Quantum Theory

Márk Mezei, Michaelmas Term 2025

About these notes

These are lecture notes for the Part A Quantum Theory course, which is a second year course in the mathematics syllabus at the University of Oxford. Sections 0-9 are largely identical to those of *Prof. James Sparks* from 2021, while section 10 on quantum key distribution is new. Starred sections/paragraphs are not examinable, either because the material is slightly off-syllabus, or because it is more difficult. There are four problem sheets. Please send any questions/corrections/comments to `mezei@maths.ox.ac.uk`.

Contents

Preamble

Classical physics – Newton's laws of mechanics and the theory of electromagnetism developed in the 19th century – fails at the atomic scale. From the beginning of the 20th century, mounting experimental evidence pointed towards the existence of a radically different theory of physics that governs the properties of atoms and their constituent particles. The deterministic and continuous nature of classical physics was replaced by a new *quantum theory*, with probability and discreteness at its heart. The theory of quantum mechanics developed in the 20th century not only successfully describes the structure of atoms and molecules, but also particle physics (such as in particle accelerators like the LHC), chemistry (such as chemical bonding), the structure of solids, superconductors, *etc.* Quantum theory also underpinned many important technological advances in the 20th century, such as the laser, the microchip (hence computers and mobile phones), and the electron microscope. Future applications include quantum cryptography and the quantum computer.

We begin the course with an overview of some key physical ideas and formulae. These developed from experiments that demonstrate the failure of classical physics at the atomic scale. An important concept here is *wave-particle duality.* As we shall see, certain aspects of an electron's behaviour can be understood by treating it as a classical point particle, while other aspects can be understood by modelling the electron as a wave. The upshot is that neither the classical particle model nor the wave model completely describe the electron, so we need something new. These ideas, together with some intuition from classical physics, will lead us to the *Schrödinger equation*, that governs the quantum *matter-waves* of particles such as the electron. We discuss general properties of the Schrödinger equation, and interpret the wave function in terms of a probability distribution. We also study in detail some of the simplest, and most important, solutions: a particle confined to a box, the harmonic oscillator (which describes small oscillations of any quantum system), and the hydrogen atom. In particular, we will derive the observed emission/absorption spectrum of the hydrogen atom, a computation that was in Schrödinger's original 1926 paper.

The development of quantum mechanics has also led to beautiful new mathematics, some of which we will see in the second half of the course. Linear algebra plays a central role, and familiar dynamical quantities, such as position, momentum, angular momentum and energy, will take new mathematical forms. This will also lead to new *physical interpretations* of these quantities, and it's here where quantum theory is famously (or infamously) counter-intuitive. The standard interpretation of quantum mechanics raises some deep philosophical questions about the role of the observer, measurement, and even the nature of reality. Most physicists take a pragmatic approach: the predictions of quantum theory, even those that might at first seem paradoxical, have been confirmed again and again in experiments over the last century. It works. But we shall return to briefly discuss some of these philosophical questions at the very end of the course.

> *What we observe is not nature itself, but nature exposed to our method of questioning.*
>
> Werner Heisenberg

Reading

- B. Zwiebach, *Mastering Quantum Mechanics: Essentials, Theory, and Applications* (MIT Press, 2022). Chapter 12.4 (great modern pedagogical reference)
- R. P. Feynman, R. B. Leighton and M. Sands, *The Feynman Lectures on Physics, Volume 3* (Addison-Wesley, 1998). Chapters 1, 2 (for physical background).
- K. C. Hannabuss, *An Introduction to Quantum Theory* (Oxford University Press, 1997). Chapters 1–4, 6–8.
- D. V. Schroeder, *Notes on Quantum Mechanics* `https://physics.weber.edu/schroeder/quantum/` (if you want to deepen your understanding with *Mathematica*)
- B. H. Bransden and C. J. Joachain, *Quantum Mechanics* (2nd edition, Pearson Education Limited, 2000). Chapters 1–7.
- A. I. M. Rae, *Quantum Mechanics* (4th edition, Taylor & Francis Ltd, 2002). Chapters 1–6.
- P. C. W. Davies and D. S. Betts, *Quantum Mechanics (Physics and Its Applications)* (2nd edition, Taylor & Francis Ltd, 1994). Chapters 1, 2, 4, 6–8.

0 Classical particles and waves

Before starting the course proper, we begin with a brief review of classical point particles and waves. This is to remind you of some concepts and formulae learned at school or in Prelims, and to set notation used later in the text.

0.1 Point particles

A point particle is an idealized object that, at any given instant of time, is located at a point in space. The position of the particle is governed by Newton's second law: if the particle has constant mass m and is acted on by a force $\mathbf{F}$, then its position vector $\mathbf{x}(t)$ obeys

$$m\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2} = \mathbf{F} . \tag{0.1}$$

In this course we shall only consider *conservative forces* where $\mathbf{F} = -\nabla V$, for some function $V = V(\mathbf{x})$ called the *potential.* The *total energy* of the particle is then the sum of the *kinetic energy* $\frac{1}{2}m|\dot{\mathbf{x}}|^2$ and *potential energy* $V = V(\mathbf{x}(t))$:

$$E = \frac{1}{2}m|\dot{\mathbf{x}}|^2 + V . \tag{0.2}$$

Here we have denoted $\dot{\mathbf{x}} = \mathrm{d}\mathbf{x}/\mathrm{d}t$. The total energy is *conserved*; that is, it is constant when evaluated on a solution to Newton's second law:

$$\frac{\mathrm{d}E}{\mathrm{d}t} = m\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}\cdot\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2} + \nabla V\cdot\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \dot{\mathbf{x}}\cdot\left(m\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2} - \mathbf{F}\right) = 0 , \tag{0.3}$$

where in the last step we used (0.1). Notice that we may also write the kinetic energy as

$$E_{\mathrm{kinetic}} = \frac{|\mathbf{p}|^2}{2m} , \tag{0.4}$$

where $\mathbf{p} = m\dot{\mathbf{x}}$ is the *momentum* of the particle.

A particle that is subjected to no forces is called a *free particle.*

* The above formulae are true (to a very good approximation) for a *non-relativistic* point particle, which means its speed $u = |\dot{\mathbf{x}}|$ is much less than the speed of light, $u \ll c =$ speed of light. For example, a *massless* particle, with $m = 0$, necessarily moves at speed c and has energy $E = c|\mathbf{p}|$, where $\mathbf{p}$ is its momentum. This will be treated in the Special Relativity option course.

0.2 Waves

Recall the classical *wave equation*

$$\frac{1}{v^2}\frac{\partial^2\phi}{\partial t^2} = \nabla^2\phi , \tag{0.5}$$

where v is the constant speed of the wave. This linear equation governs, for example, the propagation of sound or light. As basic solution we have the *complex plane wave*

$$\phi(\mathbf{x},t) = A\exp\left[\mathrm{i}(\mathbf{k}\cdot\mathbf{x} - \omega t)\right] , \tag{0.6}$$

where we have the constant *wave vector* $\mathbf{k}$, *angular frequency* ω, and (possibly complex) *amplitude* A. Substituting (0.6) into the wave equation (0.5) gives $-\omega^2/v^2 = -|\mathbf{k}|^2$, or equivalently

$$v = \frac{\omega}{|\mathbf{k}|} \ . \tag{0.7}$$

Both the real and imaginary parts of ϕ separately satisfy the wave equation, giving real solutions that are linear combinations of sines and cosines of $(\mathbf{k}\cdot\mathbf{x} - \omega t)$. In fact it is a result of Fourier analysis that *every* solution to the wave equation is a linear combination (in general involving an integral) of these plane waves.

The *wave frequency* is $\nu = \omega/2\pi$, while the *wavelength* is $\lambda = 2\pi/|\mathbf{k}|$, so that (0.7) is equivalent to saying that the wave speed is $v = \nu\lambda =$ frequency × wavelength.

1 Physical background and wave-particle duality

1.1 The photoelectric effect: waves as particles

In the mid 19th century Maxwell successfully described light as a *wave* propagating in the *electromagnetic field*. We shall not need to know anything about electromagnetic theory in this course. Nevertheless, we note in passing that, in Maxwell's theory, light propagating through a vacuum is described by two vector fields $\mathbf{E}(\mathbf{x}, t)$, $\mathbf{B}(\mathbf{x}, t)$, called the *electric* and *magnetic* fields, respectively, each Cartesian component of which satisfies the wave equation (0.5), with $v = c$ being the speed of light in vacuum. This theory of electromagnetism unified the theories of electricity, magnetism and radiation, and explained wave-like properties of light such as reflection, polarization and diffraction. However, by the beginning of the 20th century it was becoming clear that Maxwell's theory could not explain experiments at the *atomic* scale.

A simple experiment that clearly demonstrates this is the *photoelectric effect*, shown in Figure 1.

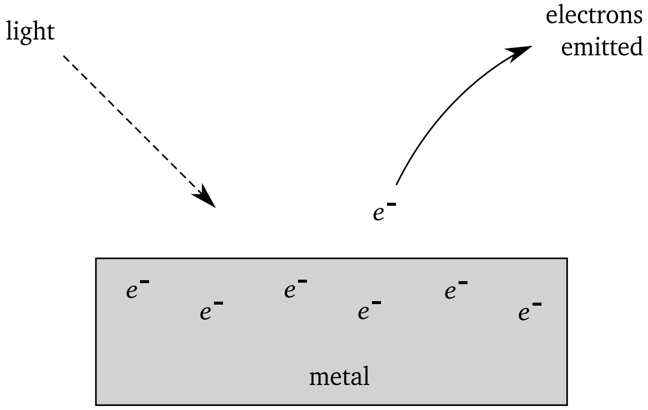

Figure 1: To observe the photoelectric effect, light is shone on a metal plate in vacuum. Electrons e^- are emitted from the surface of the metal, and their kinetic energy E_{kinetic} is measured.

Light of angular frequency ω is shone on a metal plate in vacuum. Electrons e^-, which are only weakly bound to the metal, are emitted from the surface. One measures their kinetic energy and discovers the formula

$$E_{\text{kinetic}} = -E_0 + \hbar\omega \ . \tag{1.1}$$

We also make the following observations:

(i) $E_0 > 0$ is a constant energy which depends only on the particular metal used.

(ii) The constant of proportionality $\hbar$ (usually read as "h bar") is a constant of Nature that is fundamental to quantum theory. From (1.1) we see that it has dimensions $[\hbar]$ = energy × time, or equivalently $[\hbar] = \mathrm{M}\,(\mathrm{L}\,\mathrm{T}^{-1})^2 \times \mathrm{T} = \mathrm{M}\,\mathrm{L}\,\mathrm{T}^{-1} \times \mathrm{L}$ = dimensions of *angular momentum*, where M, L and T denote dimensions of mass, length and time, respectively. Numerically, $\hbar \simeq 1.05 \times 10^{-34}\,\mathrm{J\,s}$. The combination $2\pi\hbar \equiv h$ is called *Planck's constant*, while $\hbar$ is sometimes referred to as the *reduced Planck's constant.*[1]

(iii) If the angular frequency $\omega < E_0/\hbar$, *no* e^- *are emitted.*

(iv) The formula (1.1) is independent of the *intensity* (brightness) of the light, but as the latter is increased the *number* of e^- emitted increases.

