be non-negative.

[34] This amounts to assuming that, while there is a source of electrons (far to the left) which results in a wave incident on the barrier from the left, there is no source of electrons far to the right of the barrier. Thus there is no wave incident on the barrier from the right, and D = 0.

(93) $\kappa = \dfrac{\sqrt{2m(V_0 - E)}}{\hbar}$.

Note that (92) is comprised of real exponentials, which may be compared to the complex exponentials in equations (87) and (90).

To proceed further, recall the boundary conditions for the wavefunction, that were outlined in §10. (a) Since the wavefunction must be continuous at both x = 0 and x = a, (94) $\begin{cases} \psi_{\mathrm{I}}(x=0) = \psi_{\mathrm{II}}(x=0), \\ \psi_{\mathrm{II}}(x=a) = \psi_{\mathrm{III}}(x=a). \end{cases}$

(b) Since the x-derivative of the wavefunction must be continuous at both x = 0 and x = a,

(95) $\begin{cases} \psi'_{\mathrm{I}}(x=0) = \psi'_{\mathrm{II}}(x=0), \\ \psi'_{\mathrm{II}}(x=a) = \psi'_{\mathrm{III}}(x=a). \end{cases}$

Here, a dash denotes differentiation with respect to x.

Using the constraints in (94) and (95), one obtains the following expressions for the reflection coefficient R:

(96) $R = \left[1 + \dfrac{4E(V_0 - E)}{V_0^2 \sinh^2(\kappa a)}\right]^{-1}$

and the transmission coefficient T:

(97) $T = \left[1 + \dfrac{V_0^2 \sinh^2(\kappa a)}{4E(V_0 - E)}\right]^{-1}$.

► *Exercise #29. (a) Show that (92) indeed solves the time-independent Schrödinger equation in Region II, as claimed. (b) Derive (96). (c) Derive (97). (d) Show that:*

(98) $R + T = 1$,

and then briefly interpret this result.

Remarkably, there is a nonzero probability of the particle "getting through" the rectangular potential barrier[35]. This phenomenon is known as *quantum-mechanical tunnelling*, or *barrier penetration*.

One of the most famous applications of this idea is due to Gamow, who applied it to the theory of alpha decay. Scanning tunnelling microscopy, which is a very high-resolution technique for imaging in electron microscopy, is a more contemporary application of the idea of quantum-mechanical tunnelling.

Case #2 $(E > V_0)$

In this case, the wavefunction in Region II is:

(98) $\psi_{\mathrm{II}}(x) = F\exp(iKx) + G\exp(-iKx)$,

where:

(99) $K = \dfrac{\sqrt{2m(E - V_0)}}{\hbar}$.

The boundary conditions—that both the wavefunction and its x-derivative must be continuous at x = 0 and at x = a —yield equations that allow one to write down the following expressions for the reflection and transmission coefficients:

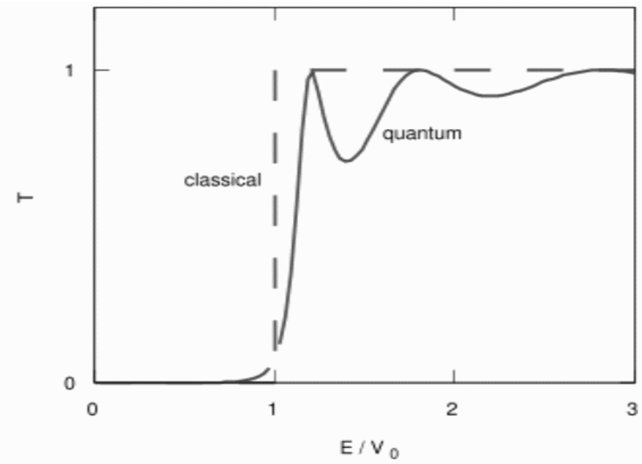
(100a) $R = \left[1 + \dfrac{4E(E - V_0)}{V_0^2 \sin^2(Ka)}\right]^{-1}$,

(100b) $T = \left[1 + \dfrac{V_0^2 \sin^2(Ka)}{4E(E - V_0)}\right]^{-1}$.

► *Exercise #30. (a) Derive (98). (b) Derive (100a). (c) Derive (100b). (d) Show that (98) again holds.*

It is evident from (100b) that the transmission coefficient is in general less than unity, because in general a part of the incident wavefunction is reflected. This differs from the classical prediction, which yields a transmission coefficient of unity.

A plot of (97) and (100b) is given below. I invite you to think about the physical meaning of the various features in this graph.



Transmission probability of a finite potential barrier for $\sqrt{2mV_0}a/\hbar = 7$. Dashed: classical result. Solid line: quantum mechanics.

[source = http://en.wikipedia.org/wiki/Rectangular_potential_barrier, accessed March 12, 2010.]

---

[35] Classically, the probability of "getting through" the barrier is zero.

§17. Infinite square well

Here, we consider the solution to the one-dimensional time-independent Schrödinger equation (85), for the case where the potential V(x) is an "infinite square well":

$$(101)\quad V(x) = \begin{cases} 0, & -a < x < a, \\ \infty, & \text{otherwise.} \end{cases}$$

This potential is sketched in Figure 11 below.

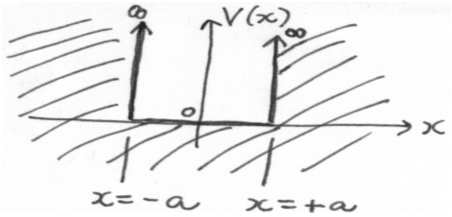


Figure 11: Infinite square well

The wavefunction must vanish when $x > a$ or $x < -a$. Continuity of the wavefunction (see "Condition I" in §10) then implies that the wavefunction vanishes at x = - a and at x = a.

Bearing the above in mind, the normalised energy eigenfunctions $\psi_n(x)$, which are labelled by the positive integer n, are:

$$(102)\quad \psi_n(x) = \begin{cases} \frac{1}{\sqrt{a}} \cos\frac{n\pi x}{2a}, & n = 1,3,5,\cdots, \\ \frac{1}{\sqrt{a}} \sin\frac{n\pi x}{2a}, & n = 2,4,6,\cdots. \end{cases}$$

The corresponding energy levels, denoted E_n, are quantised according to:

$$(103)\quad E_n = \frac{\hbar^2 \pi^2 n^2}{8ma^2}.$$

A plot, of the first four energy eigenfunctions in (102), is given below.

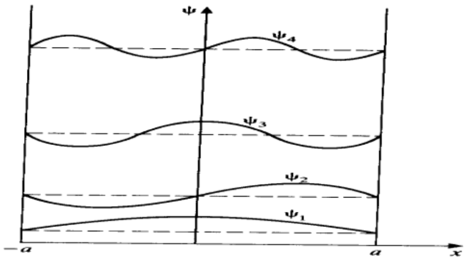


Figure 12: First four energy eigenfunctions for the infinite square well[36]

► *Exercise #31. (a) Show that the wave-functions in (102) obey the time-independent Schrödinger equation in the region* -a < x < a. *(b) Show that these wavefunctions are normalised to unity. (c) Show that these wavefunctions are continuous at both x = -a and x = a. (d) Derive (103).*

§18. Harmonic oscillator

The harmonic oscillator, treated in this section, is one of the most important constructs in quantum mechanics. Indeed, the central result of this section, namely equation (121), is applied in a vast variety of contexts in theoretical physics, from condensed matter through to quantum electrodynamics and particle cosmology. You are therefore likely to encounter the simple harmonic oscillator again and again, in your future studies in physics.

Consider the plot of a potential W(x) as a function of position x, as given by the solid line in Figure 13. At the point x = a, the potential W has a local minimum.

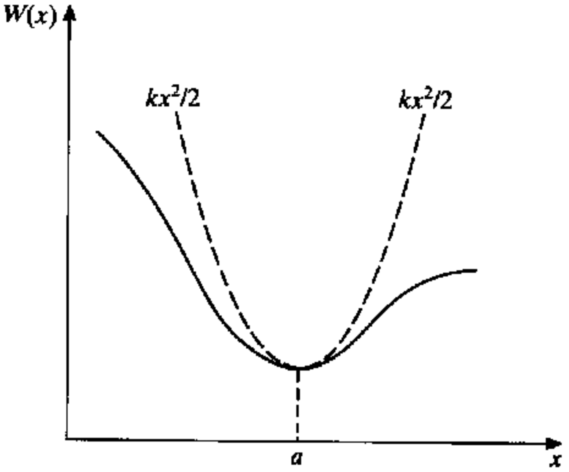


Figure 13: Plot of a continuous potential function W(x), versus position x (solid line). This potential has a local minimum at the point x = a. A quadratic Taylor approximation to the potential, about the point x = a, is given by the dotted line.[37]

If a stationary classical point particle were to be placed at the point x = a, it would remain stationary at this point. If, instead, the stationary classical particle were to be slightly displaced from its equilibrium point at x = a, and then "let go", it would exhibit oscillatory motion about x = a.

For small displacements about the equilibrium point x = a, the potential W(x) may be approximated as a

---

[36] Image taken from B.H. Bransden and C.J. Joachain, *Introduction to Quantum Mechanics*, John Wiley and Sons, New York (1989), page 156.

[37] Image taken from B.H. Bransden and C.J. Joachain, *Introduction to Quantum Mechanics*, John Wiley and Sons, New York (1989), page 166.

quadratic function of x—see dotted curve in Figure 13.

By shifting the origin of the x axis to x = a, and shifting the origin of the vertical axis to $W(x=a)$, the quadratic approximation to the potential can be written as:

$$(104)\quad V(x)=\frac{1}{2}kx^2 .$$

Here, k is a constant, and the potential is now called "V" rather than "W".

► *Exercise #32. Justify (104), using an argument based on the Taylor-series expansion of the potential* W(x)*, about the point* x = a*. Hence deduce an expression for* k*. [OPTIONAL]*

Since the potential in (104) is quadratic in x, the associated force F, as given by -dV/dx, is equal to -kx. This will be recognized as a *restoring force that is linear in the displacement of the particle from equilibrium*—as a familiar example, Hooke's law, for a particle on a spring, has such a form. The associated classical motion is sinusoidal in time, this being termed *simple harmonic motion* in a *simple harmonic oscillator*.

It is evident from the above discussion that the idea, of small oscillations of a particle about an equilibrium point, is one of great generality. For example, one may have (a) a ball on a spring oscillating about its equilibrium position, or (b) an atom in a crystal lattice oscillating about its equilibrium position, or (c) an electron in a nanometre-sized potential well oscillating about its equilibrium position, and so forth. As this list is intended to suggest, the study of the simple harmonic oscillator is applicable to a very broad range of scenarios across theoretical physics.

Now, we wish to pass from the *classical* theory of the simple harmonic oscillator, to the corresponding *quantum* theory. Accordingly, we may write down the following one-dimensional time-independent Schrödinger equation, corresponding to the potential in (104)[38]:

$$(105)\quad \left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}+\frac{1}{2}kx^2\right)\psi(x)=E\psi(x) .$$

To simplify this expression, make the following change of variables:

(106a) $\lambda=2E/(\hbar\omega)$ — scaled form of energy,

(106b) $\omega=\sqrt{k/m}$ — angular frequency of classical oscillator,

(106c) $\xi=\alpha x$ — scaled form of distance,

(106d) $\alpha=\left(\frac{mk}{\hbar^2}\right)^{1/4}=\sqrt{\frac{m\omega}{\hbar}}$.

With the above change of variables, (105) takes the form[39]:

$$(107)\quad \frac{d^2}{d\xi^2}\psi(\xi)+(\lambda-\xi^2)\psi(\xi)=0 .$$

► *Exercise #33. Show that (105) leads to (107), upon making the change of variables given in equations (106).*

It will prove convenient for us to study the "asymptotic form" of equation (107), namely the form that this equation takes when ξ is very large in magnitude.

When $|\xi|$ is very large, we may replace $\lambda-\xi^2$ with $-\xi^2$, so that (107) becomes:

$$(108)\quad \frac{d^2}{d\xi^2}\psi(\xi)=\xi^2\psi(\xi),\quad |\xi|\to\infty .$$

The above equation has the following (decaying) solutions:

(109) $\psi(\xi)=\xi^p\exp(-\xi^2/2)$, for finite p.

► *Exercise #34. (a) Show that (109) is indeed a solution to (108), for finite p. (b) Show that* $\psi(\xi)=\xi^p\exp(+\xi^2/2)$ *is also a solution, and then explain why this solution is not being considered in the present context.*

Now, (109) tells us the "asymptotic" behaviour of the wavefunction, i.e. the behaviour of the wavefunction as ξ becomes very large in magnitude. However, we are interested in the solution to (107) for *all* ξ, not just the ξ's with a large magnitude. The solution $\psi(\xi)$ to (107), whatever it is, should have the correct

---

[38] Cf. equation (85).

[39] The form of (107) is convenient because all physical constants have been subsumed into a single parameter, the scaled energy λ. The spatial wavefunction ψ now depends on the scaled distance ξ.

behaviour as $|\xi|$ tends to infinity. This will be the case if we write $\psi(\xi)$ as:

$$(110)\quad \psi(\xi) = \exp(-\xi^2/2)H(\xi),$$

where $H(\xi)$ is a finite-order polynomial known as a "Hermite polynomial"[40].

Substitute (110) into (107) to give the so-called "Hermite equation":

$$(111)\quad \frac{d^2}{d\xi^2}H(\xi) - 2\xi\frac{d}{d\xi}H(\xi) + (\lambda - 1)H(\xi) = 0.$$

► *Exercise #35. Derive (111).*

We separately seek even and odd solutions to the Hermite equation (111). Note that the even solutions to (111) obey:

$$(112a)\quad H(-\xi) = H(\xi),$$

with the odd solutions obeying:

$$(112b)\quad H(-\xi) = -H(\xi).$$[41]

Even solutions to equation (111)

The as-yet unknown even Hermite polynomials may be written as a Taylor series:

$$(113)\quad H(\xi) = \sum_{k=0}^{\infty} c_k \xi^{2k},$$

where we have kept only even powers of ξ on account of (112a).

Substitute (113) into the Hermite equation (111) to give:

$$(114)\quad \sum_{k=0}^{\infty}\begin{bmatrix} 2(k+1)(2k+1)c_{k+1} \\ -(4k+1-\lambda)c_k \end{bmatrix}\xi^{2k} = 0.$$

► *Exercise #36. Derive (114).*

Since the left side of (114) must vanish for any value of ξ, the term in square brackets must be zero, allowing us to write down:

$$(115)\quad c_{k+1} = \frac{4k+1-\lambda}{2(k+1)(2k+1)}c_k.$$

This is known as a "recursion relation"[42].

Now,

$$(116)\quad \frac{c_{k+1}}{c_k} = \frac{4k+1-\lambda}{2(k+1)(2k+1)} \to \frac{1}{k} \text{ for large k.}$$

This is "bad", insofar as it implies the following "blowup" behaviour for the wavefunction, as ξ becomes very large in magnitude:

$$(117)\quad \psi(\xi) \to \xi^{2p}\exp(\xi^2),\quad |\xi| \to \infty.$$

► *Exercise #37. Derive (117).*

To avoid the "blowup", all we need to do is to ensure that the series in (113) terminates after a finite number of terms[43]. This is easily arranged by demanding that, for some value of k (denote this particular value of k by "N"), the numerator in (115) should vanish. Indeed, if $4k+1-\lambda$ vanishes when k = N, then (115) implies $c_{N+1} = 0 \times c_N = 0$; c_{N+2}, c_{N+3}, and all remaining coefficients also vanish. Thus we have the required behaviour that (113) terminates after a finite number of terms.

Since $4k+1-\lambda$ is equal to zero when k = N,

$$(118)\quad \lambda = 4N+1,\ N = 0, 1, 2, \ldots$$

Using (106a), this becomes the following expression for the energy levels of the harmonic oscillator that are associated with the even Hermite polynomials:

$$(119)\quad E = (2N + \tfrac{1}{2})\hbar\omega,\ N = 0, 1, 2, \ldots$$

---

[40] Here, we briefly justify the construction in (110): Since $H(\xi)$ is a finite-order polynomial in ξ, we can write $H(\xi) = A + B\xi + C\xi^2 + \ldots + Z\xi^p$, where A, B, ..., Z are constants, and p is a finite integer. Equation (110) then becomes:
(i) $\psi(\xi) = \exp(-\xi^2/2)[A + B\xi + C\xi^2 + \ldots + Z\xi^p]$.
As ξ becomes very large in modulus, the term $Z\xi^p$ dominates over all of the other terms in square brackets in (i), so that the wavefunction tends towards $\psi(\xi) = Z\exp(-\xi^2/2)\xi^p$. Thus the wavefunction in (110) indeed has the correct limiting behaviour, as given in equation (109). [Note, also, that the wavefunction tends to zero as ξ becomes infinitely large in magnitude—hence we spoke of (109) as a "decaying" solution, in the main text.]

[41] Note that an even wavefunction is said to have "even parity", while an odd wavefunction is said to have an "odd parity".

[42] Suppose that you are given a particular value for the coefficient c_0 in (113). The k = 0 case of equation (115) will then tell you how to obtain the coefficient c_1 from the coefficient c_0. The k = 1 case of (115) will then tell you how to obtain the coefficient c_2 from the coefficient c_1. Proceeding *ad infinitum,* one thus (recursively) obtains all of the coefficients c_k in the expansion (113). Hence (115) is spoken of as a "recursion relation".

[43] We will then have the asymptotic behaviour as described in the footnote immediately after equation (110), in which case there is no "blowup".