The classical theory of Maxwell does not explain these observations. Instead Einstein made the following remarkable hypothesis:

> *Light of angular frequency* ω *exists in small packets, or "quanta", of energy* $E = \hbar\omega$, a relation we call the *Einstein-Planck relation.*

These packets of light are known as *photons*, and are massless particles ($m = 0$) that travel at the speed of light c (*cf.* the starred remark at the end of section 0.1). In Einstein's interpretation of the photoelectric effect, each electron e^- absorbs one photon of energy $\hbar\omega$ (*very* occasionally more than one). Part of this photon energy goes into overcoming the "binding energy" E_0 of the electron to the metal; the remainder is then converted into the observed kinetic energy of e^- when it is emitted, leading to (1.1). In particular, when $\omega < E_0/\hbar$ the photon energy $\hbar\omega$ is insufficient to overcome the binding energy E_0, explaining why no e^- are emitted in (iii) above. Increasing the intensity of the light simply increases the *number* of photons, which explains (iv).

1.2 Emission/absorption spectra of atoms: discreteness

Atoms emit and absorb light at very particular frequencies. The simplest atom is the hydrogen atom, which we shall study in more detail at the end of this course. For hydrogen these frequencies were discovered experimentally in the 19th century, see Figure 2. These frequencies were discovered

---

[1] * Planck had introduced the constant $h = 2\pi\hbar$ in earlier work on radiation emitted from certain idealized hot objects, called *black bodies.*

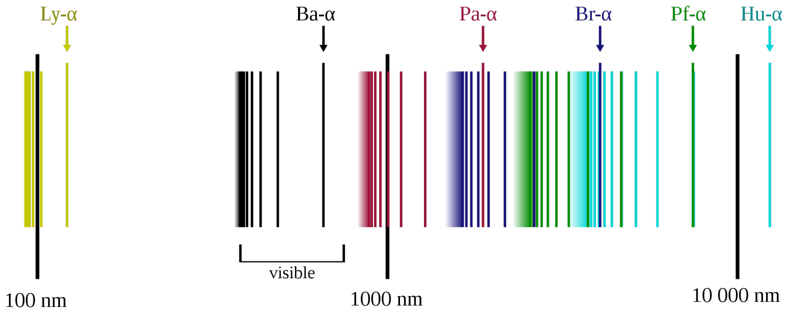

Figure 2: The spectral series of hydrogen, on a logarithmic scale (© OrangeDog on Wikipedia).

to perfectly fit the empirical formula

$$\omega_{n_1,n_2} = 2\pi R_0 \, c \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right) . \tag{1.2}$$

Here $n_1 < n_2$ are positive *integers*, and $R_0 \simeq 1.10 \times 10^7 \, \mathrm{m}^{-1}$ is *Rydberg's constant*, named after the discoverer of the empirical formula (1.2).

From Einstein's description of light in terms of photons, this implies that a hydrogen atom emits and absorbs photons of particular *energies* $\hbar\omega_{n_1,n_2}$. By conservation of energy, the energy of the hydrogen atom itself must then be changing by these amounts when a photon is emitted/absorbed. This strongly suggests that the energies of the hydrogen atom are given by

$$E_n = -\frac{2\pi R_0 \hbar c}{n^2} , \tag{1.3}$$

where n is an integer, so that for $n_1 < n_2$ the energy of the atom can change from E_{n_2} to E_{n_1} by emitting a photon of frequency ω_{n_1,n_2}. That the hydrogen atom energies are indeed *quantized* in this way – that is, taking particular discrete values, rather than being continuous – has been confirmed in many other experiments. At the end of this course we will derive (1.3) theoretically.

1.3 The double slit experiment: particles as waves

The photoelectric effect implies that light, described classically by waves satisfying the wave equation with speed $v = c$, is sometimes better described as a beam of particles, namely photons. Similarly, particles, such as the electron e^-, can display wave-like characteristics. Perhaps the best experiment that demonstrates this is the famous *double slit experiment*, performed with electrons. Thinking deeply about this setup has led Feynman to give an alternative formulation of quantum mechanics, the path integral, that is the framework of choice in most modern work on quantum field theory and that has inspired important developments in algebraic geometry and topology. You can learn about path integrals, for example, in the Part C7.1 course Theoretical Physics.

This is shown in Figures 3 and 4. A beam of electrons is fired at a double slit configuration, with a detector screen on the other side. An electron hitting the screen appears as a bright spot, and over time one can plot this as a distribution. The latter exhibits a familiar diffraction pattern, similar to that seen in the corresponding experiment with the beam of electrons replaced by a beam of light. Such diffraction patterns are explained by the interference of waves: two waves that travel through each of the slits and arrive at the same point on the detector screen have travelled different distances. These waves then either constructively or destructively interfere with each other, depending on whether the difference in these distances is an even or odd number of half-wavelengths, respectively. This is perhaps familiar to those who have taken A-level physics.

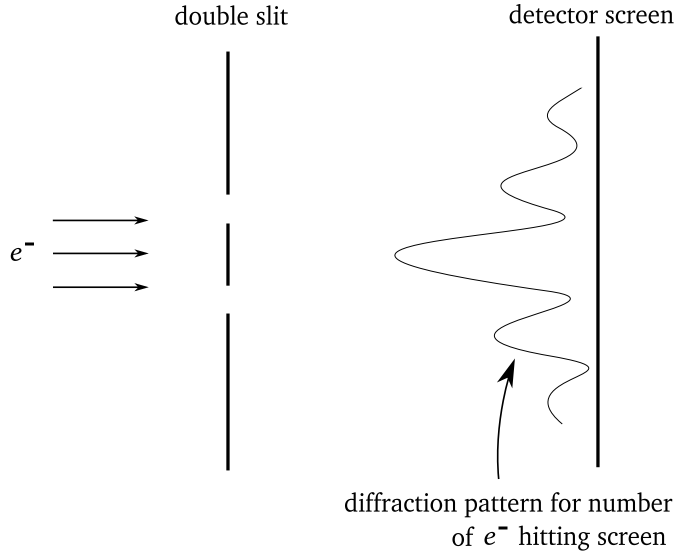

Figure 3: The double slit experiment, performed with a beam of electrons.

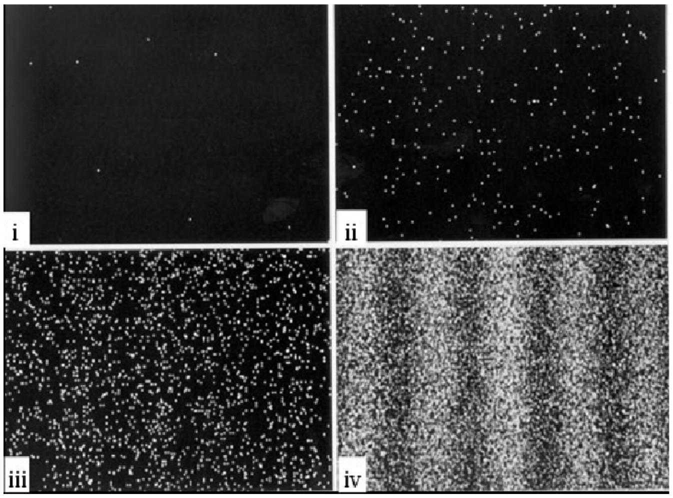

Figure 4: Actual electron hits in an experiment by a group at Hitachi (© Hitachi, Ltd): (i) 8 e^-, (ii) 270 e^-, (iii) 2,000 e^-, (iv) 160,000 e^-.

A remarkable point here is that the diffraction pattern is still observed even when only a *single* electron is passing through the slits at a time. In fact this is the case in the Hitachi experimental results shown in Figure 4 (the time lapse up to picture (iv) is 20 minutes). This implies that the electrons are not interfering with each other to cause the diffraction pattern, but rather a *single* electron is behaving like a wave. Or, more precisely, the detected distribution of electron particles is characteristic of a wave passing between the slits. Notice that we may also interpret this distribution as a *probability distribution* for where any single electron will hit the screen. In this viewpoint, it is the *probability* that displays wave-like characteristics, while the electrons themselves are always detected on the screen as localized particles. These remarks are absolutely central to wave-particle duality: the electron has both particle-like and wave-like characteristics.

The double slit experiment (and variants of it) is extremely interesting and subtle, and we refer the interested reader to the references (especially the Feynman lectures) for further discussion of its role in understanding wave-particle duality.

1.4 De Broglie's matter-waves

The experiment we have just described suggests that particles, such as electrons, are also associated with waves. De Broglie made this more precise with the claim:

> A free particle of energy E and momentum $\mathbf{p}$ is associated with a *wave* of angular frequency ω and wave vector $\mathbf{k}$ via
>
> $$\begin{aligned} E &= \hbar\omega \ , \\ \mathbf{p} &= \hbar\mathbf{k} \ . \qquad \text{(de Broglie relations)} \end{aligned} \tag{1.4}$$

Since the wavelength is $\lambda = 2\pi/|\mathbf{k}|$, we may also write the latter relation as $\lambda = 2\pi\hbar/|\mathbf{p}| = h/|\mathbf{p}|$. De Broglie's insight was that these relations should apply to *all* particles, not just massless photons. In this context, $E = \hbar\omega$ is usually referred to as a de Broglie relation, rather than the Einstein-Planck relation.

* We note that for a photon, the second relation in (1.4) is implied by the first relation $E = \hbar\omega$. This follows from the starred comment at the end of section 0.1: for a photon $E = c|\mathbf{p}|$, so that $E = \hbar\omega$ together with $c = \omega/|\mathbf{k}|$ implies that $|\mathbf{p}| = \hbar|\mathbf{k}|$. In the Special Relativity option you can learn that $(E, \mathbf{p})$ and $(\omega, \mathbf{k})$ are both 4-vectors, and indeed this was part of de Broglie's reasoning.

2 Wave mechanics

2.1 The Schrödinger equation

De Broglie had hypothesized that particles, such as the electron e^-, are associated with waves. Schrödinger set out to discover the equation that governs these *matter-waves*. He began by considering the plane wave, reviewed in section 0.2,

$$\Psi(\mathbf{x}, t) = A \exp\left[\mathrm{i}(\mathbf{k}\cdot\mathbf{x} - \omega t)\right] \ , \tag{2.1}$$

the change of notation from ϕ to Ψ is meant to emphasise that we now wish to reinterpret this plane wave as a de Broglie matter-wave. We compute

$$\begin{aligned} \mathrm{i}\hbar\frac{\partial\Psi}{\partial t} &= \hbar\omega\Psi \;=\; E\Psi \;, \\ -\mathrm{i}\hbar\nabla\Psi &= \hbar\mathbf{k}\Psi \;=\; \mathbf{p}\Psi \;, \end{aligned} \tag{2.2}$$

where in the second equality on each line we have made use of the de Broglie relations (1.4). Taking the divergence of the second equation furthermore implies that $-\hbar^2\nabla^2\Psi = \hbar^2|\mathbf{k}|^2\Psi = |\mathbf{p}|^2\Psi$, where ∇^2 is the Laplacian.

The de Broglie relations apply to a free particle, for which the particle's energy is equal to its kinetic energy (since $\mathbf{F} = \mathbf{0}$ the potential is constant, and we take this constant to be zero). If the particle has mass m, then $E = E_{\text{kinetic}} = |\mathbf{p}|^2/2m$. Using the de Broglie relations (1.4) then implies

$$\omega \;=\; \frac{E}{\hbar} \;=\; \frac{\hbar|\mathbf{k}|^2}{2m} \;. \tag{2.3}$$

This fixes the angular frequency ω in terms of the wave vector $\mathbf{k}$ and the mass m of the particle. Putting everything together, we have

$$\mathrm{i}\hbar\frac{\partial\Psi}{\partial t} \;=\; E\Psi \;=\; \frac{|\mathbf{p}|^2}{2m}\Psi \;=\; -\frac{\hbar^2}{2m}\nabla^2\Psi \;. \tag{2.4}$$

Thus we may associate to a free particle of mass m a plane wave $\Psi(\mathbf{x}, t)$, using the de Broglie relations, which then satisfies the equation

$$\mathrm{i}\hbar\frac{\partial\Psi}{\partial t} \;=\; -\frac{\hbar^2}{2m}\nabla^2\Psi \;. \tag{2.5}$$

Notice this is essentially expressing the relation (2.3), which is simply the relation between energy and momentum for a free particle.

More generally, a particle of mass m moving in a potential $V = V(\mathbf{x})$ has energy

$$E \;=\; \frac{|\mathbf{p}|^2}{2m} + V \;. \tag{2.6}$$

This led Schrödinger to

Definition / postulate A single, non-relativistic particle of mass m moving in a potential $V(\mathbf{x})$ is described by a *wave function* $\Psi(\mathbf{x}, t)$ that is governed by the *Schrödinger equation*

$$\mathrm{i}\hbar\frac{\partial\Psi}{\partial t} \;=\; -\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi \;. \tag{2.7}$$

The wave function $\Psi(\mathbf{x}, t)$ is precisely de Broglie's matter-wave.

Let us make some remarks about what we have done above, which is deceptively simple, and about the Schrödinger equation (2.7) itself:

(i) It is important to realize that we have not *derived* the Schrödinger equation, in any rigorous sense. In particular, in the last step we have taken formulae that apply to free particles, that have been suggested by a combination of experiments and theoretical arguments, and extrapolated this, using the classical formula (2.6) for energy, to an equation governing the matter-wave of a particle moving in a general potential. It turns out this equation is indeed correct, but nothing in the rather naive argument we gave really guarantees this. The real test of the Schrödinger equation is that it agrees with experiments.

(ii) The Schrödinger equation is a *linear* partial differential equation for a *complex-valued* function $\Psi(\mathbf{x},t)$. Thus if Ψ_1, Ψ_2 are solutions, then so is $\alpha_1\Psi_1 + \alpha_2\Psi_2$, for any complex constants $\alpha_1, \alpha_2 \in \mathbb{C}$. This implies that solutions form a (often infinite-dimensional) *vector space* over $\mathbb{C}$. It is precisely this superposition of wave functions that leads to interference effects, as in the double slit experiment with electrons. Notice that, in contrast with the classical wave equation (0.5), the Schrödinger equation (2.7) is complex, due to the $\mathrm{i} = \sqrt{-1}$ on the left hand side.

(iii) Although we began our exposition by discussing photons, it is important to remark that the photon is a *massless*, *relativistic* particle, and as such is ***not*** governed by the Schrödinger equation. The quantum theory of photons is a much more involved theory, known as *quantum electrodynamics*, that requires an understanding of both classical electromagnetism and Special Relativity, as well as quantum ideas. Quantum electrodynamics was the first example of a gauge theory that plays an important role in both particle physics and modern geometry. We shall only refer to the photon again in the context of emission/absorption in atoms, for which we need only the Einstein-Planck relation.

Before continuing to discuss some basic mathematical properties of the Schrödinger equation, and looking at our first example, let us pause to comment on the change of viewpoint that is already implicit in what we have said so far. Consider the classical problem of a point particle of mass m moving in a potential V. The dynamics is governed by Newton's law (0.1) with $\mathbf{F} = -\nabla V$, the solutions of which give the particle's trajectory $\mathbf{x}(t)$. For given initial conditions, say the particle's position $\mathbf{x}$ and momentum $\mathbf{p} = m\dot{\mathbf{x}}$ at time $t = t_0$, one solves for the trajectory $\mathbf{x}(t)$, which gives the particle's location and momentum at any subsequent time.

The corresponding quantum mechanical problem is very different. Given a quantum point particle of mass m moving in a potential V, we should instead solve the Schrödinger equation (2.7). We shall discuss the boundary conditions involved later, but notice immediately that the result will be some complex-valued function $\Psi(\mathbf{x},t)$. At the beginning of these lectures we emphasized that neither the classical point particle model nor the wave model of a "particle" such as an electron is adequate, and we need a new description. The wave function $\Psi(\mathbf{x},t)$ turns out to be precisely what we're looking for, and indeed it satisfies the wave-like Schrödinger equation (2.7), but how are we to understand particle-like features from a complex-valued function? In particular, you might immediately wonder what this function has to do with the particle's position at some time t. We shall address this shortly.

2.2 Stationary states

It is natural to seek separable solutions to the Schrödinger equation (2.7). Thus we write $\Psi(\mathbf{x},t) = \psi(\mathbf{x})\,T(t)$, so that (2.7) becomes (dividing through both sides by $\Psi = \psi\,T$)

$$\frac{\mathrm{i}\hbar\frac{\mathrm{d}T}{\mathrm{d}t}}{T} \;=\; \frac{-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi}{\psi}\;. \tag{2.8}$$

Since the left hand side depends only on t, while the right hand side depends only on $\mathbf{x}$, both sides must be constant. If we call this constant E (anticipating that this will be the energy of the particle), then in particular we have

$$\mathrm{i}\hbar\frac{\mathrm{d}T}{\mathrm{d}t} \;=\; ET\;, \tag{2.9}$$

which immediately integrates to

$$T(t) \;=\; \mathrm{e}^{-\mathrm{i}Et/\hbar}\;. \tag{2.10}$$

Here we have absorbed the overall multiplicative integration constant into ψ. The full wave function is thus

$$\Psi(\mathbf{x},t) \;=\; \psi(\mathbf{x})\,\mathrm{e}^{-\mathrm{i}Et/\hbar}\;. \tag{2.11}$$

Notice that the wave function (2.11) has angular frequency $\omega = E/\hbar$, so that $E = \hbar\omega$. This is precisely the de Broglie relation (1.4) between energy and angular frequency for matter-waves, allowing us to identify E with the energy of the particle. The function ψ then satisfies:

Definition The *time-independent*, or *stationary state*, *Schrödinger equation* for a particle of mass m and energy E moving in a potential $V = V(\mathbf{x})$ is

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi \;=\; E\psi\;. \tag{2.12}$$

The wave function $\Psi(\mathbf{x},t) = \psi(\mathbf{x})\,\mathrm{e}^{-\mathrm{i}Et/\hbar}$ is then called a *stationary state wave function* of energy E, although in a common abuse of language the function ψ is also often referred to as the stationary state wave function.

The usual strategy for solving the full Schrödinger equation (2.7) is to first find the stationary states $\psi(\mathbf{x})$ solving (2.12). Physically we know that a particle can have different values of the energy E, and in classical mechanics this is a continuous variable. In the quantum mechanical examples we shall study in the remainder of the course instead the allowed values of energy will be discrete, and we may label them by E_n, with $n \in \mathbb{N}$, *cf.* the hydrogen atom energy levels in (1.3). We may then take an arbitrary linear combination of these stationary states

$$\Psi(\mathbf{x}, t) \ = \ \sum_n c_n \, \psi_n(\mathbf{x}) \, \mathrm{e}^{-\mathrm{i}E_n t/\hbar} \ , \tag{2.13}$$

where $c_n \in \mathbb{C}$ are constants, and $\psi_n(\mathbf{x})$ solves (2.12) with $E = E_n$. Linearity of the Schrödinger equation (2.7) implies that (2.13) also satisfies the Schrödinger equation. Under favourable circumstances the *general* solution to the Schrödinger equation may be written this way, as we discuss later.

Recall however that for the free particle treated in section 2.1 the allowed values of E were continuous. There are many such important examples in quantum mechanics, as well as systems, where the energy spectrum has both a discrete and continuous part (e.g. the hydrogen atom). In these cases the sum in (2.13) is replaced by an integral.

2.3 One-dimensional equations

While our world is three-dimensional, it is often the case that a three-dimensional problem effectively reduces to a lower-dimensional Schrödinger equation; for example, due to symmetry reduction (see section 9), or because the particle is constrained to lie in some subspace (e.g. an electron moving in a thin wire). It is also technically more straightforward to study the *one-dimensional Schrödinger equation*

$$\mathrm{i}\hbar \frac{\partial \Psi}{\partial t} \ = \ -\frac{\hbar^2}{2m}\frac{\partial^2 \Psi}{\partial x^2} + V\Psi \ , \tag{2.14}$$

with corresponding stationary state equation

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + V\psi \ = \ E\psi \ , \tag{2.15}$$

and

$$\Psi(x, t) \ = \ \psi(x) \, \mathrm{e}^{-\mathrm{i}Et/\hbar} \ . \tag{2.16}$$

Here we have replaced the Laplacian ∇^2 by the corresponding one-dimensional operator, which is simply $\partial^2/\partial x^2$ for a particle moving on the x-axis. Equations (2.14), (2.15) govern a particle propagating on the x-axis with potential $V = V(x)$.

2.4 Particle in a box

Consider a particle inside a "box" on the x-axis. This means that the particle moves freely inside some interval $[0, a] \subset \mathbb{R}$, with $a > 0$, but cannot leave this region. One can model this by a potential function $V = V(x)$ that is zero inside the interval/box, and (formally) infinite outside:

$$V(x) \ = \ \begin{cases} 0 \ , & 0 < x < a \ , \\ +\infty \ , & \text{otherwise} \ . \end{cases} \tag{2.17}$$

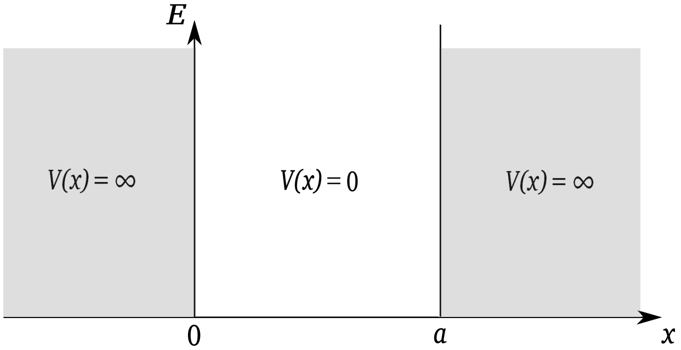

Figure 5: A particle in box $[0, a]$ on the x-axis.