Odd solutions to equation (111)

Using analogous logic to that given above, the odd solutions to equation (111) imply the following allowed energy levels for the harmonic oscillator:

(120) $E = \hbar\omega(2N + 1 + \frac{1}{2})$ $N = 0, 1, 2, ...$

► *Exercise #38. Derive (120).*

From (119) and (120), we see that the allowed energy levels of the quantum harmonic oscillator are:

(121) $E_n = \hbar\omega(n + \frac{1}{2})$, $n = 0, 1, 2, ...$

Note the subscript "n", on the energy E, to indicate the energy's dependence on the quantum number n.

Remarks:

- The minimum energy $E_0 = \frac{1}{2}\hbar\omega$, corresponding to $n = 0$ in (121), is nonzero. This non-zero minimum energy is known as the *ground state energy.* (This differs from the classical result, for which the minimum energy of the oscillator is zero.)
- The allowed energy levels are quantised, and equally spaced. (This differs from the classical result, in which the harmonic oscillator can attain a continuum of allowed energy values.)

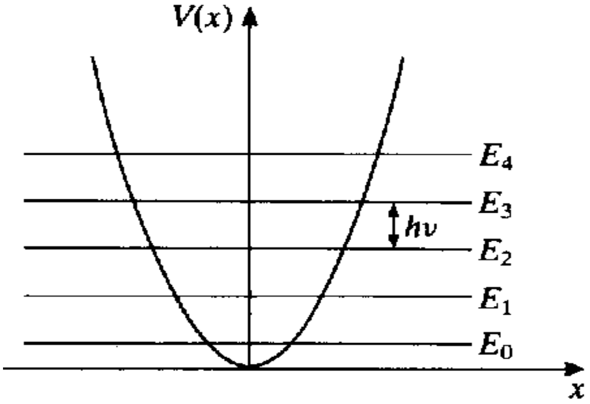


Figure 14: Allowed energy levels E_n, for the simple harmonic oscillator. Spacing between adjacent energies is $\hbar\omega = h\nu$, where ν is the frequency of the corresponding classical motion.[44]

§19. Gaussian wavepacket

For much of this section, we have been concerned with solutions to the time-*independent* Schrödinger equation in one spatial dimension. For the remaining two sections of the chapter, we will bring time into the picture and study solutions to the one-dimensional time-*dependent* Schrödinger equation.

The Gaussian wavepacket, which is an exact localized solution to the time-dependent Schrödinger equation (83), is given by:

(122)
$$\Psi(x,t) = \frac{1}{\pi^{1/4}}\sqrt{\frac{\Delta p_x/\hbar}{1+\frac{i(\Delta p_x)^2 t}{m\hbar}}} \times \exp\left[\frac{\frac{ip_0 x}{\hbar} - \left(\frac{\Delta p_x}{\hbar}\right)^2 \frac{x^2}{2} + \frac{ip_0^2 t}{2m\hbar}}{1+\frac{i(\Delta p_x)^2 t}{m\hbar}}\right].$$

Here, p_0 is the average momentum of the plane waves $\exp[i(k_x x - \omega t)]$ that are superposed to make up the Gaussian wavepacket, and Δp_x is the spread in momenta of the planes waves from which the Gaussian wavpacket is composed.[45]

► *Exercise #39. Show that (122) indeed obeys the time-dependent Schrödinger equation (83), for the case of a zero potential.*

By taking the squared modulus of (122), one obtains the following expression for the probability density associated with the Gaussian wavepacket:

(123)
$$\rho(x,t) = \frac{1}{\sqrt{\pi}} \frac{\Delta p_x/\hbar}{\sqrt{1+\frac{(\Delta p_x)^4 t^2}{m^2\hbar^2}}} \exp\left[-\frac{\left(\frac{\Delta p_x}{\hbar}\right)^2 \{x - vt\}^2}{1+\frac{(\Delta p_x)^4 t^2}{m^2\hbar^2}}\right],$$

where $v = p_0/m$ is the expectation value of the velocity of the wavepacket[46].

► *Exercise #40. Obtain (123) from (122).*

An indicative plot, of the probability density (123) associated with the Gaussian wavepacket, is given in Figure 15 below.

---

[44] Image taken from B.H. Bransden and C.J. Joachain, *Introduction to Quantum Mechanics*, John Wiley and Sons, New York (1989), page 166.

[45] Stated differently, p_0 is the expectation value of the momentum for the Gaussian wavepacket, and Δp_x is the corresponding uncertainty in this expectation value.

[46] Cf. equation (80), i.e., Ehrenfest's second equation.

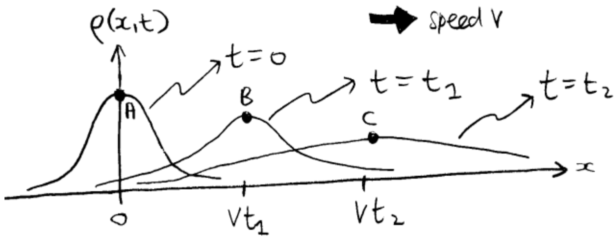


Figure 15: Moving and spreading of the Gaussian wavepacket.

Salient features of (123), as illustrated in Figure 15, include:

- At any given time t, the probability density has the form of a Gaussian (that's why it's called a "Gaussian wavepacket");
- At any given time t, the Gaussian is peaked at the position $x = vt$ (see term in braces in (123)). Thus the peak, of the Gaussian wavepacket, moves to the right at speed v.
- The width of the Gaussian wavepacket increases with time, a phenomenon known as the *spreading of the wavepacket.* Normalisation of the wavefunction implies that, as it gets wider, its height reduces, so that the area under the curve (in the Figure) remains constant.

► *Exercise #41. Show that the width* $\Delta x(t)$ *of the Gaussian wavepacket, defined by the positions at which the probability density falls to 1/e of its maximum value, is given by:*

$$(124)\quad \Delta x(t) = \frac{\hbar}{\Delta p_x}\sqrt{1+\frac{(\Delta p_x)^4 t^2}{m^2\hbar^2}}\,.$$

Note that this width increases with time, in accord with our remarks on the spreading of the wavepacket.

§20. Scattering of Gaussian from potential barrier

Here, we briefly consider the problem of barrier penetration by a Gaussian wavepacket. Since this problem does not admit an exact solution, the time-dependent Schrödinger equation must be solved numerically.

In Figure 16, we see a sequence of numerically-obtained "movie frames" of the probability density associated with a Gaussian wavepacket that is scattered from a potential barrier. Note that the expectation value of the energy, for the Gaussian wavepacket, is higher than the height of the potential barrier.

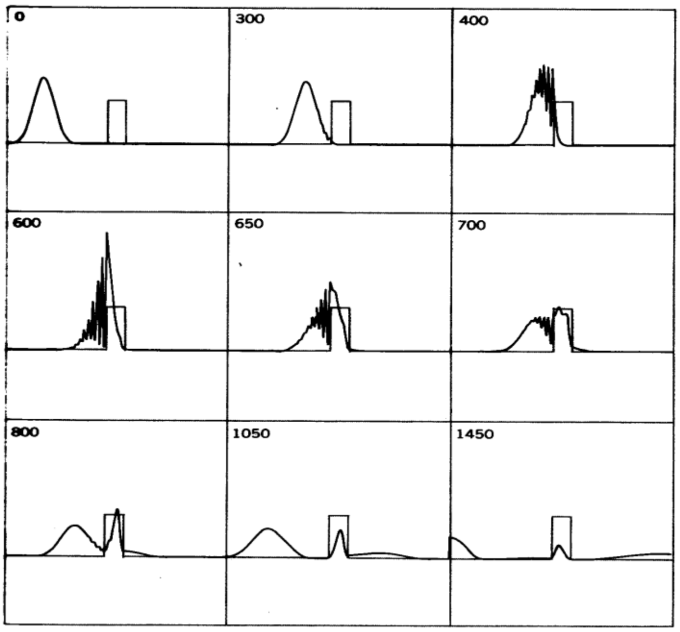


Figure 16: Scattering of a Gaussian wavepacket from a potential barrier, for the case where the mean energy of the wavepacket is larger than the height of the barrier. Time (in arbitrary units) is indicated by the number in the top left corner of each frame.[47]

- In the top left frame of Figure 16 (t=0), we see the incident Gaussian wavepacket, which is travelling to the right. At t=300, the Gaussian has moved closer towards the potential barrier (cf. second dot point on previous page).
- In the subsequent frames (t=400, 600, 650 and 700), we see some strong spatial oscillations being set up, due to the interference between the incident and reflected wavepackets.
- At t=700 we just begin to see a transmitted wavepacket emerging from the right of the barrier.
- At t=1050, the reflected wavepacket is no longer strongly interacting with the barrier; this reflected wavepacket is moving away from the barrier, and travelling to the left. The transmitted wavepacket may also be seen in this panel.
- Note also that there is "resonance" effect, whereby an appreciable amount of probability density is

---

[47] Image taken from L.I. Schiff, *Quantum Mechanics* (third edition), McGraw–Hill Book Company, New York (1968), page 107. Original source: A. Goldberg, H.M. Schey and J.L. Schwartz, *American Journal of Physics*, volume 35, page 177 (1967).

trapped within the walls of the potential barrier; this trapped probability density slowly "leaks out" as time passes. The said resonance is evident in the final four panels (t=700, 800, 1050 and 1450).

CHAPTER 5

FORMALISM

In the previous chapters, we studied *wave mechanics*, in which a particular representation of the state of a quantum system—namely the wavefunction Ψ, considered as a complex function of one or more position coordinates **r**, together with time t—occupied a prominent role.

A more general formulation of quantum mechanics is the subject of the present chapter. This formalism is more powerful than wave mechanics; indeed, wave mechanics may be considered as a special case of the more general formulation.

At the core of this general formalism is a series of postulates, some of which will be familiar from earlier chapters, whose utility is ultimately linked to the successes of the theory in consistently predicting the results of experiments.

The purpose of this chapter is to acquaint you with the more general formalism of quantum mechanics. This formalism will then be freely drawn upon in subsequent chapters of the course.

§21. State of a system

Consider a *classical* system that consists of a given number N of point particles. The "dynamical state" of the system, namely the state of the system at a given instant of time, is evidently a meaningful construct. Indeed, the dynamical state of the system, at a given time t, is completely specified by listing the position and momentum of each of the N point particles in the system. Note, also, that one would speak of the positions and momenta of the particles as the *dynamical variables* associated with the system.

On account of the Heisenberg uncertainty principles in equation (74a), one cannot specify both positions and momenta with arbitrary accuracy. Indeed, these are complementary variables, in the sense of Bohr's complementarity principle. One must therefore discard the notion of a precise dynamical state of a system, which was such a core construct in the classical era. Rather, one has the weaker notion of quantum mechanics as a statistical theory, that is, one in which the wavefunction furnishes probabilities associated with certain experimental outcomes, while being unable (in general) to precisely predict the outcomes of all experiments.

Notwithstanding the statistical character of the quantum theory, together with the deconstruction of the notion of a quantum system having a precise dynamical state in the strong classical sense of the term, one may without logical contradiction speak of the wavefunction as providing *maximal knowledge* of a given quantum system. Hence our first postulate:

POSTULATE 1: One can in principle associate, with a given quantum system, a wavefunction Ψ that captures all that is knowable about the quantum system.

* * *

Under the Born interpretation, the squared modulus of the wavefunction is related to the probability of measuring a particle to be in a given small region of space at a given time.

(a) For a single-particle quantum system described by the wavefunction $\Psi(\mathbf{r},t)$, we saw in equation (21) that:

$$(21)\quad \rho(\mathbf{r},t)d\mathbf{r} = |\Psi(\mathbf{r},t)|^2\, d\mathbf{r}$$

is the probability of finding the particle within an infinitesimal volume d**r** located at position **r**. Since the probability is unity, of finding the electron somewhere, we have the normalisation condition:

$$(22)\quad \iiint |\Psi(\mathbf{r},t)|^2\, d\mathbf{r} = 1.$$

(b) The above may be generalised. Consider an N-particle quantum system. Let $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N$ respectively denote the position coordinates associated with the particles 1 through N. Then the state of the system is described by the (many-body) wavefunction:

$$(125)\quad \Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N, t),$$

with:

$$(126)\quad \begin{aligned} &\rho(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N, t)d\mathbf{r}_1 d\mathbf{r}_2 \cdots d\mathbf{r}_N \\ &\quad = |\Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N, t)|^2\, d\mathbf{r}_1 d\mathbf{r}_2 \cdots d\mathbf{r}_N \end{aligned}$$

being interpreted as the probability of finding the first particle within an infinitesimal volume d$\mathbf{r}_1$ located at position $\mathbf{r}_1$, *and* finding the second particle within an infinitesimal volume d$\mathbf{r}_2$ located at position $\mathbf{r}_2$, *and* finding the third particle within an infinitesimal volume d$\mathbf{r}_3$ located at position $\mathbf{r}_3$, etc. The many-body generalisation of (22) may also be written down, although this will not be given here.

We shall not discuss many-body theory any further in this course. Note, however, that it becomes extremely important in many more advanced studies, such as the theory of condensed matter, together with quantum optics and the theory of Bose condensates.

* * *

Now for the second postulate, which we have met before (see §6.2):

POSTULATE 2: Quantum-mechanical states obey the superposition principle.

Thus if Ψ_A and Ψ_B are two allowed states of a quantum system, then the superposition $\alpha\Psi_A + \beta\Psi_B$ is also an allowed state, where α and β are arbitrary complex numbers.

* * *

We close this section by introducing a notation due to Dirac. Suppose that one has two single-particle wavefunctions $\Psi_1(\mathbf{r},t)$ and $\Psi_2(\mathbf{r},t)$. Then, by definition,

(127) $\langle\Psi_1|\Psi_2\rangle \equiv \iiint \Psi_1^*(\mathbf{r},t)\Psi_2(\mathbf{r},t)d\mathbf{r}$.

If, on the other hand, one has two multiparticle wavefunctions $\Psi_1(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N,t)$ and $\Psi_2(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N,t)$, then (127) generalises to:

(128)
$$\langle\Psi_1|\Psi_2\rangle \equiv \int\cdots\int d\mathbf{r}_1 d\mathbf{r}_2\cdots d\mathbf{r}_N \Psi_1^*(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N,t)\Psi_2(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N,t).$$

► *Exercise #42. Prove the following:*

(129) $\langle\Psi_1|\Psi_2\rangle = \langle\Psi_2|\Psi_1\rangle^*$

(130) $\langle c\Psi_1|\Psi_2\rangle = c^*\langle\Psi_1|\Psi_2\rangle$

(131) $\langle\Psi_3|\Psi_1+\Psi_2\rangle = \langle\Psi_3|\Psi_1\rangle + \langle\Psi_3|\Psi_2\rangle$

Here, Ψ_1, Ψ_2, Ψ_3 are wavefunctions (assume, for simplicity, that they are single-particle wavefunctions, but note that this assumption is not necessary!), and c is any complex number.

Three remarks:

(a) Two wavefunctions are said to be orthogonal if:[48]

(132) $\langle\Psi_1|\Psi_2\rangle = 0$.

(b) The normalisation condition (22) may be written in Dirac notation as:

(133) $\langle\Psi|\Psi\rangle = 1$.

(c) The symbol $\langle\Psi|$ is known as a "bra", while $|\Psi\rangle$ is known as a "ket". This notation arises since a bra and a ket, when put side by side (see e.g. left side of (133)), gives a "bracket".

§22. Operators revisited

POSTULATE 3: To each dynamical variable $\mathcal{A}$ of a quantum system, one has an associated linear operator[49] A.

We have already seen several examples of the above, such as the Schrödinger correspondence rules (30) and (31).

* * *

We now briefly revise the notion of eigen-values and eigenfunctions: a function ψ_n is said to be an eigenfunction of the linear operator A if:

(134) $A\psi_n = a_n\psi_n$.

Here, the number a_n is the eigenvalue corresponding to the eigenfunction ψ_n.

POSTULATE 4: When one measures the dynamical variable $\mathcal{A}$ associated with a given quantum system, one obtains one of the eigenvalues of the associated linear operator A.

The set of all eigenvalues of A is known as the *spectrum* of A. Since physical measurements must yield real numbers, the spectrum of A must be real if A is to represent a dynamical variable.

* * *

Earlier in the course, we argued that linear operators A, associated with a given dynamical variable $\mathcal{A}$,

---

[48] Note that analogy with vectors: two vectors are orthogonal if their dot product is zero.

[49] Note that a definition of a "linear operator" was given in the footnote to exercise 19.

must be Hermitian (see eqn (68)). In Dirac notation, the condition for an operator A to be Hermitian may be written as:

(135) $\langle \Psi | A\Psi \rangle = \langle A\Psi | \Psi \rangle$.

Equivalently, we state that an operator A is Hermitian if:

(136) $\langle \Psi_1 | A\Psi_2 \rangle = \langle A\Psi_1 | \Psi_2 \rangle$,

for any pair of square-integrable wavefunctions Ψ_1 and Ψ_2.

► *Exercise #43. Using the definition of Hermiticity in (135), show, by expressing the wave function* Ψ *in the form* $\Psi = c_1\Psi_1 + c_2\Psi_2$, *that (136) follows as a direct consequence. Therefore (135) and (136) are equivalent definitions of Hermiticity, as claimed in the main text.*

* * *

If, at a given time, a given quantum system has a wavefunction that is an eigenfunction of the operator A associated with a given dynamical variable, then the system is said to be in an *eigenstate* of the operator A.

When a quantum system is in an eigenstate of A, there is a 100% chance of measuring the value of the corresponding dynamical variable to have the value a_n, this being the eigenvalue associated with the eigenstate ψ_n (see eqn (134)).