This is also sometimes referred to as the *infinite square well potential.* See Figure 5.

Before discussing the quantum problem, let us briefly comment on the classical problem. Since $V = 0$ inside the box, the classical particle moves at some constant velocity, or equivalently constant momentum p. Since the energy $E = p^2/2m$ is conserved when the particle hits the edge of the box, after the collision p is replaced by $-p$ and the particle heads towards the other side of the box. It thus bounces backwards and forwards with constant speed $|p|/m$. Classically, notice that E may take any non-negative value (it is a continuous variable).

Now let us consider the quantum particle. As explained at the end of section 2.2, we should here solve the stationary state Schrödinger equation (2.15). *Inside the box* this differential equation rearranges to

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = -\frac{2mE}{\hbar^2}\psi \ , \tag{2.18}$$

for $x \in (0, a)$. We shall discuss boundary conditions more systematically in section 3.2, but here we note that since formally $V = \infty$ outside the box, the Schrödinger equation will make sense only if $\psi = 0$ there.[2] When we come to discuss the physical meaning of the wave function, we shall see that this assertion is very well justified physically. If we also assume that ψ is *continuous*, then we must solve (2.18) subject to the *boundary condition* $\psi(0) = \psi(a) = 0$.

The general solution to (2.18) is

$$\psi(x) = \begin{cases} A\cos\frac{\sqrt{2mE}}{\hbar}x + B\sin\frac{\sqrt{2mE}}{\hbar}x \ , & E > 0 \ , \\ A + Bx \ , & E = 0 \ , \\ A\cosh\frac{\sqrt{-2mE}}{\hbar}x + B\sinh\frac{\sqrt{-2mE}}{\hbar}x \ , & E < 0 \ . \end{cases} \tag{2.19}$$

In all cases the boundary condition $\psi(0) = 0$ immediately implies $A = 0$. Imposing this, the boundary condition $\psi(a) = 0$ then gives

$$0 = \psi(a) = \begin{cases} B\sin\frac{\sqrt{2mE}}{\hbar}a \ , & E > 0 \ , \\ Ba \ , & E = 0 \ , \\ B\sinh\frac{\sqrt{-2mE}}{\hbar}a \ , & E < 0 \ . \end{cases} \tag{2.20}$$

---

[2] This is not a very rigorous statement. We can make the discussion rigorous by simply declaring that a particle in a box by definition has $\psi = 0$ outside the box and that ψ is everywhere continuous. See also section 3.2.

Since $a > 0$, when $E \leq 0$ we see from (2.20) that $B = 0$.[3] Thus when $E \leq 0$ the only solution is $\psi \equiv 0$, which is always a physically meaningless solution to the Schrödinger equation (again, we shall discuss this more later). On the other hand, for $E > 0$ the boundary condition $\psi(a) = 0$ implies that (either $B = 0$ and $\psi \equiv 0$ or)

$$\frac{\sqrt{2mE}}{\hbar} = \frac{n\pi}{a} \, , \tag{2.21}$$

for some *integer* $n \in \mathbb{Z}$. Thus the solutions

$$\psi(x) = \psi_n(x) = \begin{cases} B \sin \frac{n\pi x}{a} \, , & 0 < x < a \, , \\ 0 \, , & \text{otherwise} \, , \end{cases} \tag{2.22}$$

are labelled by n. Notice that, without loss of generality, we may take $n > 0$.[4] The first three wave functions, for $n = 1, 2, 3$, are shown in Figure 6.

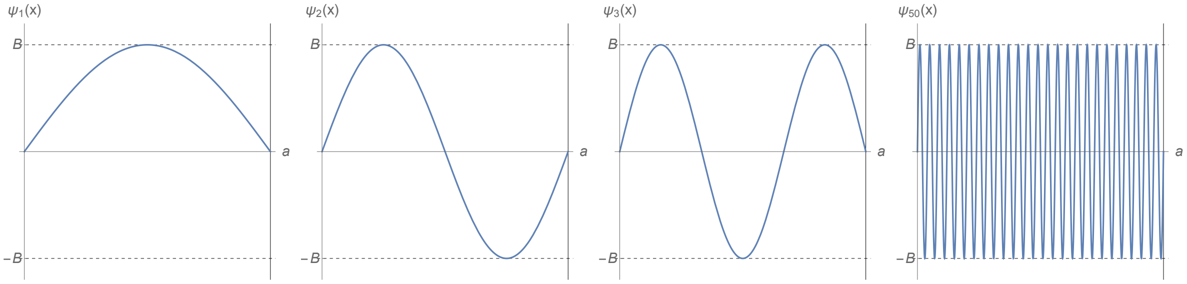

Figure 6: Wave functions for the first three and the 50th states of a particle in a box. $\psi_1(x)$ is the ground state wave function.

The associated energies are, from (2.21),

$$E = E_n = \frac{n^2 \pi^2 \hbar^2}{2ma^2} \, . \tag{2.23}$$

We see immediately that the energy is *quantized*, *i.e.* it takes values in a discrete set, here labelled naturally by a positive integer. This is in stark contrast with the energy of the classical particle, which may take any real non-negative value. There is also a lowest energy, given by setting $n = 1$.

Definition When the possible energies of a quantum system are discrete and bounded below, the lowest possible energy is called the *ground state energy* (also sometimes called the *zero point energy*). The higher energies are, in increasing order, the *first excited state energy*, *second excited state energy*, *etc.* The corresponding wave functions are called the *ground state wave function*, kth *excited state wave function.*

---

[3] In particular, note that the only real solution to $\sinh y = 0$ is $y = 0$.

[4] In more detail: note that $B \sin \frac{(-n)\pi x}{a} = (-B) \sin \frac{n\pi x}{a}$, and thus the solutions with $n < 0$ are the same as the solutions with $n > 0$, on taking the integration constant $B \to -B$.

* Are these values of energy reasonable? For the particle in a box the ground state energy is $E_1 = \pi^2\hbar^2/2ma^2$, while $E_n = n^2 E_1$. Of course, we precisely wanted a theoretical understanding of such quantized energies in order to explain the energy levels of the hydrogen atom, determined empirically as (1.3). For the particle in a box, if we take $m = m_{e^-} \simeq 9.11 \times 10^{-31}$ kg to be the mass of an electron and $a = 10^{-10}$ m to be the approximate size of an atom, we obtain

$$E_n \simeq 5.97 \times 10^{-18} n^2 \text{ J} . \tag{2.24}$$

In particular, the difference in energies between the ground state and first excited state is $E_2 - E_1 \simeq 1.79 \times 10^{-17}$ J. A photon that is emitted in a transition between these energy levels then has a wavelength $\lambda \simeq 1.12 \times 10^{-8}$ m (on the boundary between the ultraviolet and X-ray parts of the electromagnetic spectrum), which is indeed the correct order of magnitude observed in atomic transitions! Here we are effectively modelling a hydrogen atom as an electron confined to an atom-sized box, which is very crude; we shall treat the hydrogen atom more precisely in section 9. Nevertheless, the above computation is encouraging.

The full time-dependent wave functions (2.16) are

$$\Psi_n(x,t) = \psi_n(x)\, \mathrm{e}^{-\mathrm{i}E_n t/\hbar} = \begin{cases} B\sin\frac{n\pi x}{a}\, \mathrm{e}^{-\mathrm{i}n^2\pi^2\hbar t/2ma^2} , & 0 < x < a , \\ 0 , & \text{otherwise} . \end{cases} \tag{2.25}$$

Any linear combination of such wave functions, as in (2.13), satisfies the time-dependent Schrödinger equation (2.14). In particular, the space of solutions, or possible wave functions, is *infinite dimensional*. In this example the energy levels E_n, and associated stationary state wave functions ψ_n, are labelled naturally by a (positive) integer n. As we shall see throughout the course, such integers arise in many important solutions to the Schrödinger equation, and they are generally known as *quantum numbers* (although it is difficult to give a precise general definition).

2.5 Particle in a three-dimensional box

Having studied a particle in a one-dimensional box, it is now straightforward to extend this to a three-dimensional box.

Consider a particle confined to the box region given by $\{\mathbf{x} = (x,y,z) \in \mathbb{R}^3 \mid 0 \le x \le a,\ 0 \le y \le b,\ 0 \le z \le c\} \subset \mathbb{R}^3$, where the potential is zero inside the box. In other words,

$$V(x,y,z) = \begin{cases} 0 , & 0 < x < a ,\ 0 < y < b ,\ 0 < z < c , \\ +\infty , & \text{otherwise} . \end{cases} \tag{2.26}$$

As before, the stationary state wave function $\psi(\mathbf{x}) = \psi(x,y,z)$ is taken to be zero on, and outside, the boundary of the box region. Inside the box the stationary state Schrödinger equation (2.12) reduces to

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} = -\frac{2mE}{\hbar^2}\psi . \tag{2.27}$$

This may be solved by separation of variables. Equation (2.27) essentially reduces to three copies of the one-dimensional equation, with the wave functions labelled by three quantum numbers

$n_1, n_2, n_3 \in \mathbb{Z}_{>0}$. Explicitly, inside the box these are given by

$$\psi_{n_1,n_2,n_3}(x,y,z) \;=\; B\,\sin\frac{n_1\pi x}{a}\,\sin\frac{n_2\pi y}{b}\,\sin\frac{n_3\pi z}{c}\;, \tag{2.28}$$

with B again an arbitrary constant, and the corresponding energies are

$$E_{n_1,n_2,n_3} \;=\; \frac{\pi^2\hbar^2}{2m}\left(\frac{n_1^2}{a^2}+\frac{n_2^2}{b^2}+\frac{n_3^2}{c^2}\right)\;. \tag{2.29}$$

Exercise (Problem Sheet 1) Derive the wave functions (2.28) and energies (2.29) by solving (2.27) by separation of variables.

2.6 Degeneracy

Just as the full Schrödinger equation (2.7) is linear, similarly the stationary state Schrödinger equation (2.12) is also linear, but only for a *fixed* energy E. Suppose that $\psi_i(\mathbf{x})$ are *linearly independent* stationary state wave functions satisfying (2.12) with the *same* energy E, where $i = 1,\ldots,d \geq 1$. Then any linear combination

$$\psi \;=\; \sum_{i=1}^{d} c_i\psi_i \tag{2.30}$$

also solves (2.12), for any complex constants $c_i \in \mathbb{C}$. Notice that we may then view the space of solutions as a *vector space* $\mathbb{C}^d$, where $\psi_i(\mathbf{x})$ form a basis of solutions, with coefficients/components $(c_1,\ldots,c_d) \in \mathbb{C}^d$. This leads to the following definition:

Definition If the space of solutions to the stationary state Schrödinger equation (2.12) with energy E has dimension $d > 1$, we say this energy level is *d-fold degenerate*; if it is one-dimensional we say E is a *non-degenerate* energy level.