When a quantum system is *not* in an eigenstate of A, then there are a range of possible values allowed when one makes a measurement of the corresponding dynamical variable $\mathcal{A}$.[50] In making the measurement, one might obtain any of the eigenvalues a_n in the spectrum of A.[51] Expressions for the associated probabilities, of measuring any particular eigenvalue in the spectrum, will be obtained in due course.

* * *

POSTULATE 5: When one measures the dynamical variable A, the average value (expectation value) of the measurement is:[52]

(137) $\langle \Psi | A\Psi \rangle \equiv \langle \Psi | A | \Psi \rangle$.

* * *

Associated with the linear operator A, one may define the associated "adjoint operator" $A^\dagger$ (this is also known as the "Hermitian conjugate operator"). By definition,

(138) $\langle \Psi_1 | A^\dagger \Psi_2 \rangle = \langle A\Psi_1 | \Psi_2 \rangle$,

where Ψ_1 and Ψ_2 are arbitrary wavefunctions.

If a linear operator A is equal to its adjoint, i.e., (139) $A = A^\dagger$,

then A is said to be "self adjoint" or "Hermitian" (cf. (136)).

► *Exercise #44. If* A *and* B *are linear operators, and* c *is any complex number, show that:*

(140) $(cA)^\dagger = c^* A^\dagger$,

(141) $(AB)^\dagger = B^\dagger A^\dagger$.

* * *

Corresponding to the linear operator A, the inverse operator A^{-1} (if it exists) obeys:

(142) $A^{-1}A = AA^{-1} = I$.

Here, I is the so-called "unit operator". By definition, when the unit operator acts on any given wavefunction Ψ, the result is simply equal to the wavefunction:

(143) $I\Psi = \Psi$.

* * *

An operator U is said to be *unitary* if $U^\dagger$ is equal to U^{-1}:

(144) $U^\dagger = U^{-1}$.

Thus, for a unitary operator,

(145) $U^\dagger U = UU^\dagger = I$.

Note that the symbol U is usually reserved for unitary operators.

§23. Eigenfunction expansions

Recall the discussions of §5. There, we gave several examples of instances in which a complex object may be decomposed as a weighted sum of simpler objects. The "eigenfunction expansion", for a wavefunction, is an example of such a decomposition.

---

[50] Henceforth, we will use the same symbol A to denote both the dynamical variable and its corresponding linear operator.

[51] Sometimes this spectrum is discrete (e.g. in the case of the harmonic oscillator, where there was seen to be a countable infinity of eigenvalues), and sometimes it is continuous (e.g. the energy of a plane wave can take a continuous range of values). Often, there are both discrete and continuous parts to the spectrum of an operator.

[52] Note that this is just equation (66), in Dirac notation. Note, also, that the wavefunction has been assumed to be normalised.

An "eigenfunction expansion" involves expanding a given wavefunction as a weighted sum of eigenfunctions of a given operator. Here, we show how such an expansion may be obtained.

Recall (134), namely:

(134) $A\psi_n = a_n\psi_n$.

As discussed earlier, the statement "ψ_n is an eigenfunction of the operator A", means that "when A acts on the eigenfunction ψ_n, one obtains the eigenvalue a_n multiplied by the eigenfunction ψ_n." Here, n is a discrete index which is used to label the various eigenfunctions ψ_n.[53] Note, also, that A will assumed Hermitian in what follows.

Assume the eigenfunctions ψ_n to have been normalised (i.e., to obey (22)). In Dirac notation, the normalisation condition reads:

(146) $\langle\psi_n|\psi_n\rangle = 1$.

We now prove that *eigenfunctions belonging to different eigenvalues are orthogonal*. To this end, consider two different eigenfunctions ψ_i and ψ_j of A, the associated eigenvalues being termed a_i and a_j. Thus:

(147) $A\psi_i = a_i\psi_i$ and $A\psi_j = a_j\psi_j$.

Hence[54]:

$$(148)\quad \begin{aligned}(a_i - a_j)\langle\psi_i|\psi_j\rangle &= a_i\langle\psi_i|\psi_j\rangle - a_j\langle\psi_i|\psi_j\rangle\\ &= \langle a_i^*\psi_i|\psi_j\rangle - \langle\psi_i|a_j\psi_j\rangle\\ &= \langle a_i\psi_i|\psi_j\rangle - \langle\psi_i|a_j\psi_j\rangle\\ &= \langle A\psi_i|\psi_j\rangle - \langle\psi_i|A\psi_j\rangle\\ &= \langle A\psi_i|\psi_j\rangle - \langle A\psi_i|\psi_j\rangle\\ &= 0.\end{aligned}$$

Since the eigenvalues a_i and a_j are distinct, by assumption, the left side's term in round brackets is non-zero ($i \neq j$). Thus $\langle\psi_i|\psi_j\rangle = 0$ if i and j are not equal. Hence *eigenfunctions, belonging to different eigenvalues, are orthogonal to one another.*

On the other hand, $\langle\psi_i|\psi_j\rangle = 1$ when i and j are equal, on account of (146). Thus we may write:

(149) $\langle\psi_i|\psi_j\rangle = \delta_{ij}$.

Note: Eigenfunctions belonging to the *same* eigenvalues[55] are not necessary orthogonal, but they may be orthogonalised using the so-called *Schmidt (or Gram-Schmidt) orthogonalisation procedure.*[56] Thus we can consider *all* eigenfunctions, whether or not they possess distinct eigenvalues, to be orthogonal in the sense given by (149).

* * *

As we shall see, the orthogonality of the eigenfunctions is a very useful tool in constructing eigenfunction expansions. It is to this latter subject that we now turn.

POSTULATE 6: The wavefunction for a given quantum state can be decomposed as a linear combination of eigenfunctions of A, where A represents a given dynamical variable.

For the case of a discrete/countable set of eigenvalues, to which we restrict our consideration, the desired eigenfunction expansion is:

(151) $\Psi = \sum_n c_n\psi_n$.

Here, the set of numbers c_n are a series of complex *expansion coefficients*. We now need to work out how to express these coefficients in terms of the wavefunction.

To this end, pre-multiply both sides of (151) by ψ_m^*, and then integrate over all spatial coordinates. Thus (cf. (127) and (128)):

$$(152)\quad \begin{aligned}\langle\psi_m|\Psi\rangle &= \langle\psi_m|\textstyle\sum_n c_n\psi_n\rangle\\ &= \textstyle\sum_n c_n\langle\psi_m|\psi_n\rangle\\ &= \textstyle\sum_n c_n\delta_{mn}\\ &= c_m\end{aligned}$$

---

[53] In this section, we assume the eigenvalues of A to be discrete (countable). The discussion can be generalised to treat continuous (uncountable) eigenvalues, although we do not need such a generalisation for the purposes of this course. See, for example, pp. 202–5 of B.H. Bransden and C.J. Joachain, *Introduction to Quantum Mechanics*, John Wiley and Sons, New York (1989).

[54] Notes: (a) In passing from the second to the third line of (148), we have made use of the fact that a_i is real; (b) In passing from the third to the fourth line, we have made use of (134); (c) In passing from the fourth to fifth line, we have made use of the fact that A is Hermitian.

[55] Such eigenvalues are known as *degenerate* eigenvalues.

[56] We will not discuss such a procedure in this course. See any text on linear algebra, and/or speak with me, if you would like further information on this point.

Hence we have the desired expression for a given expansion coefficient c_n, in terms of the wavefunction Ψ:

(153) $c_n = \langle \psi_n | \Psi \rangle$.

Using (153), equation (151) becomes:

(154) $\Psi = \sum_n \langle \psi_n | \Psi \rangle \psi_n$.

* * *

Result (154)—and, indeed, many of the results of this section—admits a very strong geometric parallel with the theory of vectors. Here, we briefly outline this interpretation.

Consider a vector **v** in two-dimensional space. Using Cartesian coordinates (x and y), we can set up the orthogonal unit vectors $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$, which obey a direct analogue of (149), namely:

$$\hat{\mathbf{x}} \bullet \hat{\mathbf{x}} = 1,\ \hat{\mathbf{y}} \bullet \hat{\mathbf{y}} = 1,\ \hat{\mathbf{x}} \bullet \hat{\mathbf{y}} = 0,\ \hat{\mathbf{y}} \bullet \hat{\mathbf{x}} = 0.$$

The x-component of **v** is $\hat{\mathbf{x}} \bullet \mathbf{v}$, and the y-component of **v** is $\hat{\mathbf{y}} \bullet \mathbf{v}$. Hence we may write the following decomposition of **v**:

(155) $\mathbf{v} = (\hat{\mathbf{x}} \bullet \mathbf{v})\hat{\mathbf{x}} + (\hat{\mathbf{y}} \bullet \mathbf{v})\hat{\mathbf{y}}$.

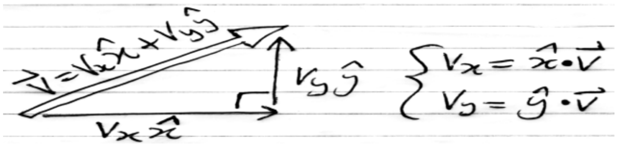


Figure 17: Decomposition of a vector v in terms of its x- and y-components.

Equations (155) and (154) are directly analogous to one another. I invite you to think more deeply about this parallel.

* * *

► *Exercise #45. Take equation (154), write the Dirac bracket explicitly as an integral over all space, and then interchange the order of the summation and the integrals. Hence, or otherwise, obtain the "closure relation":*

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

§24. Probability amplitudes

Make use of the eigenfunction expansion (151) to study the expectation value associated with a given Hermitian operator A:[57]

$$\begin{aligned}
\langle A \rangle &= \langle \Psi | A | \Psi \rangle \text{ ... now use (151)} \\
&= \left\langle \textstyle\sum_m c_m \psi_m \middle| A \middle| \sum_n c_n \psi_n \right\rangle \\
&= \sum_m \sum_n c_m^* c_n \langle \psi_m | A | \psi_n \rangle \text{ ... now use (134)} \\
&= \sum_m \sum_n c_m^* c_n \langle \psi_m | a_n | \psi_n \rangle \text{ ... now use (149)} \\
&= \sum_m \sum_n c_m^* c_n a_n \delta_{mn} \\
&= \sum_n | c_n |^2 a_n
\end{aligned} \quad (157)$$

If we let A=1 in the above equations, and note that the associated eigenvalues a_n are then all equal to 1, we obtain:

(158) $\sum_n | c_n |^2 = 1$.

This suggests that the quantities $| c_n |^2$ may be interpreted as *probabilities*. Thus the quantities c_n, as given by equation (153):

(153) $c_n = \langle \psi_n | \Psi \rangle$,

will henceforth be termed *probability amplitudes*. The number $| c_n |^2$ then gives the probability that one will measure the dynamical variable A to have the value a_n, when performing a measurement of this dynamical variable for a given quantum system that is specified by the wavefunction Ψ.

§25. Unitary transformations

Consider a vector **v** which exists in two-dimensional space. One could choose to represent this vector in terms of its coordinates along a certain pair of Cartesian x and y axes, as $\mathbf{v} = v_x \hat{\mathbf{x}} + v_y \hat{\mathbf{y}}$ (notation is the same as that used in Figure 17). Now, one might choose to rotate the Cartesian coordinate system, to a new pair of axes denoted x' and y'. Our two-dimensional vector **v**—which is a geometric object which does not "care" about which set of Cartesian axes are used to describe it—would now be written as $\mathbf{v} = v'_x \hat{\mathbf{x}}' + v'_y \hat{\mathbf{y}}'$. In rotating the coordinate system, the Cartesian coordinates $(x, y) = (v_x, v_y)$ have changed to $(x', y') = (v'_x, v'_y)$. Again, the vector **v** hasn't changed—we have merely changed its "representation".

So, one can have different ways of representing the same vector. Similarly, *one can have different ways of representing the wavefunction of quantum mechanics*. Indeed, we have already met two

---

[57] Note that the first line, of equation (157), is obtained by writing the right-hand side of (63) in Dirac notation.

different ways of representing the wavefunction, in eqns (15) and (16). With reference to these equations, we may speak of Ψ and $\tilde{\Psi}$ as two different *representations* of the wavefunction, with (15) and (16) showing how one can make a *change of representation* from Ψ to $\tilde{\Psi}$, and vice versa.

In this article, we explore the use of unitary operators U, which allow one to change from one representation of the wavefunction to another.[58]

* * *

Denote by Ψ the wavefunction in a given representation. We may then act on Ψ with a unitary operator U, to give a new representation $\hat{\Psi}$ of the wavefunction:

(159) $\hat{\Psi} \equiv \mathrm{U}\Psi$.

* * *

Now that we know how to use a unitary operator U to transform *wavefunctions* from one representation to another, we need to investigate how *operators* transform under the same unitary transformation.

To this end, suppose that we have the following operator equation:

(160) $\mathrm{A}\Psi = \Phi$,

in which the operator A acts on the wave-function Ψ to give the new wavefunction Φ.

Now, (159) tells us how to use a unitary operator U to transform from one representation (of a wavefunction) to another:

(161a) $\hat{\Psi} \equiv \mathrm{U}\Psi$, $\hat{\Phi} \equiv \mathrm{U}\Phi$.

If we pre-multiply both sides of (161a) by the inverse operator U^{-1},

(161b) $\mathrm{U}^{-1}\hat{\Phi} = \Phi, \quad \mathrm{U}^{-1}\hat{\Psi} = \Psi$

and then make use of (144), we obtain:

(162) $\Phi = \mathrm{U}^{\dagger}\hat{\Phi}, \quad \Psi = \mathrm{U}^{\dagger}\hat{\Psi}$.

Substituting expressions (162) into (160), we obtain:

(163) $\mathrm{A}\mathrm{U}^{\dagger}\hat{\Psi} = \mathrm{U}^{\dagger}\hat{\Phi}$.

Pre-multiply by U, and then make use of (144) on the right-hand side of the resulting expression, to obtain:

(164) $\mathrm{U}\mathrm{A}\mathrm{U}^{\dagger}\hat{\Psi} = \hat{\Phi}$.

This may be written as:

(165a) $\hat{\mathrm{A}}\hat{\Psi} = \hat{\Phi}$,

where:

(165b) $\hat{\mathrm{A}} = \mathrm{U}\mathrm{A}\mathrm{U}^{\dagger}$.

Comparing (165a) and (160), we see that the *operator equation is unchanged in form under the unitary transformation due to* U*, if we take wavefunctions to transform according to (159) and operators transform according to (165b).*

► *Exercise #46. (a) Show that, if* A *is Hermitian, then* $\hat{\mathrm{A}}$ *is also Hermitian.(b) Show that if* [A,B]=c, *where* A *and* B *are operators and* c *is a complex number, then* $[\hat{\mathrm{A}}, \hat{\mathrm{B}}] = \mathrm{c}$. *(c) Show that the eigenvalues of* A *are the same as the eigenvalues of* $\hat{\mathrm{A}}$. *(d) Show that the quantities* $\langle \Phi \,|\, \mathrm{A} \,|\, \Psi \rangle$ *are unchanged under the unitary transformation* U, *i.e., show that*$\langle \Phi \,|\, \mathrm{A} \,|\, \Psi \rangle = \langle \hat{\Phi} \,|\, \hat{\mathrm{A}} \,|\, \hat{\Psi} \rangle$.

Evidently, infinitely many representations are possible. Unitary transformations U allow one to convert from one representation to another.

As one example, we again recall the use of the Fourier transformation (which is a unitary transformation) to convert from Ψ to $\hat{\Psi} \equiv \tilde{\Psi}$ (see equations (15) and (16)).

► *Exercise #47. Consider a unitary transformation operator* U, *which is infinitesimally close to the unit operator* I. *Write this operator as:*

(166) $\mathrm{U} = 1 + i\varepsilon\mathrm{F}$,

where ε *is a very small real parameter. (a) By using the fact that* U *is unitary, show that* F *is Hermitian. (b) Write down the transformation law for both wavefunctions and operators, under this infinitesimal unitary transformation. [For both parts of this question, work to first order in* ε*.] [Note that* F *is known as the "generator" of the unitary transformation* U.*]*

§26. Matrix mechanics

In one of the infinitely-many possible representations of quantum mechanics, wavefunctions are represented by column vectors and the associated operators are represented by matrices. This formulation, known as *matrix mechanics*, is the subject of the present section.

Recall (160), in which the operator A acts on the wavefunction Ψ to give another wave-function Φ:

---

[58] See equations (144) and (145), to refresh your memory regarding the definition of unitary operators.

(160) $A\Psi = \Phi$.

Write the following eigenfunction expansions for the wavefunctions Ψ and Φ (cf. (151)):

(167) $\Psi = \sum_m c_m \psi_m, \quad \Phi = \sum_n d_n \psi_n$.

Here, as usual, ψ_n is a given complete set of eigenfunctions,[59] and the sets of complex numbers $\{c_m\}$ and $\{d_n\}$ are the expansion coefficients for Ψ and Φ respectively.

Subsitute (167) into (160), to give:

(168) $A \sum_m c_m \psi_m = \sum_n d_n \psi_n$.

Now bring A inside the sum on the left side:

(169) $\sum_m c_m A \psi_m = \sum_n d_n \psi_n$.

Pre-multiply both sides by ψ_p^* and integrate over all spatial coordinates:

(170) $\sum_m c_m \langle \psi_p | A \psi_m \rangle = \sum_n d_n \langle \psi_p | \psi_n \rangle$.

Let:

(171) $A_{pm} \equiv \langle \psi_p | A \psi_m \rangle$

on the left side of (170). Regarding the Dirac bracket on the right side of (170), note from equation (149) that $< \psi_p | \psi_n >= \delta_{pn}$.