For the one-dimensional particle in a box all the energy levels are non-degenerate. However, consider the three-dimensional box with equal length sides $a = b = c$, so that

$$E_{n_1,n_2,n_3} \;=\; \frac{\pi^2\hbar^2}{2ma^2}\left(n_1^2+n_2^2+n_3^2\right)\;. \tag{2.31}$$

In this case there are linearly independent wave functions with the *same* energy. For example, we may take (n_1,n_2,n_3) to be any of $(2,1,1)$, $(1,2,1)$, $(1,1,2)$, all of which have the same energy $E = 6\pi^2\hbar^2/2ma^2$. The corresponding wave functions are given in (2.28), and are respectively proportional to $\sin\frac{2\pi x}{a}\sin\frac{\pi y}{b}\sin\frac{\pi z}{c}$, $\sin\frac{\pi x}{a}\sin\frac{2\pi y}{b}\sin\frac{\pi z}{c}$, $\sin\frac{\pi x}{a}\sin\frac{\pi y}{b}\sin\frac{2\pi z}{c}$, which are linearly independent. There is thus a three-fold degeneracy in the number of quantum stationary states with this energy. We shall see other examples of degenerate energy levels, with $d > 1$, later in the lectures. Note that the reason for the degeneracy is the permutation symmetry between n_1, n_2, n_3 in the expression of the energy, which originates from the equality of wall length or ultimately in the symmetry of the problem. In generic systems we do not expect any degeneracy of the spectrum, and if there is, there is almost always a symmetry explanation for it. We will see a striking example of this lore for the hydrogen atom in section 9.

Notice that when there are degenerate energy levels (2.13) can be written more precisely as

$$\Psi(\mathbf{x},t) \;=\; \sum_n \sum_{i=1}^{d_n} c_{n,i}\,\psi_{n,i}(\mathbf{x})\,\mathrm{e}^{-\mathrm{i}E_n t/\hbar}\;. \tag{2.32}$$

Here we have assumed that the energy level E_n has degeneracy $d_n \geq 1$, and $\psi_{n,i}(\mathbf{x})$ are linearly independent stationary states of energy E_n, $i = 1, \ldots, d_n$.

2.7 Particle on a circle

We consider a variation of the particle in a box problem: a free particle moving on a circle of radius R. This setup is modelled by a one-dimensional Schrödinger equation with the potential set to zero and the spatial coordinate x periodically identified, $x \sim x + 2\pi R$, and wave functions obeying $\Psi(t,x) = \Psi(t, x+2\pi R)$.

The stationary state Schrödinger equation takes the form,

$$-\frac{\hbar^2}{2m}\partial_{xx}\psi(x) \;=\; E\psi(x)\,. \tag{2.33}$$

This is the same equation as for the particle in a box, however here we imposed periodicity

$$\psi(x) = \psi(x+2\pi R)\,. \tag{2.34}$$

This is different from the box problem (with $a = 2\pi R$), where we had $\psi(0) = \psi(2\pi R) = 0$.[5] It is the boundary conditions that lead to the discreteness of the energy levels.

Going through the same steps that lead to (2.19) for the box, we find that the ground state is the constant function:

$$\psi_{\text{g.s.}} \;=\; A\,, \qquad E_{\text{g.s.}} \;=\; 0\,, \tag{2.36}$$

and the excited states are

$$\psi(x) \;=\; A\cos\frac{\sqrt{2mE}}{\hbar}x + B\sin\frac{\sqrt{2mE}}{\hbar}x\;, \qquad E \;=\; E_n \;=\; \frac{n^2\hbar^2}{2mR^2} \qquad (n\in\mathbb{Z}^+)\,. \tag{2.37}$$

Since sine and cosine are linearly independent functions, we find that each excited state is doubly degenerate, $d_n = 2$. However the ground state is nondegenerate.

---

[5] Periodicity can be rewritten to resemble the box problem: Let us restrict $x \in [0, 2\pi R)$, then the periodicity condition can be imposed as

$$\psi(0) = \psi(2\pi R)\,, \qquad \psi'(0) = \psi'(2\pi R)\,, \tag{2.35}$$

where the condition on derivatives is imposed to rule out periodic functions with discontinuous derivatives like the $2\pi R$-periodic extension of $\sin\left(\frac{x}{2R}\right)$, which has a kink at $x = 2\pi R\mathbb{Z}$. Kinks are disallowed, since they lead to the nonexistence of $\partial_{xx}\psi(x)$ that is the left hand side of (2.33).

3 The Born interpretation

We have now met the Schrödinger equation and solved it in the simplest interesting example, namely a particle confined to a box. We have seen that this leads to quantized energy levels, and that by crudely modelling a hydrogen atom as an electron confined to an atom-sized box, we obtain energies of the correct order of magnitude seen in atomic transitions.

An immediate question is: *what is the physical meaning of the wave function* $\Psi(\mathbf{x},t)$ *that we are solving for?* Comparing to the corresponding classical problem, described at the end of section 2.1, we may also ask: where is the particle at time t? In this section we shall answer these questions.

3.1 Probability density

In order to motivate the interpretation that follows, we begin by going back to the double slit experiment in section 1.3. In fact let us begin by discussing the corresponding experiment with light (also called *Young's experiment*). In this case the *intensity* of the light hitting the detector screen forms an interference pattern, and in classical electromagnetic theory this intensity is proportional to the *absolute value squared of the amplitude of the wave.* Unfortunately we won't have time to explain this in detail here, but this fact would have been well-known to the pioneers of quantum theory in the early 20th century. In the double slit experiment with electrons, we instead plot the spatial distribution of electrons hitting the detector screen over some long period of time, and then reinterpret this as a *probability distribution* for where any given electron will hit the screen.

If we now conflate these observations, we are led to the hypothesis that the *probability density function* for an electron hitting the detector screen is given by the *absolute value squared of the de Broglie matter-wave* associated to the electron. Of course, the detector screen could be anywhere, and the de Broglie matter-wave is precisely the wave function $\Psi(\mathbf{x},t)$ appearing in the Schrödinger equation. Hence we arrive at:

Definition / postulate The function

$$\rho(\mathbf{x},t) \equiv |\Psi(\mathbf{x},t)|^2 , \tag{3.1}$$

is a *probability density function* for the position of the particle, where $\Psi(\mathbf{x},t)$ is the particle's wave function.

This interpretation of Schrödinger's wave function is due to Born.

The assertion (3.1) is equivalent to the statement that the probability of finding the particle in a volume $D \subset \mathbb{R}^3$ is given by

$$\mathbb{P}_\Psi(D) = \iiint_D |\Psi(\mathbf{x},t)|^2 \, \mathrm{d}^3 x , \tag{3.2}$$

where we have denoted $\mathrm{d}^3x = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$. This probability depends both on the region D, and also on the wave function $\Psi(\mathbf{x},t)$ satisfying the Schrödinger equation. We have already seen for the particle in a box that the space of solutions to the Schrödinger equation is an infinite-dimensional vector space, with basis (2.25). A given solution is said to describe the *state* of the particle, so that the probabilities (3.2) depend on the state. Notice $\mathbb{P}_\Psi(D)$ also depends in general on time t, although this is suppressed in the notation.

Of course, the above assertions immediately raise some issues. In particular, the probability of finding the particle *somewhere* in $\mathbb{R}^3$ should equal 1, at any time. Thus for (3.1), or equivalently (3.2), to make sense, the wave function must be *normalized* in the following sense:

Definition A wave function Ψ is said to be *normalizable* if

$$0 < \iiint_{\mathbb{R}^3} |\Psi(\mathbf{x},t)|^2 \, \mathrm{d}^3x < \infty \, , \qquad \forall \text{ time } t \, . \tag{3.3}$$

Moreover, if

$$\iiint_{\mathbb{R}^3} |\Psi(\mathbf{x},t)|^2 \, \mathrm{d}^3x = 1 \, , \qquad \forall \text{ time } t \, , \tag{3.4}$$

then Ψ is said to be a *normalized wave function.*

Similar definitions apply in one dimension. For example, the normalized condition (3.4) becomes

$$\int_{-\infty}^{\infty} |\Psi(x,t)|^2 \, \mathrm{d}x = 1 \, . \tag{3.5}$$

Let us make some remarks:

(i) Notice that the Born interpretation says nothing about where the particle actually is if we do not measure its position. Instead we compute the probability of finding the particle in a particular region, should we make such a measurement. This is *very* different from classical mechanics, and is not very intuitive. We shall see later in section 5 that measurement of position is a special case of more general *quantum measurement postulates*, which apply to any physical observable – for example, momentum, angular momentum, and energy. In each case the value of the observable is simply not defined unless we measure it, and we will only ever deal with *probabilities* of obtaining particular values for those measurements.

(ii) The normalized condition (3.4) fixes the freedom to multiply a solution to the Schrödinger equation by a complex constant, up to a constant *phase* $\mathrm{e}^{\mathrm{i}\varphi}$. The latter is in fact not physical, so that wave functions differing by a constant phase are *physically equivalent.* Notice also that the both the normalizable and normalized conditions rule out the *trivial solution* $\Psi \equiv 0$.

(iii) For the particle in a box, the wave functions (2.25) are normalizable:

$$\int_{-\infty}^{\infty} |\Psi_n(x,t)|^2 \, \mathrm{d}x = |B|^2 \int_0^a \sin^2 \frac{n\pi x}{a} \, \mathrm{d}x = \frac{1}{2} a |B|^2 \, . \tag{3.6}$$

The total wave functions

$$\Psi_n(x,t) = \begin{cases} \sqrt{\frac{2}{a}} \, \sin \frac{n\pi x}{a} \, \mathrm{e}^{-\mathrm{i}n^2\pi^2\hbar t/2ma^2} \, , & 0 < x < a \, , \\ 0 \, , & \text{otherwise} \, , \end{cases} \tag{3.7}$$

with $B = \sqrt{2/a}$, are then *normalized.* Note that a particle being "confined to a box" may be interpreted as saying that there is *zero probability* of finding it anywhere outside the box, or in other words that the wave function is identically zero outside the box.

(iv) More generally, notice that for a stationary state of energy E, we have

$$\rho(\mathbf{x},t) \;=\; |\Psi(\mathbf{x},t)|^2 \;=\; \left|\psi(\mathbf{x})\,\mathrm{e}^{-\mathrm{i}Et/\hbar}\right|^2 \;=\; |\psi(\mathbf{x})|^2 \;. \tag{3.8}$$

Thus $\Psi(\mathbf{x},t)$ is normalized for all time provided $\psi(\mathbf{x})$ is normalized, *i.e.*

$$\iiint_{\mathbb{R}^3} |\psi(\mathbf{x})|^2\,\mathrm{d}^3x \;=\; 1 \;. \tag{3.9}$$

(v) Recall that the plane wave

$$\Psi(\mathbf{x},t) \;=\; A\,\exp\left[\mathrm{i}\left(\mathbf{k}\cdot\mathbf{x} - \frac{\hbar|\mathbf{k}|^2 t}{2m}\right)\right] \;, \tag{3.10}$$

satisfies the free Schrödinger equation ($V = 0$), and was interpreted as describing a free particle of mass m and momentum $\mathbf{p} = \hbar\mathbf{k}$ – see the discussion around equation (2.3). However, clearly $|\Psi(\mathbf{x},t)|^2 = |A|^2$ is not integrable over $\mathbb{R}^3$, and thus the plane wave is *not* normalizable! We shall return to discuss this further in sections 5 and 6.[6]

Example (Particle in a box) Let us apply the above ideas to the particle in a one-dimensional box. The normalized wave functions are (3.7). The corresponding probability density functions are then, for $x \in [0,a]$, given by

$$\rho_n(x) \;=\; \frac{2}{a}\sin^2\frac{n\pi x}{a} \;=\; \frac{1}{a}\left(1-\cos\frac{2n\pi x}{a}\right) \;, \tag{3.11}$$

and identically zero outside the box. Plots of the probability density functions for the ground state and first two excited states are shown in Figure 7.

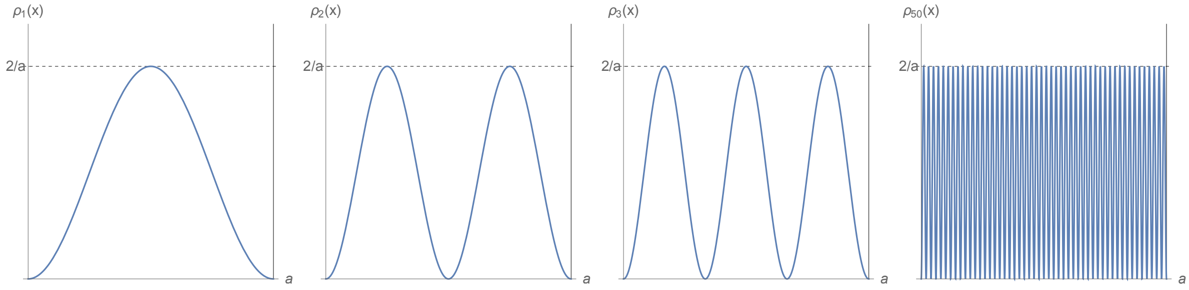

Figure 7: Probability density functions for the first three and the 50th states of a particle in a box.

As usual in probability theory, we may define the *distribution function* as

$$F_n(x) \;\equiv\; \int_0^x \rho_n(y)\,\mathrm{d}y \;=\; \frac{x}{a} - \frac{1}{2n\pi}\sin\frac{2n\pi x}{a} \;, \tag{3.12}$$

---

[6] * In fact the plane wave is also sometimes interpreted as a *beam* of particles of mass m and momentum $\mathbf{p}$, although exploring this further is beyond our syllabus. Readers who would like to learn more about this topic might like to read Chapter 5 of the book by Hannabuss.

which is the probability that the particle lies in the interval $[0, x]$. Notice in particular that the first term, $\frac{x}{a}$, is the result for the *uniform distribution*, where the particle is equally likely to be found anywhere in the box. This is interpreted as the *classical result*. By this we mean that if we are ignorant of the state of the classical particle before we observe its position, then because it moves at constant speed back and forth across the box, it is equally likely to be found anywhere. The second, oscillatory term in (3.12) is then a *quantum contribution*, or *quantum correction*, to the classical result.

Let us compute some example probabilities using these formulae. The probability that the particle is within $\frac{a}{4}$ distance from the centre of the box is given by

$$\begin{aligned} \int_{\frac{a}{4}}^{\frac{3a}{4}} |\psi_n(x)|^2 \, \mathrm{d}x &= F_n\left(\frac{3a}{4}\right) - F_n\left(\frac{a}{4}\right) \ , \\ &= \frac{1}{2} + \frac{1}{n\pi} \sin \frac{n\pi}{2} \ , \\ &= \begin{cases} \frac{1}{2} \ , & n \text{ even} \ , \\ \frac{1}{2} + \frac{(-1)^{(n-1)/2}}{n\pi} \ , & n \text{ odd} \ . \end{cases} \end{aligned} \tag{3.13}$$

In particular, we see that this *approaches the classical result* of $\frac{1}{2}$, for the uniform distribution, as $n \to \infty$. The tendency of quantum results to approach those of the corresponding classical theory for large quantum numbers is called the *correspondence principle*.

We may similarly use the probability density $\rho = |\Psi|^2$ to compute *expectation values*:

Definition In quantum mechanics the *expectation value of a function of position* $f(\mathbf{x})$ is denoted

$$\mathbb{E}_\Psi(f(\mathbf{x})) \equiv \iiint_{\mathbb{R}^3} f(\mathbf{x}) \, |\Psi(\mathbf{x}, t)|^2 \, \mathrm{d}^3 x \ , \tag{3.14}$$

or similarly in one dimension

$$\mathbb{E}_\Psi(f(x)) \equiv \int_{-\infty}^{\infty} f(x) \, |\Psi(x, t)|^2 \, \mathrm{d}x \ . \tag{3.15}$$

Notice that this expectation value depends on the wave function/state Ψ, and in general is a function of time t.

Example For the particle in a box we may use (3.12) to compute the expected value of the particle's position

$$\begin{aligned} \mathbb{E}_{\Psi_n}(x) &= \int_0^a x \, \rho_n(x) \, \mathrm{d}x \ , \\ &= [x \, F_n(x)]_0^a - \int_0^a F_n(x) \, \mathrm{d}x \ , \\ &= a - \left[\frac{x^2}{2a} + \frac{a}{4n^2\pi^2} \cos \frac{2n\pi x}{a} \right]_0^a \ , \\ &= \frac{1}{2} a \ , \end{aligned} \tag{3.16}$$

agreeing with the classical result for the uniform distribution.

3.2 The continuity equation and boundary conditions

In this section we consider more carefully the *boundary conditions* involved in the Schrödinger equation.

Proposition 3.1 *The Schrödinger equation implies the* continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} \, = \, 0 \ , \tag{3.17}$$

where $\rho(\mathbf{x},t) = |\Psi(\mathbf{x},t)|^2$ *is the probability density we have already met, and the vector field*

$$\mathbf{j}(\mathbf{x},t) \, \equiv \, \frac{\mathrm{i}\hbar}{2m}\left(\Psi(\mathbf{x},t)\overline{(\nabla\Psi)(\mathbf{x},t)} - \overline{\Psi(\mathbf{x},t)}(\nabla\Psi)(\mathbf{x},t)\right) \ , \tag{3.18}$$

is called the probability current **j**.

Proof This is a direct computation:

$$\begin{aligned}
\frac{\partial}{\partial t}\left|\Psi(\mathbf{x},t)\right|^2 \, &= \, \left(\frac{\partial}{\partial t}\overline{\Psi(\mathbf{x},t)}\right)\Psi(\mathbf{x},t) + \overline{\Psi(\mathbf{x},t)}\frac{\partial}{\partial t}\Psi(\mathbf{x},t)\,, \\
&= \, \overline{\left[-\frac{\mathrm{i}}{\hbar}\left(-\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi\right)\right]}\Psi + \overline{\Psi}\left[-\frac{\mathrm{i}}{\hbar}\left(-\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi\right)\right], \quad \text{(Schrödinger)} \\
&= \, \frac{\mathrm{i}}{\hbar}\left(-\frac{\hbar^2}{2m}\nabla^2\overline{\Psi} + V\overline{\Psi}\right)\Psi - \overline{\Psi}\,\frac{\mathrm{i}}{\hbar}\left(-\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi\right) \ , \quad (V \text{ is real}) \\
&= \, \frac{\mathrm{i}\hbar}{2m}\left(\overline{\Psi}\,\nabla^2\Psi - \Psi\,\nabla^2\overline{\Psi}\right) \ , \\
&= \, \frac{\mathrm{i}\hbar}{2m}\nabla\cdot\left(\overline{\Psi}\,\nabla\Psi - \Psi\,\nabla\overline{\Psi}\right) \, = \, -\nabla\cdot\mathbf{j}\ .\blacksquare
\end{aligned} \tag{3.19}$$

This leads to the following result:

Proposition 3.2 *Suppose that for all time* t, $\mathbf{j}(\mathbf{x},t)$ *satisfies the boundary condition that it tends to zero faster than* $1/|\mathbf{x}|^2$ *as* $|\mathbf{x}| = r \to \infty$. *Then*

$$\iiint_{\mathbb{R}^3} |\Psi(\mathbf{x},t)|^2 \, \mathrm{d}^3 x \tag{3.20}$$

is independent of t. *In particular, if* Ψ *is normalized at some time* $t = t_0$, *it is normalized* $\forall t$.

Proof Let S be a closed surface that bounds a region $D \subset \mathbb{R}^3$. Then

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\iiint_D |\Psi(\mathbf{x},t)|^2\,\mathrm{d}^3x \, &= \, \iiint_D \frac{\partial\rho}{\partial t}\,\mathrm{d}^3x \ , \qquad &&\text{(derivative through the integral)} \\
&= \, \iiint_D (-\nabla\cdot\mathbf{j}\,)\,\mathrm{d}^3x \ , &&\text{(continuity equation)} \\
&= \, -\iint_S \mathbf{j}\cdot\mathrm{d}\mathbf{S} \ , &&\text{(Divergence Theorem)}\ .
\end{aligned} \tag{3.21}$$

Now take S to be a sphere of radius $r > 0$, centred on the origin, so that D is a ball. Then the outward pointing unit normal vector to S is $\mathbf{n} = \dfrac{\mathbf{x}}{r}$, and by definition

$$\mathbf{j}\cdot\mathrm{d}\mathbf{S} \, = \, \mathbf{j}\cdot\mathbf{n}\,r^2\,\mathrm{d}S_{\text{unit}} \ , \tag{3.22}$$

where $\mathrm{d}S_{\text{unit}} = \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi$ is the area element on a *unit* radius sphere, in spherical polar coordinates. In general, the function $\mathbf{j}\cdot\mathbf{n}$ will depend on r and the angular variables θ,ϕ on the sphere (and on time t). Provided $\mathbf{j}\cdot\mathbf{n} = o(1/r^2)$, uniformly in the angular coordinates, then using (3.22) the surface integral in (3.21) tends to zero as $r\to\infty$, and hence

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint_{\mathbb{R}^3} |\Psi(\mathbf{x},t)|^2\,\mathrm{d}^3x \;=\; 0\;, \tag{3.23}$$

which implies that the expression in (3.20) is independent of t. ■

Notice that (3.21) says that for any region $D\subset\mathbb{R}^3$ with closed boundary $S=\partial D$ we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathbb{P}_\Psi(D) \;=\; -\iint_S \mathbf{j}\cdot\mathrm{d}\mathbf{S}\;. \tag{3.24}$$

That is, the rate of change of probability that the particle is in the region D is minus the *flux* of the probability current $\mathbf{j}$ *out* of the bounding surface S. The vector field $\mathbf{j}$ thus gives the “flow” of probability. For example, if the particle is very likely to be found in some region D_1 at time t_1, but then at a later time is instead very likely to be found in region D_2 at time $t_2 > t_1$, then the probability current $\mathbf{j}$ will point from D_1 towards D_2.