Hence (170) becomes:

(172) $\sum_m c_m A_{pm} = \sum_n d_n \delta_{pn}$.

In the sum on the right side of this equation, the only non-zero term is d_p. We thereby arrive at the central result for this section

(173) $\sum_m A_{pm} c_m = d_p$.

This has the form of a matrix equation, as can be seen by writing (173) in the following form:

(174) $$\begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \end{pmatrix}.$$

Remarks:

- The expansion coefficients c_m, in the eigenfunction expansion (167) for Ψ , have been placed inside a column vector. A similar thing has been done with the expansion coefficients for Φ . Thus, in the present formulation of quantum mechanics, the wavefunctions are considered to be represented by column vectors.
- The coefficients A_{pm}, as defined in (171), are known as the *matrix elements* of the operator A. In (174), this operator is denoted by a matrix, which operates on one column vector (wavefunction) to give another column vector (wavefunction).
- Different choices for the complete "basis set" $\{ \psi_n \}$ will yield different equivalent representations of matrix mechanics.
- If the basis set $\{ \psi_n \}$ is such that every element of the set is an eigenfunction of A, the matrix in (174) will be diagonal.

► *Exercise #48. Show that a Hermitian operator is represented by a Hermitian matrix, in the matrix formulation of quantum mechanics. [Note, in this context, that the Hermitian conjugate of a given matrix is obtained by first transposing the said matrix, and then replacing each element by its complex conjugate:* $A^\dagger_{mn} = A^*_{nm}$. *A matrix is Hermitian if it is equal to its Hermitian conjugate:* $A_{mn} = A^*_{nm}$ *.] [Cf. equations (138) and (139).]*

§27. The state vector

Throughout this chapter, we have developed the powerful notion that there are infinitely many different ways of formulating quantum mechanics. Notwithstanding the fact that there are an infinite plethora of representations for quantum mechanics, the underlying physical theory is independent of the representation one chooses.

Just as we may use the symbol **v** to represent a vector without "descending" to a particular description of this entity in terms of a given coordinate system, we use the ket $|\Psi\rangle$ to denote the wavefunction of a given physical system in a representation-free manner. This is called the *state vector* of the quantum system.

One may change, from one allowed repre-sentation to another, using the methods in §25.

---

[59] Note that these eigenfunctions are *not* considered to be eigenfunctions of the operator A, in general. Rather, the set $\{ \psi_n \}$ is in general considered to represent the orthogonal and normalized set of eigenfunctions of some "other" Hermitian operator B.

§28. Creation and destruction operators

We revisit the theory of the harmonic oscillator (cf. §18) from the point of view of the so-called *creation and destruction operators*. This elegant method is due to Dirac.

Remarkably, *one is able to deduce the energy levels of the harmonic oscillator without needing to invoke a particular representation of the wavefunction*. Rather we need only work with a certain commutation relation—in this context we note, from exercise 46(b), that commutation relations are unchanged under unitary transformations that convert one representation of the wavefunction into another valid representation.

The Hamiltonian $\hat{H}$ for the harmonic oscillator is given by a sum of kinetic and potential energies:[60]

$$(175)\quad \hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2} m\omega^2 \hat{q}^2,$$

where $\hat{p}$ and $\hat{q}$ are momentum and position operators[61] satisfying:

$$(176)\quad [\hat{q}, \hat{p}] = i\hbar,$$

m is the mass of the particle, and ω is the angular frequency (of the oscillator, when considered classically).

We now introduce the operators:

$$(177)\quad \hat{a} \equiv \frac{m\omega\hat{q} + i\hat{p}}{\sqrt{2\hbar m\omega}}, \quad \hat{a}^\dagger = \frac{m\omega\hat{q} - i\hat{p}}{\sqrt{2\hbar m\omega}}.$$

Equations (176) and (177) together imply that:

$$(178)\quad [\hat{a}, \hat{a}^\dagger] = 1.$$

►*Exercise #49: Derive equation (*178*).*

The Hamiltonian, expressed in terms of $\hat{a}$ and $\hat{a}^\dagger$, becomes:

$$(179)\quad \hat{H} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \tfrac{1}{2}\right).$$

►*Exercise #50: Derive equation (*179*).*

With a view to interpreting this result, let us consider the operator:

$$(180)\quad \hat{N} \equiv \hat{a}^\dagger\hat{a}.$$

The expectation value of $\hat{N}$ is non-negative, as is readily shown:

$$(181)\quad \left\langle \Psi \middle| \hat{N} \middle| \Psi \right\rangle = \left\langle \Psi \middle| \hat{a}^\dagger\hat{a} \middle| \Psi \right\rangle = \langle \hat{a}\Psi \mid \hat{a}\Psi \rangle \geq 0.$$

Let $|\beta\rangle$ be an eigenvector of $\hat{N}$ with eigenvalue β:

$$(182)\quad \hat{N}|\beta\rangle = \beta|\beta\rangle.$$

Hence:

$$(183)\quad \left\langle \beta \middle| \hat{N} \middle| \beta \right\rangle = \langle \beta|\beta|\beta \rangle = \beta\langle \beta \mid \beta \rangle \geq 0,$$

where the last step follows from (181). Since $\langle \beta \mid \beta \rangle \geq 0$, the inequality at the end of (183) implies that $\beta \geq 0$. Hence $\hat{N}$ has a lowest non-negative eigenvalue, which we denote by β_0:

$$(184)\quad \beta_0 \geq 0.$$

Put this result to one side for the moment. Now, using equations (178) and (182), it can be shown that:

$$(185a)\quad \hat{N}\hat{a}|\beta\rangle = (\beta - 1)\hat{a}|\beta\rangle,$$

$$(185b)\quad \hat{N}\hat{a}^\dagger|\beta\rangle = (\beta + 1)\hat{a}^\dagger|\beta\rangle.$$

►*Exercise #51: Derive equations (185).*

We see from equations (185) that both $\hat{a}|\beta\rangle$ and $\hat{a}^\dagger|\beta\rangle$ are eigenfunctions of $\hat{N}$, with eigenvalues of $\beta - 1$ and $\beta + 1$ respectively.

Next, let $\beta = \beta_0$ in (185a), which gives:

$$(186)\quad \hat{N}\hat{a}|\beta_0\rangle = (\beta_0 - 1)\hat{a}|\beta_0\rangle.$$

Suppose[62] that $\hat{a}|\beta_0\rangle \neq 0$; as a consequence of this assumption, (186) implies that $\hat{a}|\beta_0\rangle$ is an eigenfunction of $\hat{N}$ with eigenvalue $\beta_0 - 1$. This eigenvalue is smaller than the smallest eigenvalue β_0. Our assumption has led to a logical contradiction, therefore the initial assumption is incorrect, therefore:

$$(187)\quad \hat{a}|\beta_0\rangle = 0.$$

Furthermore, from (180) and (182) we have:

---

[60] This can be obtained directly from (105), if one makes use of (106b), together with the Schrödinger correspond-ence rule for momentum in equation (31).

[61] For the remainder of this article, a "hat" (caret) over a quantity indicates it to be an operator.

[62] Note that we have just commenced a proof by contradiction.

(188) $\hat{a}^\dagger\{\hat{a}|\beta_0\rangle\} = \beta_0|\beta_0\rangle$.

The quantity in braces is zero, on account of (187), hence the left side of (188) is zero. The ket $|\beta_0\rangle$ on the right side of (188) is non-zero, therefore:

(189) $\beta_0 = 0$.

It follows from (185b) that the eigenvalues n of $\hat{N}$ are the integers:

(190) $n = 0,1,2,3...$

We denote the eigenvectors of eigenvalue n by $|n\rangle$, and so (182) becomes:

(191) $\hat{N}|n\rangle = n|n\rangle,\ n = 0,1,2,3...$

We demand that the states $|n\rangle$ be normalised:

(192) $\langle n\,|\,n\rangle = 1$.

It follows that the states $|n\pm1\rangle$ defined by:

(193a) $\hat{a}|n\rangle = \sqrt{n}|n-1\rangle$,

(193b) $\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$,

are also normalised.

►*Exercise #52: Derive equations (*193*). Show that the states, constructed using these equations, are normalised.*

The normalised eigenfunctions $|n\rangle$ of $\hat{N}$ can be constructed from the so-called "vacuum state" $|0\rangle$ in the following manner:

(194) $|n\rangle = \frac{(\hat{a}^\dagger)^n}{\sqrt{n!}}|0\rangle,\ n = 0,1,2,3....$

►*Exercise #53: Derive equation (194).*

From (179) and (191), we see that the states $|n\rangle$ are eigenfunctions of the harmonic oscillator Hamiltonian, with energy eigenvalues:

(195) $E_n = \hbar\omega(n+\frac{1}{2}),\ n = 0,1,2,3...$

The operators $\hat{a}$ and $\hat{a}^\dagger$ are called *destruction* and *creation* operators respectively, on account of equations (193)[63]. Bearing (195) in mind, we see that the ket $|n\rangle$ *represents a state with n quanta of energy;* this state is known as the "number state" or "Fock state". We can therefore meaningfully call $\hat{N} \equiv \hat{a}^\dagger\hat{a}$ the "number operator", since its eigenvalue tells us the number n of quanta in a given "number state" $|n\rangle$ (cf. (180) and (191)). The operator $\hat{a}$ acts on $|n\rangle$ to give $\sqrt{n}|n-1\rangle$, thus destroying a quantum of energy from the harmonic oscillator (cf. (193a)); similarly, $\hat{a}^\dagger$ acts on $|n\rangle$ to give $\sqrt{n+1}|n+1\rangle$, thus creating a quantum of energy (cf. (193b)). Since the operators $\hat{a}$ and $\hat{a}^\dagger$ take us up and down the scale of energy, they are sometimes collectively referred to as "ladder operators".[64]

Note that such "creation and destruction operators" play a hugely important role in more advanced studies on quantum mechanics, quantum field theory and condensed matter physics, such as you will encounter in later years of study. It is for this reason that such material has been included here.

CHAPTER 6

ANGULAR MOMENTUM IN QUANTUM-MECHANICAL SYSTEMS

§29. Orbital angular momentum

Consider a classical point particle of mass m and momentum **p**, which is located at position **r** with respect to some origin O. As you will have seen in at least one earlier course, the *angular momentum* **L** of the particle (henceforth this will be called the *orbital angular momentum*) is given by:

(196) $\mathbf{L} = \mathbf{r}\times\mathbf{p}$.

If we introduce Cartesian components via:

(197a) $\mathbf{r} = (x, y, z)$,

(197b) $\mathbf{p} = (p_x, p_y, p_z)$,

(197c) $\mathbf{L} = (L_x, L_y, L_z)$,

then the classical expression (196) becomes:

(198a) $L_x = yp_z - zp_y$,

(198b) $L_y = zp_x - xp_z$,

---

[63] The following terms are interchangeable: "destruction operator", "absorption operator", "annihilation operator".

[64] Cf. the "ladder" in Figure 14.

(198c) $L_z = xp_y - yp_x$.

We re-iterate that all of the above expressions are classical!

* * *

To convert the above classical expressions for angular momentum into their corresponding quantum-mechanical counterparts, we use the Schrödinger correspondence rule (31), namely:

(31) $\mathbf{p} \to -i\hbar\nabla$.

Substituting into (196), we obtain the following expression for the *quantum-mechanical vector* ***operator*** *for orbital angular momentum:*

(199) $\mathbf{L} = -i\hbar\,\mathbf{r}\times\nabla$.

If one prefers to make use of Cartesian components, one can make the substitutions:

(200a) $p_x \to -i\hbar\dfrac{\partial}{\partial x}$,

(200b) $p_y \to -i\hbar\dfrac{\partial}{\partial y}$,

(200c) $p_z \to -i\hbar\dfrac{\partial}{\partial z}$,

in equations (198), to obtain the following Hermitian operators:[65]

(201a) $L_x = -i\hbar\left(y\dfrac{\partial}{\partial z} - z\dfrac{\partial}{\partial y}\right)$,

(201b) $L_y = -i\hbar\left(z\dfrac{\partial}{\partial x} - x\dfrac{\partial}{\partial z}\right)$,

(201c) $L_z = -i\hbar\left(x\dfrac{\partial}{\partial y} - y\dfrac{\partial}{\partial x}\right)$.

* * *

► *Exercise #54. Please show that the commutator, of any two Cartesian components of the orbital angular momentum, is related to the "other" Cartesian component via:*[66]

(202a) $[L_x, L_y] = i\hbar L_z$,

(202b) $[L_y, L_z] = i\hbar L_x$,

(202c) $[L_z, L_x] = i\hbar L_y$.

Hint: The fundamental commutator in equation (71a), together with the results of Exercise #19, might be useful.

Next, recall the discussions of §12.2. There, we saw the link between non-commuting operators and an associated uncertainty principle (eqn (72))—i.e., we saw that *if two operators do not commute then it is not possible to simultaneously measure their associated dynamical variables with arbitrary precision.*

Bearing the above in mind, the physical meaning of equations (202) becomes clear: *it is not possible to simultaneously measure the three Cartesian components of the orbital angular momentum*. As has so often been the case in this course, this observation is contrary to the classical doctrine.

* * *

To proceed further, let us introduce the *squared magnitude of the orbital angular momentum operator*, via:

(203) $\mathbf{L}^2 = L_x^2 + L_y^2 + L_z^2$.

► *Exercise #55. Please show that* $\mathbf{L}^2$ *commutes with each Cartesian component of the orbital angular momentum, i.e. show that:*

(204) $[\mathbf{L}^2, L_x] = [\mathbf{L}^2, L_y] = [\mathbf{L}^2, L_z] = 0$.

Thus *it* ***is*** *possible to simultaneously measure both* $\mathbf{L}^2$ *and any* ***one*** *Cartesian component of the orbital angular momentum, with arbitrary precision*. We shall be returning to this point in the next section, in our discussions on the *spherical harmonics*, which are simultaneous eigenfunctions of $\mathbf{L}^2$ and L_z.

* * *

It shall prove convenient to transform from Cartesian coordinates (x,y,z) to spherical polar coordinates (r,θ,ϕ).

---

[65] You may want to prove to yourself that these operators are indeed Hermitian, as claimed in the main text.

[66] Note that the three formula in (202) are all related to one another, by "cyclic permutation" of the variables x,y and z.

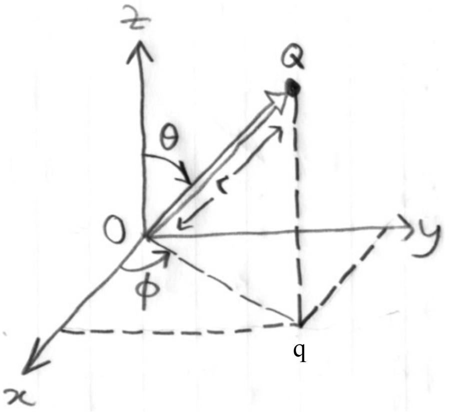


Figure 18: Spherical polar coordinates (r,θ,ϕ) of a point Q.[67]

As shown in Figure 18, spherical polar coordinates are interpreted as follows:

- r is the radial distance of a given point Q from the origin of coordinates O;
- θ is the angle that the line from O to Q makes with the z axis;
- ϕ, the "azimuthal angle", is the angle that the line from O to q makes with the positive x axis, where q=(x,y,z=0) is the projection of Q=(x,y,z) onto the xy plane.

Given the above diagram, one can immediately write down the following transformation law which relates spherical polar coordinates (r,θ,ϕ) to Cartesian coordinates (x,y,z):

(205a) $x = r\sin\theta\cos\phi$,

(205b) $y = r\sin\theta\sin\phi$,

(205c) $z = r\cos\theta$.

If we express the operators L_x, L_y, L_z and $\mathbf{L}^2$ in spherical polar coordinates, we obtain:

$$\text{(206a)}\quad L_x = -i\hbar\left(-\sin\phi\frac{\partial}{\partial\theta} - \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right),$$

$$\text{(206b)}\quad L_y = -i\hbar\left(\cos\phi\frac{\partial}{\partial\theta} - \cot\theta\sin\phi\frac{\partial}{\partial\phi}\right).$$

$$\text{(206c)}\quad L_z = -i\hbar\frac{\partial}{\partial\phi},$$

$$\text{(206d)}\quad \mathbf{L}^2 = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right].$$

► *Exercise #56. Derive equations (206).*

§30. The spherical harmonics

As we saw in the previous section, it is possible to simultaneously measure, with arbitrary precision, both $\mathbf{L}^2$ and any **one** Cartesian component of the orbital angular momentum. Thus, for example, it is possible to simultaneously measure both $\mathbf{L}^2$ and L_z with arbitrary precision. Consequently, it is possible to construct functions that are eigenfunctions of both $\mathbf{L}^2$ and L_z. One such set of functions, introduced here, is known as the *spherical harmonics* $Y_{lm}(\theta,\phi)$.

* * *

The spherical harmonics are simultaneous eigenfunctions of both $\mathbf{L}^2$ and L_z:

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

(208) $L_z Y_{lm}(\theta,\phi) = m\hbar Y_{lm}(\theta,\phi)$.

Here,

- m is an integer known as the *magnetic quantum number*, and
- *l* is a positive integer known as the *orbital angular momentum quantum number*.

We state without proof that for a given *l*, which may taken any one of the values 0, 1, 2, etc., the allowed values of m are:

(209) $m = -l, -l+1, \ldots, l-1, l$.

Note also that the spherical harmonics form an orthogonal, normalised and complete set over the unit-radius sphere coordinatized by the angles (θ,ϕ). The property of completeness ensures that one is able to expand any function $s(\theta,\phi)$ as a weighted sum of spherical harmonics:

$$\text{(210)}\quad s(\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} s_{lm} Y_{lm}(\theta,\phi).$$

Explicitly, the spherical harmonics are:[68]

---

[67] From this diagram, we note that: (i) r can take any value from zero to infinity; (ii) θ can take any value from zero to π radians; (iii) ϕ can take any value from zero to 2π radians.

[68] A derivation of this expression is given in most texts on quantum mechanics, such as B.H. Bransden and C.J. Joachain, *Introduction to Quantum Mechanics*, John Wiley and Sons, New York (1989), Chapter 6.

(211a)

$$Y_{lm}(\theta,\phi) = (-1)^{m}\sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}}P_l^{m}(\cos\theta)\exp(im\phi), \quad m \geq 0,$$

when $m \geq 0$; when m is negative, one instead uses:

(211b) $Y_{lm}(\theta,\phi) = (-1)^{m} Y^{*}_{l,|m|}(\theta,\phi), \quad m < 0$. The symbols P_l^m denote *associated Legendre functions*, which are given by:

(212) $$P_l^m(w) = \frac{(1-w^2)^{|m|/2}}{2^l l!}\left(\frac{d}{dw}\right)^{l+|m|}(w^2-1)^l.$$

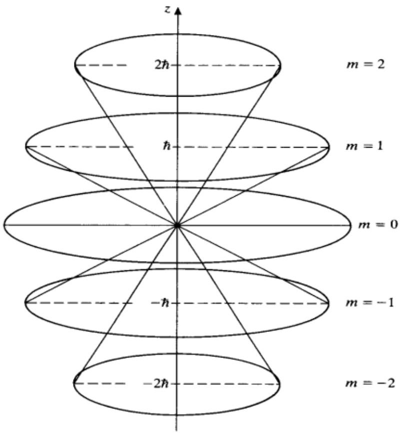
§31. Vector model for angular momentum

Here, we make some interpretative comments regarding the equations of the previous section.

* * *

From (207), we see that the bigger the positive integer l is, the larger is the squared magnitude $\mathbf{L}^2$ of the orbital angular momentum. Very loosely, increasing l is correlated with an increased magnitude of the "swirling" in the probability density associated with a given quantum state.

From (208), we see that the eigenvalue of L_z is proportional to the magnetic quantum number m. This can take both positive and negative values, indicating that probability density can "swirl" around the z axis in both the positive (counter-clockwise) and negative (clockwise) directions, implying that the associated z component of angular momentum may be both positive and negative.



The vector model of angular momentum, illustrated here for the orbital angular momentum $\mathbf{L}$ when $l = 2$. The vector $\mathbf{L}$, of length $\sqrt{l(l+1)}\hbar$, precesses about the axis of quantisation, the $(2l + 1)$ allowed projections of $\mathbf{L}$ on this axis being given by $m\hbar$, with $m = -l, -l+1, \ldots, +l$.

Figure 19: Vector model for angular momentum, for the case $l = 2$.[69]

The above admittedly vague ideas are closely related to the *vector model for angular momentum* that is given in Figure 19.

With a view to developing this model, consider a quantum state of fixed orbital angular momentum l, represented by the integers (l,m). Now, from (207) we can think of $\hbar\sqrt{l(l+1)}$ as indicating the "length" of the angular momentum vector, while from (208) we can think of $m\hbar$ as the projection of this vector onto the z axis.[70]

Now, from (209) we see that m can take any integer values from $-l$ to $+l$. For example, if l=2 then the "length" of the angular-momentum vector is $\hbar\sqrt{2(2+1)} = \hbar\sqrt{6}$, with the $2l + 1 = (2)(2) + 1 = 5$ allowed projections onto the z axis being $-2\hbar, -\hbar, 0, \hbar, 2\hbar$.

Since the x and y components of the angular momentum are not considered to be known with arbitrary precision, we can view the angular momentum vector as *precessing* about the z axis, tracing out cones such as those shown in Figure 19. Thus, while both the length of the vector and its z projection are precisely known and constant, both the x and y components of the angular momentum have a nonzero degree of uncertainty associated with them.

We close this section by stressing that the vector model of angular momentum cannot be taken too seriously!

§32. General angular momentum

In the previous section we dealt with a particular type of angular momentum, namely orbital angular momentum. Other types of angular momenta are also possible. Of these other types of angular momentum, it is the *spin* angular momentum that will be of most interest to us, at least in the present course.

The *spin angular momentum*, a term synonymous with both *spin* and *intrinsic angular momentum*, is defined as follows:

> The spin of a particle is equal to the angular momentum that particle possesses when it is at rest.[71]

---

[69] Image taken from B.H. Bransden and C.J. Joachain, *Introduction to Quantum Mechanics*, John Wiley and Sons, New York (1989), page 269.

[70] Note that we are free to orient this axis (known as the *quantization axis*) in any direction.

[71] This definition breaks down for particles, such as photons, that cannot be brought to rest! Please speak with me if you would like more information on this point.

As a very loose analogy—which will break down if pushed much further than the discussion given here—consider the Earth as it orbits the Sun. At a given instant, the total angular momentum of the earth will be the vector sum of its "**r** cross **p**" orbital angular momentum (cf. (196)), together with the "spinning like a top" angular momentum due to the fact that the Earth is rotating about an axis passing through the north and south poles.

The analogy breaks down since contemporary quantum theory considers the electron to be a point particle, so that its spin is not thought of as due to it "spinning about its axis". The term "spin", which is certainly suggestive of the idea of "spinning about an axis", is therefore considered something of a misnomer.

* * *

We digress! The central point of this section is as follows: Orbital angular momentum is not the only type of angular momentum, as the example of spin shows. It is therefore useful to *consider the quantum theory of angular momentum in a more general setting*, which can then be specialised to particular types of angular momentum (spin, orbital etc.) as necessary. The resulting "general theory of angular momentum" is the topic of the present section.

* * *

We have already encountered, on several occasions, the importance of commutation relations in the quantum theory. In particular, we saw in Exercise #46(b) that commutation relations are unchanged under unitary transformations that convert one representation of the wavefunction into another valid representation. Working with commutation relations alone therefore entails an associated high degree of generality, since one is not thereby "committed" to a particular representation of a given quantum system.

Another motivator, to develop a general theory of quantum-mechanical angular momenta that is based on commutation relations alone, arises from the fact that the spin angular momentum is **not** considered to be due to the manner in which a given wavefunction is smeared through physical three-dimensional space. Thus we do **not** expect to be able, in general, to write our general angular momentum operators in terms of derivatives with respect to spatial variables (cf. eqns (201) and (206) for orbital angular momentum, which **were** written in terms of such spatial derivatives).

* * *

Motivated by the above lengthy preamble, and inspired by the commutation relations given in equations (202), we state that *a vector operator* **J**, *with Cartesian components* $(\mathrm{J_x},\mathrm{J_y},\mathrm{J_z})$, *is an angular momentum operator if the said components are Hermitian operators that obey the following commutation relations:*

(213a) $[\mathrm{J_x},\mathrm{J_y}]=\mathrm{i}\hbar\mathrm{J_z}$,

(213b) $[\mathrm{J_y},\mathrm{J_z}]=\mathrm{i}\hbar\mathrm{J_x}$,

(213c) $[\mathrm{J_z},\mathrm{J_x}]=\mathrm{i}\hbar\mathrm{J_y}$.

Using the above commutation relations alone, one can show that:

(214) $\mathbf{J}^2=\mathrm{J_x^2}+\mathrm{J_y^2}+\mathrm{J_z^2}$ (cf. (203))

commutes individually with each of the $\mathrm{J_x},\mathrm{J_y}$ and $\mathrm{J_z}$, i.e. that (cf. (204)):

(215) $[\mathbf{J}^2,\mathrm{J_x}]=[\mathbf{J}^2,\mathrm{J_y}]=[\mathbf{J}^2,\mathrm{J_z}]=0.$

Thus we can construct simultaneous eigenfunctions of $\mathbf{J}^2$ and any **one** of its Cartesian components (say, $\mathrm{J_z}$).

Denoting the said simultaneous eigenfunctions by $|\,\mathrm{jm}\rangle$, we then have (cf. (207) and (208)):

(216a) $\mathbf{J}^2|\,\mathrm{jm}\rangle=\mathrm{j(j+1)}\hbar^2|\,\mathrm{jm}\rangle$,

(216b) $\mathrm{J_z}|\,\mathrm{jm}\rangle=\mathrm{m}\hbar|\,\mathrm{jm}\rangle$.

We state the following general results without proof, noting merely that they can all be derived using equations (213) through (216):[72]

- The permitted values of the angular momentum quantum number j are:

  (217) $\mathrm{j}=0,\tfrac{1}{2},1,\tfrac{3}{2},2,\tfrac{5}{2},\cdots$

  (This may be compared to the case of orbital angular momentum, for which we saw that j = *l* could only take on the values 0, 1, 2 etc.)

- For a specified j, the allowed values of m are:

  (218) $\mathrm{m}=-\mathrm{j},-\mathrm{j}+1,\cdots,\mathrm{j}-1,\mathrm{j}.$

  (Note that this has the same form as (209), which was obtained for the special case of orbital angular momentum.)

► *Exercise #57. Show that (215) is indeed a consequence of equations (213), as claimed above.*

* * *

---

[72] For a proof of these statements, see, for example, A. Messiah, *Quantum Mechanics*, volume II, North Holland Publishing Company, Amsterdam (1961), pp. 508–514.

For the case where j represents a spin angular momentum, note the following special cases of (217):

- spin = 0 for pions;
- spin = 1/2 for quarks;
- spin = 1/2 for protons;
- spin = 1/2 for neutrons;
- spin = 1/2 for electrons;
- spin = 1 for photons;
- spin = 3/2 for Ω^- particles.

Equation (217) suggests that the spin of *all* particles is either "integral" (0, 1, 2 etc.) or "half integral" (1/2, 3/2, 5/2, etc.).

In the so-called *Standard Model of Elementary Particle Physics*, hadrons (defined as any particle that feels the strong nuclear force) are considered to be *composite particles made up from quarks*. The quarks are considered to have spin 1/2.

From (218), the magnetic quantum number m, for a quark spin, can either be +1/2 ("spin up") or -1/2 ("spin down").

With reference to the above list, we note:

- Pions are considered to be made up of a quark ("q") plus an anti-quark ("$\overline{q}$"), one of which has spin up and the other of which has spin down. This gives a total spin of zero for the pion, as suggested by Figure 20(a).
- More generally, in the Standard Model one has a class of hadrons known as *mesons*, made from a quark-antiquark pair. If the quark and anti-quark have their spins anti-parallel then the resulting meson will have spin zero; if the quark and anti-quark have their spins parallel then the resulting meson will have spin one.
- An isolated quark, if it existed, would have spin 1/2. Free quarks have not been observed in nature, despite extensive searches, with theories having been developed for why this is the case.
- Both protons and neutrons are considered to be composite particles made from three quarks. As suggested by Figure 20(b), two of these are considered to be spin up, and one to be spin down, with respect to some axis; the total spin, of both the proton and the neutron, is then +1/2.
- More generally, in the Standard Model one has a class of hadrons known as *baryons*, which are considered to be made from three quarks.[73] If any two quarks are parallel with one another and anti-parallel to the third quark, the spin of the resulting baryon will be 1/2 (see Figure 20(b)); if, on the other hand, the three quark spins are aligned, then the spin of the resulting particle will be 3/2 (see Figure 20(c)).

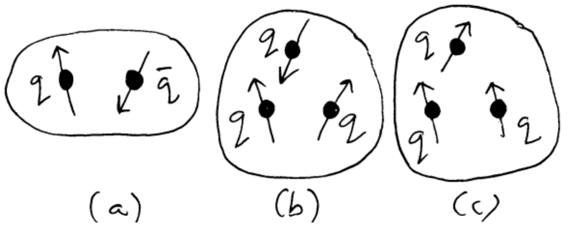


Figure 20: Quark model of hadrons

§33. General angular momentum, continued

In this section, which is rather formal, we justify two of the central results of the previous section: equations (217) and (218). We shall see that these results are obtained in a representation-free way as logical consequences of the commutation relations (213), together with the Hermitian character of J_x, J_y and J_z. The method is due to Dirac.[74]

§33.1 Raising and lowering operators for angular momentum

Let us introduce the (non-Hermitian) operators:

(219a) $J_+ \equiv J_x + iJ_y$,

(219b) $J_- \equiv J_x - iJ_y$.

One can readily show that J_+ and J_- satisfy the following commutation relations:

(220a) $[\mathbf{J}^2, J_\pm] = 0$,

(220b) $J_\pm J_\mp = \mathbf{J}^2 - J_z^2 \pm \hbar J_z$,

(220c) $[J_z, J_\pm] = \pm\hbar J_\pm$.

Equations (220) are intermediate results, each of which will be used in the discussions below.

► *Exercise #58. Derive (220a), (220b) and (220c). Hint: all the results you need are in equations (213), (214) and (219). Further, the results of Exercise #19 may be helpful.*

* * *

Recalling the eigenfunctions $|jm\rangle$ introduced in (216), we can calculate:

---

[73] One also has anti-baryons, made of three anti-quarks.

[74] Cf. §28, which is also due to Dirac, and which has some parallels with the present discussion.

$$\mathbf{J}^2\left\{J_\pm\left|jm\right\rangle\right\} = \mathbf{J}^2 J_\pm\left|jm\right\rangle \text{ ...now use (220a)}$$

$$= J_\pm\mathbf{J}^2\left|jm\right\rangle \text{...now use (216a)}$$

(221) $$= J_\pm j(j+1)\hbar^2\left|jm\right\rangle$$

$$= j(j+1)\hbar^2 J_\pm\left|jm\right\rangle$$

$$= j(j+1)\hbar^2\left\{J_\pm\left|jm\right\rangle\right\},$$

and:

$$J_z\left\{J_\pm\left|jm\right\rangle\right\} = J_z J_\pm\left|jm\right\rangle \text{ ...now use (220c)}$$

$$= (J_\pm J_z \pm \hbar J_\pm)\left|jm\right\rangle \text{...now use (216b)}$$

(222) $$= (J_\pm m\hbar \pm \hbar J_\pm)\left|jm\right\rangle$$

$$= (m\pm 1)\hbar J_\pm\left|jm\right\rangle$$

$$= (m\pm 1)\hbar\left\{J_\pm\left|jm\right\rangle\right\}.$$

In words, equation (221) states that: [75]

> $J_\pm\left|jm\right\rangle$ is an eigenfunction of the operator $\mathbf{J}^2$, with eigenvalue $j(j+1)\hbar^2$,

while equation (222) states that: [76]

> $J_\pm\left|jm\right\rangle$ is an eigenfunction of the operator J_z, with eigenvalue $(m\pm 1)\hbar$.

This immediately leads us to the following interpretation of the operator J_+:

> J_+ may be viewed as a *raising operator* which acts on the state $\left|jm\right\rangle$, thereby raising the z-projection of angular momentum from $m\hbar$ to $(m+1)\hbar$, while leaving unchanged the "squared length" $j(j+1)\hbar^2$ of the angular momentum vector;

and the operator J_-:

> J_- may be viewed as a *lowering operator* which acts on the state $\left|jm\right\rangle$, thereby lowering the z-projection of angular momentum from $m\hbar$ to $(m-1)\hbar$, while leaving unchanged the "squared length" $j(j+1)\hbar^2$ of the angular momentum vector.

§33.2 Proof of equation (218)

Regarding the just-mentioned interpretation of J_+ and J_- as raising and lowering operators, we now show that the process, of raising (or lowering) the z-projection m of the angular momentum, cannot be continued indefinitely.

To do this, we shall obtain an upper bound on the modulus of m.

Consider the expectation value of the operator $\mathbf{J}^2$. From (214), we have:

(223) $$\left\langle\mathbf{J}^2\right\rangle = \left\langle J_x^2\right\rangle + \left\langle J_y^2\right\rangle + \left\langle J_z^2\right\rangle,$$

where angular brackets denote expectation value (cf. eqn (66)). The Hermitian character—of J_x, J_y and J_z—implies that each term, on the right side of (223), is greater than or equal to zero. For example,

$$\left\langle J_x^2\right\rangle \equiv \left\langle\Psi\middle|J_x J_x\Psi\right\rangle$$

(224) $$= \left\langle J_x^\dagger\Psi\middle|J_x\Psi\right\rangle$$

$$= \left\langle J_x\Psi\middle|J_x\Psi\right\rangle$$

$$\geq 0.$$

Since each term on the *right side* of (223) is greater than or equal to zero, the *left side* of (223) must be greater than or equal to zero. As a corollary, we have:

(225) $$\left\langle\mathbf{J}^2\right\rangle = \left\langle J_x^2\right\rangle + \left\langle J_y^2\right\rangle + \left\langle J_z^2\right\rangle \geq \left\langle J_z^2\right\rangle.$$

In the above inequality, let us calculate expectation values using the eigenfunctions $\left|jm\right\rangle$ introduced in (216). Hence we are led to the following chain of logic:

$$\left\langle\mathbf{J}^2\right\rangle \geq \left\langle J_z^2\right\rangle$$

$$\left\langle jm\middle|\mathbf{J}^2\middle|jm\right\rangle \geq \left\langle jm\middle|J_z^2\middle|jm\right\rangle \text{...now use (216a, b)}$$

(226) $$\left\langle jm\middle|j(j+1)\hbar^2\middle|jm\right\rangle \geq \left\langle jm\middle|(m\hbar)^2\middle|jm\right\rangle$$

$$j(j+1)\hbar^2\left\langle jm\middle|jm\right\rangle \geq (m\hbar)^2\left\langle jm\middle|jm\right\rangle$$

$$j(j+1) \geq m^2$$

$$|m| \leq \sqrt{j(j+1)}$$

This is the promised upper bound on the modulus of m.