Proposition 3.2 is important for the consistency of the Born interpretation. Suppose we are given the wave function $\Psi(\mathbf{x},t_0)$ at some initial time $t=t_0$. If this is normalizable, then by multiplying by an appropriate constant we may assume that it is normalized, so $\iiint_{\mathbb{R}^3}|\Psi(\mathbf{x},t_0)|^2\,\mathrm{d}^3x = 1$. The wave function $\Psi(\mathbf{x},t)$ at some later time $t>t_0$ is obtained by solving the time-dependent Schrödinger equation (2.7) with this initial condition. Proposition 3.2 guarantees that this solution is also normalized, so that $\rho=|\Psi|^2$ may be interpreted as a probability density function for all time t. We discuss this initial value problem further in section 3.3 below.

We now state more formally the conditions that solutions to the Schrödinger equation should satisfy:

(i) *The wave function $\Psi(\mathbf{x},t)$ should be a continuous, single-valued function.* This condition ensures that the probability density $\rho=|\Psi|^2$ is single-valued and has no discontinuities. We already imposed the continuity property for the particle in a box.

(ii) *Ψ should be normalizable, i.e. the integral of $|\Psi|^2$ over all space should be finite, and non-zero.* As already mentioned, this condition may be relaxed for free particles.

(iii) *$\nabla\Psi$ should be continuous everywhere, except where there is an infinite discontinuity in the potential V.* This condition follows since a finite discontinuity in $\nabla\Psi$ implies an infinite discontinuity in $\nabla^2\Psi$, and thus from the Schrödinger equation an infinite discontinuity in V. Again, we encountered precisely this behaviour for the particle in a box.

3.3 Measurement of energy

In section 3.1 we discussed the measurement of position in quantum mechanics. The only other dynamical quantity that we have discussed in any detail so far is energy, and in this section we give a brief description of its measurement in quantum mechanics, focusing on the example of a particle in a box. We shall give a more general treatment of quantum measurement, that applies also to other physical quantities, in section 5.

In section 2.2 we found separable solutions to the time-dependent Schrödinger equation (2.7). The corresponding stationary states (2.11) have definite frequency, and hence definite energy E. For example, the complete set of stationary state wave functions for a particle in a box is given by (3.7). As already mentioned, since the time-dependent Schrödinger equation is *linear*, any linear combination (2.13) of such stationary state wave functions also solves the Schrödinger equation. For the particle in a box, (2.13) is simply a *Fourier sine series*. From Prelims you know that any sufficiently well-behaved[7] function $f : [0, a] \to \mathbb{C}$ with $f(0) = f(a) = 0$ can be expanded as a Fourier sine series

$$f(x) = \sum_{n=1}^{\infty} c_n \sqrt{\frac{2}{a}} \sin \frac{n\pi x}{a} = \sum_{n=1}^{\infty} c_n \, \psi_n(x) \ , \tag{3.25}$$

for appropriate $c_n \in \mathbb{C}$. We can then consider the following *initial value problem* for the Schrödinger equation for the particle in a box. Suppose that the wave function at time $t = 0$ is given by $\Psi(x, 0) = f(x)$. Using the fact that the stationary state wave functions $\psi_n(x) = \sqrt{\frac{2}{a}} \sin \frac{n\pi x}{a}$ are *orthonormal*

$$\int_0^a \psi_m(x) \, \psi_n(x) \, \mathrm{d}x = \delta_{mn} \ , \tag{3.26}$$

we may compute the coefficients c_n in the Fourier sine series (3.25) as

$$\int_0^a f(x) \psi_m(x) \, \mathrm{d}x = \sum_{n=1}^{\infty} c_n \int_0^a \psi_n(x) \psi_m(x) \, \mathrm{d}x = c_m \ . \tag{3.27}$$

Here in the first equality we have substituted the expansion (3.25), and in the second equality we have used (3.26). The functions $\psi_n(x)$ are stationary states of energy E_n, with E_n given by (2.23), so restoring the phase factors $\mathrm{e}^{-\mathrm{i}E_n t/\hbar}$ we immediately obtain a corresponding solution to the time-dependent Schrödinger equation using (2.13):

$$\Psi(x, t) = \sum_{n=1}^{\infty} c_n \, \psi_n(x) \, \mathrm{e}^{-\mathrm{i}E_n t/\hbar} \ . \tag{3.28}$$

By construction this solves the time-dependent Schrödinger equation with initial condition $\Psi(x, 0) = f(x)$, where the coefficients c_n are given in terms of $f(x)$ by (3.27).

---

[7]* For the application to wave functions below we require the wave function to be *normalizable*. For the particle in a box example, this implies that $\int_0^a |f(x)|^2 \, \mathrm{d}x < \infty$. In fact this is then sufficient for the Fourier series (3.25) to converge almost everywhere to $f(x)$.

Definition / interpretation Suppose that the *normalized* wave function for a particle in a box is given by (3.28). Then the *probability* of measuring the energy of the particle to be E_n is $|c_n|^2$.

This definition makes sense, since

$$1 = \int_0^a |\Psi(x,t)|^2 \, \mathrm{d}x = \sum_{m,n=1}^{\infty} \overline{c}_m \, c_n \, \mathrm{e}^{-\mathrm{i}(E_n - E_m)t/\hbar} \int_0^a \overline{\psi}_m(x) \, \psi_n(x) \, \mathrm{d}x = \sum_{n=1}^{\infty} |c_n|^2 \ . \tag{3.29}$$

Here in the second equality we have substituted the expansion (3.28), while the last equality uses (3.26), noting that $\psi_m(x)$ is real.

Notice that if the particle's wave function is a stationary state of energy E_n, then by definition the only non-zero coefficient in (3.28) is c_n, and the probability of measuring the particle's energy to be E_n is then $|c_n|^2 = 1$. It thus makes sense to say that a stationary state (2.11) describes a particle of *definite energy* E. On the other hand, wave functions of the general form (3.28), involving two or more stationary states of different energies, do not describe a particle with a definite energy. In some sense, the particle is in a linear combination of states with different energies, but the value of the energy itself is then not well-defined. All we can say is that if we make a measurement of the energy, there is a probability $|c_n|^2$ of obtaining the value E_n.

Example Suppose that the particle in a box $[0, a]$ has an initial wave function

$$\Psi(x,0) = \begin{cases} Ax(a-x) \ , & 0 < x < a \ , \\ 0 \ , & \text{otherwise} \ . \end{cases} \tag{3.30}$$

Find the *normalized* wave function $\Psi(x,t)$, and compute the probability that the particle is found in the ground state.

To normalize $\Psi(x,0)$ we first compute

$$\int_0^a |\Psi(x,0)|^2 \, \mathrm{d}x = |A|^2 \int_0^a x^2(a-x)^2 \, \mathrm{d}x = |A|^2 \left[a^2 \frac{x^3}{3} - 2a \frac{x^4}{4} + \frac{x^5}{5} \right]_0^a = |A|^2 \frac{a^5}{30} \ . \tag{3.31}$$

Recalling that the constant phase in A is arbitrary, without loss of generality we may take $A = \sqrt{30/a^5}$ to normalize $\Psi(x,0)$. The coefficients c_n in the Fourier expansion of $\Psi(x,0)$ are then given by (3.27), so using the stationary states $\psi_n(x) = \sqrt{\frac{2}{a}} \sin \frac{n\pi x}{a}$ we calculate

$$\begin{aligned} c_n &= \int_0^a \Psi(x,0) \psi_n(x) \, \mathrm{d}x = \sqrt{\frac{2}{a}} \sqrt{\frac{30}{a^5}} \int_0^a x(a-x) \sin \frac{n\pi x}{a} \, \mathrm{d}x \\ &= \frac{2\sqrt{15}}{a^3} \left\{ \left[-\left(\frac{a}{n\pi} \right) x(a-x) \cos \frac{n\pi x}{a} \right]_0^a + \left(\frac{a}{n\pi} \right) \int_0^a (a-2x) \cos \frac{n\pi x}{a} \, \mathrm{d}x \right\} \\ &= \frac{2\sqrt{15}}{a^3} \left\{ \left[\left(\frac{a}{n\pi} \right)^2 (a-2x) \sin \frac{n\pi x}{a} \right]_0^a + 2 \left(\frac{a}{n\pi} \right)^2 \int_0^a \sin \frac{n\pi x}{a} \, \mathrm{d}x \right\} \\ &= 2 \frac{2\sqrt{15}}{a^3} \left(\frac{a}{n\pi} \right)^3 \left[-\cos \frac{n\pi x}{a} \right]_0^a \\ &= \frac{4\sqrt{15}}{n^3 \pi^3} \left[1 - (-1)^n \right] \\ &= \begin{cases} 0 \ , & n \text{ even} \ , \\ \frac{8\sqrt{15}}{n^3\pi^3} \ , & n \text{ odd} \ . \end{cases} \end{aligned} \tag{3.32}$$

Here we have integrated by parts on each line. From (3.28) the wave function at time t is hence

$$\Psi(x,t) \;=\; \sqrt{\frac{30}{a}}\left(\frac{2}{\pi}\right)^3 \sum_{n=1,3,5,7,\ldots} \frac{1}{n^3}\sin\frac{n\pi x}{a}\,\mathrm{e}^{-\mathrm{i}n^2\pi^2\hbar t/2ma^2}\ , \tag{3.33}$$

where we have substituted for the energy levels E_n using (2.23). Finally, the probability that the particle is in the ground state is simply

$$\mathbb{P}_\Psi(\text{ground state}) \;=\; |c_1|^2 \;=\; \left(\frac{8\sqrt{15}}{\pi^3}\right)^2 \;\simeq\; 0.9986\ . \tag{3.34}$$

The fact that this is so close to 1 is because the initial wave function is so closely approximated by the ground state wave function – see Figure 6.

Although we have focused on the particle in a box in this subsection, in fact the above discussion of energy generalizes to other quantum mechanical systems, and also to other dynamical observables. We shall return to this topic in section 5.

4 The harmonic oscillator

The quantum harmonic oscillator is ubiquitous in physics, and as such is probably the most important solution to the Schrödinger equation. In this section we give a treament based on solving the Schrödinger equation explicitly, but shall later see that there are more powerful algebraic methods that allow us to derive the same results much more elegantly.