* * *

In light of (226), there is a certain *maximum* value of the magnetic quantum number m, that can be attained for a specified j; similarly, there is a certain *minimum* value of the magnetic quantum number m, that can be attained for a specified j.

Let $m_{\max}(j)$ denote the said maximum value, with $m_{\min}(j)$ denoting the corresponding minimum value. In this subsection, we show that:

---

[75] Cf. (216a), which says that |jm> is an eigenfunction of $\mathbf{J}^2$, with (the same) eigenvalue $j(j+1)\hbar^2$.

[76] Cf. (216b), which says that |jm> is an eigenfunction of J_z, with (different) eigenvalue $m\hbar$.

(227) $m_{\max}(j) = j,$

(228) $m_{\min}(j) = -j.$

We begin our derivation, of the above pair of equations, with the observation that the state $|\,jm_{\max}>$ will be annihilated by J_+:[77]

(229) $J_+\left|jm_{\max}\right\rangle = 0$.

Applying J_- to both sides,

(230) $J_-J_+\left|jm_{\max}\right\rangle = 0$,

so that:

$$\begin{aligned} 0 &= J_-J_+\left|jm_{\max}\right\rangle \text{...now use (220b)} \\ (231) \quad &= \left(\mathbf{J}^2 - J_z^2 - \hbar J_z\right)\left|jm_{\max}\right\rangle \text{...now use (216a,b)} \\ &= \{j(j+1) - m_{\max}^2 - m_{\max}\}\hbar^2\left|jm_{\max}\right\rangle. \end{aligned}$$

The quantity in braces must be zero, yielding:

(232) $j(j+1) - m_{\max}^2 - m_{\max} = 0$.

This quadratic in $m_{\max}$ may be solved for $m_{\max}$ using the quadratic formula, yielding the two solutions $m_{\max}(j) = j$ and $m_{\max}(j) = -j-1$. Rejecting the second such solution as unphysical, we obtain equation (227), as required.

► *Exercise #59. By beginning with the observation that* J_- *annihilates* $\left|jm_{\min}\right\rangle$*, obtain equation (228).*

* * *

Summarising the above arguments, we see that, for a specified j, the allowed values of the magnetic quantum number are exactly as was claimed in equation (218), namely:

(218) $m = -j, -j+1, \cdots, j-1, j.$

§33.3 Proof of equation (217)

The final task, for the present section, is to prove equation (217), which states that j must be either integral[78] or half-integral.[79]

To this end, consider a given state $\left|jm\right\rangle$, in which m is equal to neither its maximum permissible value of j, nor its minimum permissible value of –j. Apply the raising operator J_+ a total of p times (p is a positive integer) to this state, until m *has* attained its maximum permissible value of j. Thus:

(233) $m + p = j$.

Put this result to one side, for the moment. Start, once again, with the state $\left|jm\right\rangle$ mentioned above. Apply the lowering operator J_- a total of q times (q is a positive integer) to this state, until m has attained its *minimum* permissible value of -j. We therefore have:

(234) $m - q = -j$.

Next, let us subtract (234) from (233), giving:

(235) $p + q = 2j$.

Thus,

(236) $j = \dfrac{p+q}{2}$.

Since both p and q are positive integers, p+q will also be a positive integer, leading immediately to (217), as required.

§34. Matrix mechanics and angular momentum

Recall the discussions of §26, which dealt with the topic of *matrix mechanics*. Here, we very briefly revise the salient features of matrix mechanics, before applying these results to the study of angular momentum.[80]

* * *

In section §26, we considered the equation:

(160) $A\Psi = \Phi$,

in which an operator A acts on the wavefunction Ψ to give a new wavefunction Φ. We saw that the above equation could be written in matrix form as:

$$(174)\quad \begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \end{pmatrix}.$$

Here,

- the wavefunction Ψ is represented by the column vector containing the numbers $c_1, c_2, \ldots$, these being the coefficients in the following expansion of Ψ over a given complete set of functions $\{\psi_m\}$ (see (167)):

$$\Psi = \sum_m c_m \psi_m \,;$$

---

[77] If this were not the case, we would have a logical contradiction, for then $m_{\max}$ would be the maximum value of m, *and* it wouldn't be the maximum value.

[78] I.e., chosen from the set 0, 1, 2, 3, etc.

[79] I.e., chosen from the set 1/2, 3/2, 5/2 etc.

[80] As we shall argue in the *next* section, the methods of matrix mechanics are essential for the treatment of spin angular momentum. This serves to motivate our discussions on the matrix mechanics of angular momenta, in the present section.

- the wavefunction Φ is represented by the column vector containing the numbers $d_1, d_2, \ldots$, these being the coefficients in the following expansion of Φ (see eqn (167)):

$$\Phi = \sum_n d_n \psi_n \ ;$$

- the operator A is represented by the matrix of coefficients A_{pm}, known as *matrix elements*, which are given by:

$$(171) \quad A_{pm} \equiv \langle \psi_p | A\psi_m \rangle .$$

* * *

We now seek expressions for the matrix elements of the operators $\mathbf{J}^2$, J_z, J_+, J_-, J_x and J_y, using the states $|jm\rangle$ as a complete basis. In what follows, we assume this basis to be both orthogonal and normalised (i.e., we consider it to be "orthonormal"), thus:

$$(237) \quad \langle j'm' | jm \rangle = \delta_{jj'}\delta_{mm'} .$$

(a) Matrix elements of $\mathbf{J}^2$

The required matrix elements are:

$$(238) \quad \begin{aligned} &\langle j'm' | \mathbf{J}^2 | jm \rangle_{\text{...now use (216a)}} \\ &= \langle j'm' | j(j+1)\hbar^2 | jm \rangle \\ &= j(j+1)\hbar^2 \langle j'm' | jm \rangle_{\text{...now use (237)}} \\ &= j(j+1)\hbar^2 \delta_{jj'}\delta_{mm'} , \end{aligned}$$

corresponding to a diagonal matrix.

(b) Matrix elements of J_z

The required matrix elements are:

$$(239) \quad \begin{aligned} &\langle j'm' | J_z | jm \rangle_{\text{...now use (216b)}} \\ &= \langle j'm' | m\hbar | jm \rangle \\ &= m\hbar \langle j'm' | jm \rangle_{\text{...now use (237)}} \\ &= m\hbar \delta_{jj'}\delta_{mm'} , \end{aligned}$$

corresponding once again to a diagonal matrix.

(c) Matrix elements of J_+ *and* J_-

Prior to calculating the matrix elements of J_+ and J_-, we shall have need of a few preliminary calculations.

Now, we have already seen—in §33.1—that J_+ [J_-] may be viewed as a raising [lowering] operator which acts on the state $|jm\rangle$, thereby raising [lowering] the z-projection of angular momentum from $m\hbar$ to $(m+1)\hbar$ [$(m-1)\hbar$], while leaving unchanged the "squared length" $j(j+1)\hbar^2$ of the angular momentum vector. While the state $|jm\rangle$ may be safely taken as normalised (see (237)), the "raised" and "lowered" states will not in general be normalised. Thus we write:

$$(240) \quad J_\pm | jm \rangle = N_\pm(j,m) | j, m \pm 1 \rangle ,$$

where $N_\pm(j,m)$ is a number. Form the Hermitian conjugate of (240):[81]

$$(241) \quad \langle jm | J_\pm^\dagger = \langle j, m \pm 1 | N_\pm^*(j,m)$$

and then take the inner product of (241) with (240), to see that:

$$(242) \quad \langle jm | J_\pm^\dagger J_\pm | jm \rangle = \langle j, m \pm 1 | N_\pm^*(j,m) N_\pm(j,m) | j, m \pm 1 \rangle .$$

Now, we note from (219), together with the fact that J_x, J_y, J_z are Hermitian, that:

$$(243) \quad J_\pm^\dagger = J_\mp .$$

Hence (242) becomes:

$$(244) \quad \begin{aligned} &\langle jm | J_\mp J_\pm | jm \rangle \\ &= |N_\pm(j,m)|^2 \langle j, m \pm 1 | j, m \pm 1 \rangle \\ &= |N_\pm(j,m)|^2 . \end{aligned}$$

Taking the square root of the above expression, and adopting the convention that $N_\pm(j,m)$ be real and positive, we arrive at:

$$(245) \quad N_\pm(j,m) = \sqrt{\langle jm | J_\mp J_\pm | jm \rangle} .$$

Next, let us make use of expression (220b) for $J_\pm J_\mp$, to simplify (245) still further:

$$(246) \quad \begin{aligned} N_\pm(j,m) &= \sqrt{\langle jm | \mathbf{J}^2 - J_z^2 \mp \hbar J_z | jm \rangle} \\ &= \sqrt{\langle jm | j(j+1)\hbar^2 - (m\hbar)^2 \mp \hbar(m\hbar) | jm \rangle} \\ &= \hbar\sqrt{j(j+1) - m^2 \mp m} \\ &= \hbar\sqrt{j(j+1) - m(m \pm 1)} . \end{aligned}$$

This result completes the "preliminary calculations" mentioned a little earlier. We are now ready to calculate the desired matrix elements of J_+ and J_-.

The required matrix elements are:

---

[81] This is done by replacing all operators with their Hermitian conjugates, all complex numbers with their complex conjugates, all kets with their corresponding bras, and all bras with their corresponding kets. Also, the order is inverted, in which both kets and operators appear.

$$\langle j'm'|J_\pm|jm\rangle \text{...now use (240)}$$

$$= \langle j'm'|N_\pm(j,m)|j,m\pm1\rangle \text{...now use (246)}$$

$$(247) \quad = \hbar\sqrt{j(j+1)-m(m\pm1)}\langle j'm'|j,m\pm1\rangle$$

$$\text{...now use (237)}$$

$$= \hbar\sqrt{j(j+1)-m(m\pm1)}\ \delta_{jj'}\delta_{m\pm1,m'}.$$

We note that this does not correspond to a diagonal matrix, in contrast to the matrix elements calculated earlier in this section.

(d) Matrix elements of J_x *and* J_y

We note from (219) that:

(248) $J_x = \frac{1}{2}(J_+ + J_-)$,

(249) $J_y = \frac{1}{2i}(J_+ - J_-)$.

Thus the matrix elements of J_x and J_y are readily obtained from the matrix elements for J_+ and J_-.

Indeed, from (248), we have:

$$\langle j'm'|J_x|jm\rangle$$

$$= \tfrac{1}{2}\left(\langle j'm'|J_+|jm\rangle + \langle j'm'|J_-|jm\rangle\right)$$

$$(250) \quad \text{...now use (247)}$$

$$= \tfrac{1}{2}\hbar\sqrt{j(j+1)-m(m+1)}\ \delta_{jj'}\delta_{m+1,m'}$$

$$+ \tfrac{1}{2}\hbar\sqrt{j(j+1)-m(m-1)}\ \delta_{jj'}\delta_{m-1,m'}$$

Similarly,

$$\langle j'm'|J_y|jm\rangle$$

$$(251) \quad = \tfrac{1}{2i}\hbar\sqrt{j(j+1)-m(m+1)}\ \delta_{jj'}\delta_{m+1,m'}$$

$$- \tfrac{1}{2i}\hbar\sqrt{j(j+1)-m(m-1)}\ \delta_{jj'}\delta_{m-1,m'}$$

Again, we note that the corresponding matrices are not diagonal.

► *Exercise #60. (a) In equation (229), we claimed that* $J_+|jm_{\max}\rangle = 0$ *(i.e., that* J_+ *annihilates* $|jm_{\max}\rangle$*). Show that this is in fact a consequence of (240) and (246). (b) Similarly, in the previous exercise we claimed that* J_- *annihilates* $|jm_{\min}\rangle$*. Show that this, too, is a consequence of (240) and (246).*

§35. Spin angular momentum

We stated, in a footnote at the start of the previous section, that a matrix formulation is essential for treating spin. Indeed, this is why we outlined the matrix formulation of general angular momenta, in the previous section. Here, we show how to apply the results of the previous section, to the important special case of *spin* angular momentum.

§35.1 A change of notation

A small change of notation is in order, this change of notation being used exclusively to refer to *spin* angular momenta:

$$(252) \quad \begin{cases} \mathbf{J} \text{ becomes } \mathbf{S} \\ J_x \text{ becomes } S_x \\ J_y \text{ becomes } S_y \\ J_z \text{ becomes } S_z \\ j \text{ becomes } s \\ m \text{ becomes } m_s \\ |jm\rangle \text{ becomes } \chi_{s,m_s} \end{cases}$$

Thus, for the case of spin angular momentum, equations (213a), (213b), (213c), (214), (215), (216a), (216b), (217), (218), (219a), (219b), (237), (238), (239), (247), (250) and (251) respectively become:

(253a) $[S_x, S_y] = i\hbar S_z$,

(253b) $[S_y, S_z] = i\hbar S_x$,

(253c) $[S_z, S_x] = i\hbar S_y$,

(254) $\mathbf{S}^2 = S_x^2 + S_y^2 + S_z^2$,

(255) $[\mathbf{S}^2, S_x] = [\mathbf{S}^2, S_y] = [\mathbf{S}^2, S_z] = 0$,

(256a) $\mathbf{S}^2\chi_{s,m_s} = s(s+1)\hbar^2\chi_{s,m_s}$,

(256b) $S_z\chi_{s,m_s} = m_s\hbar\chi_{s,m_s}$,

(257) $s = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \cdots$,

(258) $m_s = -s, -s+1, \cdots, s-1, s$,

(259a) $S_+ \equiv S_x + iS_y$,

(259b) $S_- \equiv S_x - iS_y$,

(260) $\langle \chi_{s',m_s'}|\chi_{s,m_s}\rangle = \delta_{ss'}\delta_{m_s m_s'}$,

(261) $\langle \chi_{s',m_s'}|\mathbf{S}^2|\chi_{s,m_s}\rangle = s(s+1)\hbar^2\delta_{ss'}\delta_{m_s m_s'}$,

(262) $\langle \chi_{s',m_s'}|S_z|\chi_{s,m_s}\rangle = m_s\hbar\delta_{ss'}\delta_{m_s m_s'}$,

$$(263) \quad \langle \chi_{s',m_s'}|S_\pm|\chi_{s,m_s}\rangle = \hbar\sqrt{s(s+1)-m_s(m_s\pm1)}\ \delta_{ss'}\delta_{m_s\pm1,m_s'},$$

$$\begin{aligned}&\left\langle \chi_{s',m_s'} \middle| S_x \middle| \chi_{s,m_s} \right\rangle \\ (264)\qquad &= \tfrac{1}{2}\hbar\sqrt{s(s+1)-m_s(m_s+1)}\;\delta_{ss'}\delta_{m_s+1,m_s'} \\ &\quad + \tfrac{1}{2}\hbar\sqrt{s(s+1)-m_s(m_s-1)}\;\delta_{ss'}\delta_{m_s-1,m_s'}\,,\end{aligned}$$

$$\begin{aligned}&\left\langle \chi_{s',m_s'} \middle| S_y \middle| \chi_{s,m_s} \right\rangle \\ (265)\qquad &= \tfrac{1}{2i}\hbar\sqrt{s(s+1)-m_s(m_s+1)}\;\delta_{ss'}\delta_{m_s+1,m_s'} \\ &\quad - \tfrac{1}{2i}\hbar\sqrt{s(s+1)-m_s(m_s-1)}\;\delta_{ss'}\delta_{m_s-1,m_s'}\,.\end{aligned}$$

The interpretation, of each of the above equations, follows directly from the discussions given earlier, in the context of the theory of general angular momenta.

§35.2 Example: Spin one

An example may be useful in illustrating the meaning of some of the formulae written down above. Accordingly, let us study the case of *spin one* (s = 1), this being one of the spins that are allowed by (257).

From (251), when s = 1, there are three allowed values for the "magnetic" quantum number m_s, namely:

$$(266)\quad m_s = -1, 0, 1\,.$$

While we shall prove it in a few moments,[82] it appears natural to represent the three allowed spin states $\chi_{s,m_s} = \chi_{1,-1}, \chi_{1,0},$ and $\chi_{1,1}$ by the following column vectors:

$$(267)\quad \chi_{1,1} = \begin{pmatrix}1\\0\\0\end{pmatrix},\quad \chi_{1,0} = \begin{pmatrix}0\\1\\0\end{pmatrix},\quad \chi_{1,-1} = \begin{pmatrix}0\\0\\1\end{pmatrix}.$$

Operators, which act on these three-component column vectors in order to yield a new three-component column vector, are then represented via 3×3 square matrices, the elements of which are the *matrix elements* of the said operator. (a) From the s = 1 case of (261), the matrix representing $\mathbf{S}^2$ is:[83]

$$(268)\quad \mathbf{S}^2 = 2\hbar^2\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}.$$

(b) From the s = 1 case of (262), the matrix representing S_z is:

$$(269)\quad S_z = \hbar\begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}.$$

(c) From the s = 1 case of (264) and (265) respectively, we have the following matrices representing S_x and S_y:

$$(270)\quad S_x = \frac{\hbar}{\sqrt{2}}\begin{pmatrix}0&1&0\\1&0&1\\0&1&0\end{pmatrix},$$

$$(271)\quad S_y = \frac{\hbar}{\sqrt{2}}\begin{pmatrix}0&-i&0\\i&0&-i\\0&i&0\end{pmatrix}.$$

► *Exercise #61. (a) Derive equations (268) through (271), by showing that they are indeed the claimed* s = 1 *special cases of the relevant formulae from §35.1. (b) Show that the matrices, in equations (268) through (271), indeed obey equations (253) through (256).*[84]

We close this sub-section by considering the question of how to write down the wavefunction for a spin-one particle. Since the spin variable can take the three different allowed z-projections that are given in equation (266), it is evident that we need *three* complex scalar wavefunctions to describe a spin-one particle. Thus we write the wavefunction $\Psi(\mathbf{r},t)$, for the spin-one particle, as (cf. (267)):

$$\begin{aligned}&\Psi(\mathbf{r},t)\\ (272)\quad &= \Psi_1(\mathbf{r},t)\begin{pmatrix}1\\0\\0\end{pmatrix} + \Psi_0(\mathbf{r},t)\begin{pmatrix}0\\1\\0\end{pmatrix} + \Psi_{-1}(\mathbf{r},t)\begin{pmatrix}0\\0\\1\end{pmatrix}\\ &= \begin{pmatrix}\Psi_1(\mathbf{r},t)\\ \Psi_0(\mathbf{r},t)\\ \Psi_{-1}(\mathbf{r},t)\end{pmatrix}.\end{aligned}$$

Here,

- $\Psi_1(\mathbf{r},t)$ is a complex scalar wavefunction describing the $m_s = 1$ projection of the spin-one particle,

---

[82] See exercise #61, and its accompanying footnote.

[83] Note that we use the same symbol $\mathbf{S}^2$, to refer to both the operator *per se* and to its matrix representation. Similar considerations apply to the remainder of this paragraph.

[84] Note that the proof, that the column vectors in (267) obey (256b), serves to demonstrate that these were the correct vectors to represent the various spin states of the spin-one particle. (Alternatively, the column vectors in (267) may be constructed by determining the normalised eigenvectors of the matrix in (269).)

- $\Psi_0(\mathbf{r},t)$ is a complex scalar wavefunction describing the $m_s = 0$ projection of the spin-one particle, and
- $\Psi_{-1}(\mathbf{r},t)$ is a complex scalar wavefunction describing the $m_s = -1$ projection of the spin-one particle.

► *Exercise #62. How does the Born interpretation (see §6.1) generalise for the case of the spin-one wavefunction in (272)?*

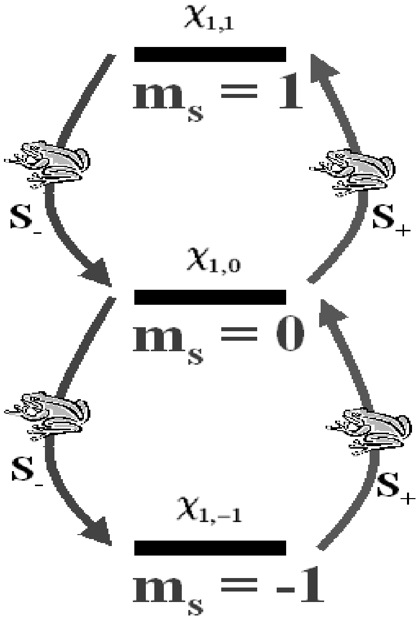


Figure 21: Diagram to aid understanding of special case of spin s = 1, discussed in this sub-section.[85]

§36. Spin half

Here we consider the very important case of spin-half angular momentum (cf. eq (257)). To motivate the importance of the study of spin-half angular momentum, we note that spin-half particles[86] include:

- electrons e^- and anti-electrons e^+;
- quarks (up u, down d, strange s, charmed c, top t, bottom b, anti-up $\bar{u}$, anti-down $\bar{d}$, anti-strange $\bar{s}$, anti-charmed $\bar{c}$, anti-top $\bar{t}$ and anti-bottom $\bar{b}$), from which all strongly interacting particles[87] ("hadrons") are believed to be composed;[88]
- muons μ^- and anti-muons μ^+;
- tauons τ^- and anti-tauons τ^+;
- neutrinos (electron neutrinos ν_e, muon neutrinos ν_μ, tauon neutrinos ν_τ, electron anti-neutrinos $\bar{\nu}_e$, muon anti-neutrinos $\bar{\nu}_\mu$ and tauon anti-neutrinos $\bar{\nu}_\tau$).

When $s = \frac{1}{2}$, equation (258) indicates there are only two allowed z-projections of the spin,[89] corresponding to:

(273) $m_s = -\frac{1}{2}, +\frac{1}{2}$.

The two allowed spin states:

(274a) $\chi_{s,m_s} = \chi_{1/2,-1/2} \equiv \chi_{\mathrm{down}}$

and:

(274b) $\chi_{s,m_s} = \chi_{1/2,1/2} \equiv \chi_{\mathrm{up}}$

are represented by the following column vectors (cf. (267)):[90]

(275) $\chi_{\mathrm{up}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \chi_{\mathrm{down}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$

The above column vectors are given the special name of *spinors*.

Operators, which act on these two-component column vectors in order to yield a new two-component column vector, are represented via 2×2 square

---

[85] Note that such "frog hopping" diagrams can be constructed for all other allowed values of s. They provide a very useful pictorial device for visualising the physics underlying the mathematics.

[86] Note that particles with half-integral spin (such as those being considered here) are called *fermions*, while those with integer spins are called *bosons*. To augment the long list of (spin-half) fermions given in the main text, we note that the following are bosons: photons, gravitons, gluons, W and Z particles, and mesons.

[87] I.e., particles which feel the strong nuclear force.

[88] Example: protons and neutrons are each believed to be comprised of three quarks (proton = up + up + down, neutron = up + down + down). More generally, all known hadrons fall into one of the following four categories: (a) *baryons*, which are considered to be made from three quarks; (b) *anti-baryons*, which are considered to be made from three anti-quarks; (c) *mesons*, which are considered to be made from a quark-antiquark pair; and (d) *penta-quarks*, which are made from four quarks and one anti-quark (and which can therefore be considered as a bound state of a baryon with a meson).

[89] For a historically-important and striking experimental demonstration of this two-fold quantisation of the z-projection of a spin-half angular momentum, you may want to look up an account of the *Stern–Gerlach experiment* using silver atoms.

[90] They are also sometimes called *two-component spinors* or *Pauli spinors*, to distinguish them from the *four-component spinors* (also called *Dirac spinors* or *bi-spinors*) that arise in studies on relativistic quantum mechanics. Please see me if you are interested in finding out more about this point.

matrices, the elements of which are the matrix elements of the said operator.

Now, make use of expressions (261) through (265), in order to explicitly write down the matrix representations of the operators $\mathbf{S}^2, S_z, S_\pm, S_x$ and S_y. This yields:

$$(276)\quad S_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$(277)\quad S_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},$$

$$(278)\quad S_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$(279)\quad \mathbf{S}^2 = \frac{3\hbar^2}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$(280)\quad S_+ = \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

$$(281)\quad S_- = \hbar\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

► *Exercise #63. (a) Derive equations (276) through (281), by showing that they are indeed the claimed* s = 1/2 *special cases of the relevant formulae from §35.1. (b) Show that the matrices, in equations (276) through (279), indeed obey equations (253) through (256). (c) Show that the matrix representation of* S_+ *"converts" the spinor* χ_{down} *into* χ_{up}*, and that the matrix representation of* S_- *"converts"* χ_{up} *into* χ_{down}. *(d) Show that the matrix representation of* S_+ *annihilates the spinor* χ_{up}*, and that the matrix representation of* S_- *annihilates the spinor* χ_{down}.

* * *

We now consider the question of how to write down the wavefunction for a spin-half particle. Since the spin variable can take the two different allowed z-projections that are given in equation (273), it is evident that we need *two* complex scalar wavefunctions to describe a spin-half particle. Thus we write the wavefunction $\Psi(\mathbf{r},t)$, for the spin-half particle, as the following *spinor field* (cf. (272)):

$$(282)\quad \begin{aligned} \Psi(\mathbf{r},t) &= \Psi_{\text{up}}(\mathbf{r},t)\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \Psi_{\text{down}}(\mathbf{r},t)\begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \Psi_{\text{up}}(\mathbf{r},t) \\ \Psi_{\text{down}}(\mathbf{r},t) \end{pmatrix}. \end{aligned}$$

Here,

- $\Psi_{\text{up}}(\mathbf{r},t)$ is a complex scalar wavefunction describing the $m_s = 1/2$ ("up") projection of the spin-half particle, and
- $\Psi_{\text{down}}(\mathbf{r},t)$ is a complex scalar wavefunction describing the $m_s = -1/2$ ("down") projection of the spin-half particle.

* * *

► *Exercise #64. To get rid of the factors of* $\hbar/2$ *that appear in equations (276) through (278),one often sees the introduction of the so-called* ***Pauli matrices****, defined via:*

$$(283)\quad S_x = \frac{\hbar}{2}\sigma_x,\quad S_y = \frac{\hbar}{2}\sigma_y,\quad S_z = \frac{\hbar}{2}\sigma_z.$$

Show that the Pauli matrices obey the following algebra:

$$(284)\quad [\sigma_x,\sigma_y] = 2i\sigma_z,$$

$$(285)\quad [\sigma_y,\sigma_z] = 2i\sigma_x,$$

$$(286)\quad [\sigma_z,\sigma_x] = 2i\sigma_y,$$

$$(287)\quad \sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1.$$

Show also that the Pauli matrices obey the following anti-commutation relations:

$$(288)\quad \{\sigma_i,\sigma_j\} = 2\delta_{ij},$$

where δ_{ij} *is the Kronecker delta, and the anti-commutator* $\{A,B\}$ *of operators* A *and* B *is defined via:*

$$(289)\quad \{A,B\} \equiv AB + BA.$$

CHAPTER 7

THE HYDROGEN ATOM

Here we treat the famous three-dimensional problem of solving the Schrödinger equation for the case of a hydrogen atom.

§37. Central potentials

Notwithstanding the discussions on spin angular momentum in the previous section, here we shall ignore the spin of the electron.[91] Thus, rather than working with the two-component spinor wavefunction given in equation (282), we can instead work with the more familiar complex scalar wavefunction, as we have done for the majority of the course.
The starting point, for our analysis, will be the time-independent Schrödinger equation, as given in equation (47):

$$(47)\quad \left[-\frac{\hbar^2}{2m_e}\nabla^2 + V(\mathbf{r})\right]\psi(\mathbf{r}) = E\psi(\mathbf{r}).$$

Here, m_e is the mass of the electron,[92] E is the energy corresponding to the spatial wavefunction $\psi(\mathbf{r})$, $\mathbf{r}$ denotes position in three-dimensional space, and $V(\mathbf{r})$ is a time-independent scalar potential.

Now, in the context of a simple model of the hydrogen atom in which one has a single electron orbiting an immobile point nucleus of charge –e, where e is the charge on an electron, the potential V will depend only on the magnitude of **r**, denoted r, if the point nucleus is taken to coincide with the origin of coordinates. Such potentials $V(r)$ are known as *central potentials*. Thus we may write V(r) instead of $V(\mathbf{r})$ in (47), yielding:

$$(290a)\quad H\psi(\mathbf{r}) = E\psi(\mathbf{r}),$$

where the Hamiltonian H is:[93]

$$(290b)\quad H = -\frac{\hbar^2}{2m_e}\nabla^2 + V(r).$$

As was the case in our discussions on orbital angular momentum, it will prove convenient for us to work in spherical polar coordinates (r,θ,ϕ) (see Fig. 18, together with equations (205a,b,c)). Accordingly, we seek to write (290b) in spherical polar coordinates. We invoke the well-known form taken by the three-dimensional Laplacian ∇^2 in spherical polar coordinates, namely:[94]

$$(291)\quad \nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right\}.$$

Bearing in mind expression (206d) for the square $\mathbf{L}^2$ of the orbital angular momentum operator, which is proportional to the term in braces above, we can write:

$$(292)\quad \nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{\mathbf{L}^2}{\hbar^2 r^2}.$$

Thus, for the case of a central time-independent potential, equation (290) becomes the following time-independent Schrödinger equation:

$$(293)\quad \left[\frac{-\hbar^2}{2m_e r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\mathbf{L}^2}{2m_e r^2} + V(r)\right]\psi(\mathbf{r}) = E\psi(\mathbf{r}).$$

► *Exercise #65. Show that the operators* $H, \mathbf{L}^2$ *and* L_z *all commute with one another (i.e. show that the commutator, of any two of these operators, is zero.).*

Since the operators H, $\mathbf{L}^2$ and L_z all commute with one another, it is possible to find solutions to (293) which are simultaneously eigenfunctions of all three of these operators. Thus we seek solutions, each of which have a particular energy E, a particular orbital angular momentum quantum number *l*, and a particular magnetic quantum number m. We may therefore write our desired spatial wavefunctions as $\psi_{Elm}(\mathbf{r})$, which obey the three eigenvalue equations:

$$(294a)\quad H\psi_{Elm}(\mathbf{r}) = E\psi_{Elm}(\mathbf{r}),$$

$$(294b)\quad \mathbf{L}^2\psi_{Elm}(\mathbf{r}) = l(l+1)\hbar^2\psi_{Elm}(\mathbf{r}),$$

$$(294c)\quad L_z\psi_{Elm}(\mathbf{r}) = m\hbar\psi_{Elm}(\mathbf{r}).$$

Bearing in mind our earlier studies on the spherical harmonics $Y_{lm}(\theta,\phi)$ (see §30), which were constructed so as to obey eigenvalue equations of the same form as (294b) and (294c), we claim that the desired spatial wavefunctions $\psi_{Elm}(\mathbf{r})$ may be written as:

---

[91] This neglect, of the spin of the electron, results in the neglect of certain small spin-sensitive correction terms to the energy levels of the hydrogen atom.

[92] We put a subscript on the electron mass, so that this quantity is not confused with the magnetic quantum number m, which shall soon appear in our equations describing the hydrogen atom.

[93] Reminder: for the purposes of this course, "Hamiltonian" can be taken as synonymous with "the total energy of a system". For more on Hamiltonians, see the paragraph of text immediately after exercise #12, together with references cited therein.

[94] Note that (291) can be proven using the same mathematical tools as those used during the proof of equation (206d), in the solutions to exercise #56. Equation (291) is also proven in almost all texts on vector analysis.

(295) $\psi_{El\mathrm{m}}(\mathbf{r}) = R_{El\mathrm{m}}(r)Y_{l\mathrm{m}}(\theta,\phi)$.

Here, $R_{El\mathrm{m}}(r)$ is known as a *radial wavefunction*, since it depends only on the radial variable r.

► *Exercise #66. Show that (295) indeed satisfies (294b) and (294c), as claimed in the main text. Hint: both* L^2 *and* L_z *act only on the angular variables* θ *and* ϕ*, with neither acting on the radial variable* r.

With a view to obtaining the differential equation obeyed by the radial wavefunction, substitute (295) into (293). This yields:

(296)
$$\left[\frac{-\hbar^2}{2m_e r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)+\frac{\mathbf{L}^2}{2m_e r^2}+V(r)\right]R_{El\mathrm{m}}(r)Y_{l\mathrm{m}}(\theta,\phi) = ER_{El\mathrm{m}}(r)Y_{l\mathrm{m}}(\theta,\phi).$$

Making use of (207), and remembering from (206d) that $\mathbf{L}^2$ does not act on r, we obtain:

(297)
$$\left[\frac{-\hbar^2}{2m_e r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)+\frac{l(l+1)\hbar^2}{2m_e r^2}+V(r)\right]R_{El\mathrm{m}}(r)Y_{l\mathrm{m}}(\theta,\phi) = ER_{El\mathrm{m}}(r)Y_{l\mathrm{m}}(\theta,\phi).$$

We may now cancel the spherical harmonic from both sides of this equation, to obtain an ordinary differential equation governing the form of the radial wavefunction:

(298)
$$\left[\frac{-\hbar^2}{2m_e r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right)+\frac{l(l+1)\hbar^2}{2m_e r^2}+V(r)\right]R_{El}(r) = ER_{El}(r).$$

Note that, since the magnetic quantum number m does not feature in the above equation, the "m" subscript has been dropped from R—the radial wavefunction has no dependence on the magnetic quantum number. Also, since R is a function of a single variable (r), the partial derivatives have been replaced by ordinary derivatives, inside the square brackets.

It will prove convenient to introduce the *scaled radial wavefunction* $u_{El}(r)$, as follows:

(299) $u_{El}(r) \equiv rR_{El}(r)$.

Since we do not want R to blow up at the origin, u must vanish at the origin:

(300) $u_{El}(r=0) = 0$.

Substituting (299) into (298), we obtain the following ordinary differential equation for the scaled radial wavefunction $u_{El}(r)$:

(301) $$\left[\frac{-\hbar^2}{2m_e}\frac{d^2}{dr^2}+V(r)+\frac{l(l+1)\hbar^2}{2m_e r^2}\right]u_{El}(r) = Eu_{El}(r).$$

► *Exercise #67. Derive (301).*

Interestingly, equation (301) has exactly the same mathematical form as the 1+1-dimensional time-independent Schrödinger equation (47), if we interpret:

(302) $$V(r)+\frac{l(l+1)\hbar^2}{2m_e r^2}$$

as an "effective potential" associated with non-zero orbital angular momentum quantum numbers *l*. The angular-momentum-induced contribution to the effective potential, for non-zero *l*, diverges at the origin and tapers off with radial distance r; loosely, this can be thought of as a "centrifugal barrier".

► *Exercise #68. Show that, in the limit as* r *tends to zero from above,* $u_{El}(r)$ *is proportional to* r^{l+1}. *Loosely interpret this result, in the context of the "centrifugal barrier" idea, that has just been sketched in the main text.*

§38. The hydrogen atom

We are now ready to apply the key results, of the previous section's discussions on the three-dimensional time-independent Schrödinger equation in the presence of a time-independent central scalar potential V(r), to the special case of the hydrogen atom. In particular, the goal of the present section is to solve the said Schrödinger equation for the *energy levels* of the hydrogen atom.

For the purposes of the present analysis, the hydrogen atom will be taken to consist of a single electron orbiting a point-like nucleus. The electron and the nucleus are both taken to have electric charges of magnitude e, that are opposite in sign. Further, we will make the approximation that the nucleus has a mass that is very large in comparison to that of the electron, allowing us to assume the nucleus to be essentially stationary.[95]

---

[95] Note that this assumption is easily lifted by introducing the so-called *reduced mass*. Please see me if you would like to discuss this point further.

Choose the origin, of spherical polar coordinates (r, θ, ϕ), to coincide with the location of the nucleus. The scalar electrostatic potential energy V(r) is then given by:

$$(303)\quad V(r) = -\frac{e^2}{4\pi\varepsilon_0 r},$$

where r is the radial distance, from the nucleus to a point in the space surrounding the nucleus, and ε_0 is the permittivity of free space.

Note that the above potential energy is *negative*, this being indicative of the fact that the electron in *bound* to the nucleus (see §6.4.). Further, recalling from first-year electrostatics that the magnitude of the force F between two point particles (in this case, the electron and proton) is equal to the magnitude of the gradient of the potential energy associated with their interaction, we have, from (303):

$$(304)\quad |F| = |F(r)| = \left|\frac{\partial}{\partial r}V(r)\right| = \frac{e^2}{4\pi\varepsilon_0 r^2}.$$

This is the familiar Coulomb inverse-square force law.

Having made some interpretive comments on the central Coulomb potential in (303), we may substitute this potential into equation (301) for the scaled radial wavefunction. Thus:

$$(305)\quad \left[\frac{-\hbar^2}{2m_e}\frac{d^2}{dr^2} - \frac{e^2}{4\pi\varepsilon_0 r} + \frac{l(l+1)\hbar^2}{2m_e r^2}\right]u_{El}(r) = Eu_{El}(r).$$

Multiplying through by $-2m_e/\hbar^2$, and then shifting all terms over to the left side, one obtains:

$$(306)\quad \left[\frac{d^2}{dr^2} + \frac{m_e e^2}{2\pi\varepsilon_0\hbar^2 r} - \frac{l(l+1)}{r^2} + \frac{2m_e E}{\hbar^2}\right]u_{El}(r) = 0.$$

Now, in the present context, E is the total energy of the electron, namely the sum of its kinetic and potential energies. If E is positive, then the electron has enough energy to “fly away to infinity”—with infinite separation between the stationary proton and the electron, the potential energy of the electron will be negligible, and thus it will have positive, purely kinetic, energy E. Thus, if E is positive, the electron is not bound to the nucleus; the hydrogen atom is ionised.

Conversely, if we take:

$$(307)\quad E < 0,$$

then the electron “does not have enough energy to fly away to infinity”. The electron is thus *bound* to the hydrogen atom. We confine our discussions to this case.

As was the case in our studies on the harmonic oscillator (see equations (206)), it is convenient to simplify the appearance of (306) by introducing suitable scaled variables. In the present case, let us make the following change of variables:

$$(308a)\quad \rho \equiv \frac{\sqrt{-8m_e E}}{\hbar}r,$$

$$(308b)\quad \lambda = \frac{e^2}{4\pi\varepsilon_0\hbar}\sqrt{\frac{-m_e}{2E}},$$

so that (306) becomes:

$$(309)\quad \left[\frac{d^2}{d\rho^2} - \frac{l(l+1)}{\rho^2} + \frac{\lambda}{\rho} - \frac{1}{4}\right]u_{El}(\rho) = 0.$$

► *Exercise #69. Derive (309).*

When the scaled distance ρ is very large (i.e., as ρ tends to infinity), we may neglect the second and third terms in the square brackets. Thus (309) has the following “asymptotic form”:

$$(310)\quad \left[\frac{d^2}{d\rho^2} - \frac{1}{4}\right]u_{El}(\rho) = 0, \quad \rho \to \infty.$$

As one can readily show by direct substitution, the solutions to this equation are:

$$(311)\quad u_{El}(\rho) = \exp(\pm\rho/2), \quad \rho \to \infty.$$

► *Exercise #70. Justify the statement that has just been made, regarding the $\rho \to \infty$ behaviour of the scaled radial wavefunction $u_{El}(\rho)$.*

Rejecting the positive exponent in (311) as unphysical—because it “blows up” as the scaled radial distance ρ tends to infinity—we conclude that $u_{El}(\rho)$ behaves like $\exp(-\rho/2)$ as $\rho \to \infty$. Thus we can write $u_{El}(\rho)$ as the product of $\exp(-\rho/2)$ and a finite-order polynomial.

Bearing in mind the result of Exercise #68, which pertains to the asymptotic behaviour of $u_{El}(\rho)$ as ρ tends to *zero* from above, we may then immediately write down:

$$(312)\quad u_{El}(\rho) = \rho^{l+1}\exp(-\rho/2)\sum_{k=0}^{\infty} g_k\rho^k\,,$$

where $c_0 \neq 0$.

Notes:

- When ρ tends to zero, the exponential in (312) tends to unity, and we are left with, to leading order in ρ (i.e. the smallest power of ρ), the asymptotic form $c_0\rho^{l+1}$. This is consistent with the result of Exercise #68.
- When ρ tends to infinity, and assuming there to be no arbitrarily large powers in the summation in (312)[96], we are left with an asymptotic form proportional to $\exp(-\rho/2)$, as required.

To proceed further, substitute (312) into (309). Thus:

$$(313)\quad \left[\frac{d^2}{d\rho^2} - \frac{l(l+1)}{\rho^2} + \frac{\lambda}{\rho} - \frac{1}{4}\right] e^{-\rho/2}\sum_{k=0}^{\infty} g_k\rho^{k+l+1} = 0\,.$$

Working through the algebra—which involves calculating the above derivatives, cancelling the exponential factor, and then combining the remaining summations into a single summation—one arrives at:

$$(314)\sum_{k=0}^{\infty}\left\{\begin{array}{l}[k(k+1)+(2l+2)(k+1)]g_{k+1}\\ +(\lambda-l-1-k)g_k\end{array}\right\}\rho^k = 0\,.$$

► *Exercise #71. Show that (314) indeed follows from (313).*

The quantity in braces must vanish, leading us to the following recursion relation:

$$(315)\quad g_{k+1} = \frac{l+1+k-\lambda}{k(k+1)+(2l+2)(k+1)}g_k\,.$$

From the above formula, we see that:

$$(316)\quad \frac{g_{k+1}}{g_k} \to \frac{1}{k} \text{ as } k \to \infty\,.$$

This is identical in form to (116). Bearing in mind (117)—as proven in Exercise #37—we conclude that the summation, in (312), behaves as $\rho^{2p}\exp(\rho^2)$ as ρ tends to infinity.

This "blowup" is physically unacceptable, and can be avoided using the now-familiar trick that the series in (312) will terminate if the numerator, of the fraction in (315), is zero for some k (call it N).[97]

Thus we have the condition:

$$(317)\quad l+1+N-\lambda = 0, \quad N = 0,1,2,\cdots.$$

Now, we know from our studies on angular momentum, that l can take on the values 0, 1, 2 etc. Thus $l+1+N$ can take on the integer values 1, 2, 3, etc. Letting $l+1+N = n$, where n is known as the *principal quantum number*, (317) becomes:

$$(318)\quad \lambda = n, \quad n = 1,2,3,\cdots.$$

With the help of (308b), this leads immediately to our final result:

$$(318)\quad E_n = -\frac{m_e e^4}{32\pi^2\varepsilon_0^2\hbar^2 n^2}, \quad n = 1,2,3,\cdots,$$

Note that we have put an "n" subscript, on the energy E, to indicate that this energy depends on the quantum number n.

Remarkably, the above energy levels are independent of both the orbital angular momentum quantum number *l*, and the magnetic quantum number m.

We have seen equation (318) before, for it is exactly the same result for the hydrogen-atom energy levels as that which is obtained using the Bohr model of the hydrogen atom—see Exercise #3.

If the Bohr and Schrödinger theory give the same results, why bother with the extra complexity of the Schrödinger theory? Two answers:

- The Schrödinger theory is very readily extended to incorporate extra effects that are needed for a very precise description of the hydrogen spectrum. Bohr theory is much more difficult to extend, so as to incorporate these effects.
- The Schrödinger equation and its generalisations are applicable to a much broader range of phenomenon than just the hydrogen atom! Examples include:
    - condensed matter physics;
    - electron diffraction;
    - cold electron emission;
    - simple model for alpha decay;
    - laser cooling;
    - quantum computing;

---

[96] A piece of sloppiness: One might object that we have formally included infinitely many terms in the summation, which may spoil the quoted asymptotic behaviour. While this is true, we point out that the said series will be truncated to a finite-order polynomial, for reasons that will become apparent later in this section.

[97] Recall that an analogous argument was used during our studies on the harmonic oscillator.

- Aharonov-Bohm effect;
- energy spectra of molecules;
- simple nuclear systems;
- electron-photon interactions;
- Zeeman effect;
- scattering theory.

Non-linear generalizations of the Schrödinger theory allow one to describe such systems as:

- superfluids;
- superconductors;
- Bose-Einstein condensates.

CHAPTER 8

BEYOND THIS COURSE

§39. Introduction

In many respects, this course is an "enabler" which develops a good deal of formalism and associated key concepts that will be of repeated use to you in much of your remaining physics studies. In developing this course, I have kept this aim in mind.

The course will have succeeded, in my opinion, if the following two core goals have been realized:

> (a) As a result of studying the previous chapters, you are well prepared when, in a future course that builds upon this one, you meet "familiar friends" learned in the present course, and feel comfortable and adventurous and confident in applying these tools to new contexts;
>
> (b) As a result of studying the sketch in the present chapter, and reviewing earlier chapters in the light of this chapter's pointers, you have some feel for the over-arching "big picture" of the myriad ways in which the core tools and concepts of this course pervade the broader framework of theoretical physics.

§40. An ending, and a beginning

Much of the material of §§1 and 2 will have covered familiar ground, regarding some of the key ideas of quantum mechanics as covered in the earlier years of university, and also in the later years of high school. One exception to this mood of recapitulation might have been the "lemon tree" argument presented in §1, which took an alternative perspective to the usual "quantum mechanics is weird" argument, by posing the idea that there is no *a priori* reason why the microscopic world should evolve in accord with an "everyday common sense" that is developed over length/time/mass scales that are typically many orders of magnitude shorter/faster/lighter than the familiar phenomena that may be directly registered by our senses. An immediate corollary of this "dangers of extrapolation" observation is that many of the most exciting developments in physics are counterintuitive to a degree, insofar as they challenge our "everyday common sense" by using evidence-based scientific enquiry and mathematical reasoning to transcend and surprise the prejudices of everyday common sense. Familiar general examples include Quantum Mechanics' violation of the common sense of "everyday physics" as quantified in Classical Mechanics, the violation of everyday common-sense notions of time by Special Relativity, and the violation of everyday common-sense notions of Euclidean geometry by General Relativity. Familiar specific examples include superfluidity, superconductivity, the incorporation of both Special and General Relativity into the Global Positioning system, highly directional synchrotron radiation from relativistic electrons orbiting a circular ring, black holes, quantized spectral lines, and quantum computation.

Chapter 2 introduced the wavefunction, a construct which you will have been familiar with from previous courses. We began by considering the plane wave in §4, subsequently superposing a continuum of such plane waves (in §5, via a Fourier integral) so as to synthesize an arbitrary wavefunction. Note that this process of synthesis—whereby a given complicated physical construct (in this case, a quantum-mechanical wavefunction) is expressed as a weighted superposition of an infinite succession of "basis functions" (in this case, the set of all plane waves)—is a common theme throughout physics. The Fourier integral is used in many settings, from General Relativity to Classical Optics, from Classical Electromagnetism to Ultracold Condensed-Matter Physics, to synthesize certain physical systems as superpositions of plane waves. If the set of basis functions is not restricted to plane waves, then the utility of "weighted superpositions" (cf. equation (15)) reaches still further throughout the fabric of theoretical physics. I invite you to identify several such superpositions, which do not utilize plane waves, from your current experience of physics (cf. the concept of an "eigenfunction expansion" in §23).

The historically-challenging task of interpreting the wavefunction, which for some physicists remains challenging to this day, was discussed in §6. In addition to providing the Born interpretation of the wavefunction (in §6.1), we encountered the correspondence principle (§6.3) and the concept of bound and unbound states (§6.4). (a) Note that a modified form of the correspondence principle is very often used, in a suitably broad sense of the term, by theoretical physicists whenever their physical reasoning leads them into "uncharted territory" where their intuition begins to fail. Namely, the demand that "the newer, higher theory should reduce to the older, less general theory, in the domain of validity of the latter" often guides physicists to seek how their "new" theory reduces to an older and more familiar theory as a limiting special case. Such a process, which is reminiscent of the Correspondence Principle's demand that "quantum mechanics must reduce to classical mechanics in the domain of validity of the latter", is a powerful analytical and explicatory tool in the physicist's arsenal of tools for scientific discovery. As two examples of many that could be chosen to exemplify this line of reasoning, we mention the honours-level demonstrations that General Relativity contains Special Relativity as a limiting case, and that the relativistic quantum-mechanical equation for the electron (known as the Dirac equation) reduces to the non-relativistic quantum-mechanical equation for the electron (namely the Schrödinger equation, which was obtained using a chain of reasoning proceeding from the non-relativistic Newton-type expression in equation 32 of these notes). (b) Also, regarding unbound states, about which very little was said in the present course, note that much of quantum scattering theory (in the contexts, for example, of the scattering of nuclei from one another in nuclear physics, the scattering of high-energy particles from one another in particle physics, or in the scattering of fast electrons from crystals in electron diffraction physics) is based on superposing unbound-state solutions to the relevant quantum equations of motion (such as the Schrödinger equation, the Dirac equation, and the Klein–Gordon equation).

The Schrödinger equation was discussed at length in Chapter 3. Many of the key themes developed there arise again in more sophisticated settings in later courses. For example, the discussion of the time-dependent Schrödinger equation in §6 is paralleled by the corresponding time-dependent quantum-mechanics equations discussed in honours physics (i.e., the previously mentioned Dirac and Klein–Gordon equations), with the corresponding time-independent equations, boundary conditions and considerations regarding probability density all paralleling to some degree the discussions in Chapter 3. The role of quantum-mechanical operators, also discussed in Chapter 3, plays a very significant role in more advanced studies on Quantum Mechanics, including third-year-level Advanced Quantum Mechanics, and honours-level Quantum Mechanics, Advanced Quantum Mechanics, and Quantum Field theory.

The formalism developed in Chapter 5 will also be of great use in many of your further studies in theoretical physics, in third year and beyond. The concepts of state vectors (§§21 and 27), Hermitian operators (§22), eigenfunction expansions (§23), probability amplitudes (§24), unitarity (§25), and creation and destruction operators (§28)—all of these are "bread and butter" tools of theoretical physics which arise again and again in a variety of settings, from the quantum optics of ultracold Bose-condensed gases through to the scattering of high-energy particles, from excitations in quantum superfluids through to changes of representation in the context of quantum scattering theory, from the study of weak gravitational waves to the modal expansion of a light field in a laser cavity. As this random series of examples is intended to suggest, *the tools developed in this course will be of use to you throughout your career in physics; your understanding of and familiarity with these tools will continue to deepen with time as you see them applied again and again in a variety of contexts.* As a subtext to the above, let me mention that I do feel a little guilty that I am not able to develop a greater number of particular applications of the formalism discussed in the present course, due to time constraints.

It may have come as something of a surprise that a whole chapter (Chapter 6) was devoted to angular momentum. When you have a moment, I challenge you to go to the Science library, and pick up random high-level textbooks on one or more of the following topics: (a) quantum field theory; (b) relativistic quantum mechanics; (c) high-energy particle physics; (d) quantum statistical mechanics; (e) quantum optics; (f) quantum collision theory; (g) nuclear physics. With a few of these books on the desk in front of you, look up "angular momentum" and "spin" in each index. The diversity and number of "hits" will give you a better vision of the vast importance of angular momentum and spin in contemporary physics.

At the end of the previous chapter, we mentioned that "The Schrödinger equation and its generalizations are applicable to a much broader range of phenomenon than just the hydrogen atom." This serves to highlight two salient points: (a) One could spend a lifetime exploring solutions to the Schrödinger equation in a variety of physics fields, and indeed a very large amount of original contemporary research is devoted to finding solutions to particular special cases of the Schrödinger equation in settings as diverse as the scattering of fast electrons from crystals, the study of electrons in quantum dots, and the dynamics of electrons in conducting metal lattices. (b) Notwithstanding this rich field of study in solving the Schrödinger equation, whether for one-electron or many-electron systems (we focused on one-electron systems in the present course, but the "many body theory" of many-electron systems is also hugely important (e.g. in condensed matter physics)), some systems require *modified* forms of the Schrödinger equation for their accurate description. Examples of such "modified Schrödinger equations" include the "Gross–Pitaevskii equation" used to model Bose–Einstein condensates, the "complex Ginzburg–Landau equation" used to model high-temperature superconducting films, the differential equations used to model certain superfluids, and certain differential equations used in modeling solitons in the context of the nonlinear optics of highly energetic optical beams.