4.1 The one-dimensional harmonic oscillator

Consider a classical particle of mass m moving in one dimension under the influence of a potential $V(x)$. Near to a critical point x_0 of V, where $V'(x_0)=0$, we have the Taylor expansion

$$V(x) \;=\; V(x_0)+\tfrac{1}{2}V''(x_0)(x-x_0)^2+O\left((x-x_0)^3\right)\ . \tag{4.1}$$

If $V''(x_0)>0$ then x_0 is a *local minimum* of the potential. Without loss of generality let us choose coordinates where the critical point x_0 is at the origin. Then to lowest order the potential near to $x_0=0$ is

$$V(x) \;=\; V(0)+\frac{1}{2}m\omega^2x^2\ , \tag{4.2}$$

where we have defined $\omega^2=V''(0)/m$. The dynamics is not affected by the value of $V(0)$, since an additive constant drops out of the force $F=-V'$. This leads to:

Definition The *harmonic oscillator potential* for angular frequency ω is

$$V(x) \;=\; \frac{1}{2}m\omega^2x^2\ . \tag{4.3}$$

The above analysis shows that *any* system near to a point of stable equilibrium is described by this potential. Classically, we have the force $F = -V' = -m\omega^2 x$, so that Newton's second law reads

$$m\ddot{x} = -m\omega^2 x . \tag{4.4}$$

This has general solution $x(t) = A\cos\omega t + B\sin\omega t$, called *simple harmonic motion.* A classical particle in the potential (4.3) thus oscillates around the minimum with angular frequency ω. See Figure 8.

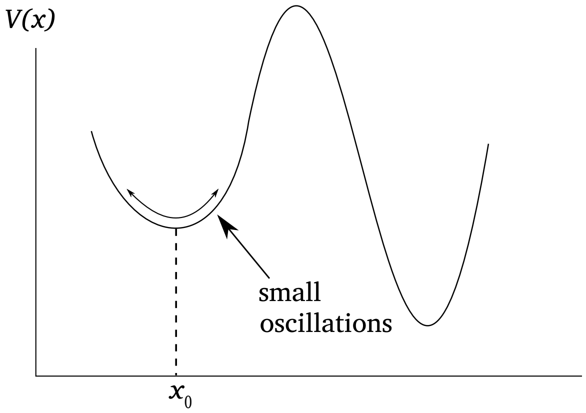

Figure 8: A particle performing small oscillations around a point of stable equilibrium of any potential $V(x)$ is described to lowest order by a harmonic oscillator.

We now turn to the corresponding quantum mechanical problem. The stationary state Schrödinger equation (2.15) of energy E is

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + \frac{1}{2}m\omega^2 x^2\psi = E\psi . \tag{4.5}$$

The first thing to do is to redefine variables so as to remove the various physical constants and work with *dimensionless variables*:

$$\epsilon \equiv \frac{2E}{\hbar\omega} , \qquad \xi \equiv \sqrt{\frac{m\omega}{\hbar}}x , \tag{4.6}$$

so that (4.5) becomes

$$-\frac{\mathrm{d}^2\chi}{\mathrm{d}\xi^2} + \xi^2\chi = \epsilon\chi , \tag{4.7}$$

where we have defined

$$\psi(x) \equiv \chi(\xi) = \chi\left(\sqrt{\frac{m\omega}{\hbar}}x\right) . \tag{4.8}$$

Although at first sight (4.7) looks like a fairly simple ODE, depending on a single constant ϵ (a dimensionless version of the energy E), in fact it is not so simple to solve. However, it is not difficult to spot that $\chi(\xi) = \mathrm{e}^{\mp\xi^2/2}$ solve (4.7) with $\epsilon = \pm 1$. To see this, we compute

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\left(\mathrm{e}^{\mp\xi^2/2}\right) = \mp\xi\,\mathrm{e}^{\mp\xi^2/2} , \qquad \frac{\mathrm{d}^2}{\mathrm{d}\xi^2}\left(\mathrm{e}^{\mp\xi^2/2}\right) = (\xi^2 \mp 1)\,\mathrm{e}^{\mp\xi^2/2} . \tag{4.9}$$

As discussed in section 3, we are only interested in *normalizable* solutions to the Schrödinger equation. For a stationary state, this means that

$$\int_{-\infty}^{\infty} |\psi(x)|^2 \, \mathrm{d}x \; < \; \infty \; . \tag{4.10}$$

Via the change of variable (4.6), (4.8), for $\chi(\xi) = \mathrm{e}^{\mp \xi^2/2}$ the left hand side of (4.10) is

$$\sqrt{\frac{\hbar}{m\omega}} \int_{-\infty}^{\infty} \mathrm{e}^{\mp \xi^2} \, \mathrm{d}\xi \; , \tag{4.11}$$

which is finite only for the minus sign, giving a Gaussian integral. Thus the solution $\chi(\xi) = \mathrm{e}^{+\xi^2/2}$, with $\epsilon = -1$, is *not* normalizable, and we thus discard it, while the solution $\chi(\xi) = \mathrm{e}^{-\xi^2/2}$, with $\epsilon = 1$, is normalizable.[8]

It turns out to be a good idea to again change variables by defining

$$\chi(\xi) \; \equiv \; f(\xi) \, \mathrm{e}^{-\xi^2/2} \; , \tag{4.12}$$

so that

$$\begin{aligned} \frac{\mathrm{d}\chi}{\mathrm{d}\xi} &= \left(\frac{\mathrm{d}f}{\mathrm{d}\xi} - \xi f \right) \mathrm{e}^{-\xi^2/2} \; , \\ \frac{\mathrm{d}^2\chi}{\mathrm{d}\xi^2} &= \left[\frac{\mathrm{d}^2 f}{\mathrm{d}\xi^2} - \xi \frac{\mathrm{d}f}{\mathrm{d}\xi} - f - \xi \left(\frac{\mathrm{d}f}{\mathrm{d}\xi} - \xi f \right) \right] \mathrm{e}^{-\xi^2/2} \; , \end{aligned} \tag{4.13}$$

and hence (4.7) becomes

$$\frac{\mathrm{d}^2 f}{\mathrm{d}\xi^2} - 2\xi \frac{\mathrm{d}f}{\mathrm{d}\xi} + (\epsilon - 1) \, f \; = \; 0 \; . \tag{4.14}$$

If anything, this looks worse than (4.7), so you might wonder why we bothered with (4.12)! The reasoning here is that for large values of $|\xi|$ the $\xi^2\chi$ term in (4.7) will dominate over the $\epsilon\chi$ term; thus for large $|\xi|$ one expects solutions for different ϵ to have the same behaviour to leading order. This is why we have written $\chi(\xi)$ in (4.12) as the above normalizable $\epsilon = +1$ solution times some other function $f(\xi)$.

It is still not clear how to solve (4.14) exactly, so as usual for this type of equation we try a power series solution.[9] Thus we write

$$f(\xi) \; = \; \sum_{k=0}^{\infty} a_k \, \xi^k \; , \tag{4.15}$$

and compute

$$\xi \frac{\mathrm{d}f}{\mathrm{d}\xi} \; = \; \sum_{k=0}^{\infty} k \, a_k \, \xi^k \; , \tag{4.16}$$

---

[8] Getting ahead of ourselves a little, we shall see later that this normalizable solution we have "spotted" is in fact the *ground state* wave function, with lowest energy $E = \frac{1}{2}\hbar\omega$ (setting $\epsilon = 1$ in (4.6)).

[9] ***** This is a topic that could have been included in the "core" Part A Differential Equations course, but has been left to the DE2 option. The point $\xi = 0$ is an *ordinary point* of (4.14), and then *Fuchs' theorem* guarantees that any solution may be expressed as a convergent Taylor series (4.15). The interested reader is referred to appendix A. Having said this, most quantum mechanics textbooks suppress these details.

and

$$\frac{\mathrm{d}^2 f}{\mathrm{d}\xi^2} \;=\; \sum_{k=0}^{\infty}(k+1)(k+2)\, a_{k+2}\, \xi^k \; . \tag{4.17}$$

Notice that the last expression involves a relabelling of the original sum. By substituting (4.15), (4.16) and (4.17) into (4.14) we obtain

$$\sum_{k=0}^{\infty} \left[(k+1)(k+2)\, a_{k+2} - 2k\, a_k + (\epsilon - 1)\, a_k\right] \xi^k \;=\; 0 \; . \tag{4.18}$$

The coefficient of every power of ξ must be separately zero, so we obtain the recurrence relation

$$a_{k+2} \;=\; \frac{2k+1-\epsilon}{(k+1)(k+2)}\, a_k \; . \tag{4.19}$$

The even and odd powers are then decoupled, giving rise to the two linearly independent series solutions

$$\begin{aligned} f_{\text{even}}(\xi) \;&=\; a_0 \left[1 + \frac{(1-\epsilon)}{2!}\xi^2 + \frac{(5-\epsilon)(1-\epsilon)}{4!}\xi^4 + \cdots \right] \; , \\ f_{\text{odd}}(\xi) \;&=\; a_1 \left[\xi + \frac{(3-\epsilon)}{3!}\xi^3 + \frac{(7-\epsilon)(3-\epsilon)}{5!}\xi^5 + \cdots \right] \; . \end{aligned} \tag{4.20}$$

Definition In one dimension a stationary state wave function satisfying $\psi(-x) = \pm\,\psi(x)$ is said to describe an $\left\{ \begin{array}{c} \textit{even} \\ \textit{odd} \end{array} \right\}$ *parity state.*

Recalling equations (4.6), (4.8) and (4.12), the total stationary state wave functions are given by $f_{\text{even/odd}}(\xi)\, \mathrm{e}^{-\xi^2/2}$, which then describe even and odd parity states, respectively.

We must now determine when these solutions are normalizable. It turns out that this is the case if and only if the series in (4.20) *terminate*. We shall not provide a full, rigorous proof of this here, but instead sketch the proof. This is purely for reasons of time. A more detailed treatment may be found in the book by Hannabuss.

Notice first that either of the series in (4.20) terminating is equivalent to the statement that $f(\xi)$ is a *polynomial*, which in turn is equivalent to the statement that $a_{n+2} = 0$ for some integer $n \geq 0$, since then the recurrence relation (4.19) implies that $a_k = 0$ for all $k \geq n+2$. Note that a_0 and a_1 are necessarily non-zero, otherwise $f_{\text{even/odd}}(\xi) \equiv 0$, respectively.

Suppose that either series in (4.20) does *not* terminate. Then all the coefficients are non-zero, and the ratio

$$\frac{a_{k+2}}{a_k} \sim \frac{2}{k} \; , \qquad \text{as } k \to \infty \; , \tag{4.21}$$

where the $\sim$ symbol here by definition means that the ratio $\left(\frac{a_{k+2}}{a_k}\right) / \left(\frac{2}{k}\right) \to 1$ as $k \to \infty$. Compare this asymptotic behaviour of the power series coefficients with that of the function e^{ξ^2}:

$$\mathrm{e}^{\xi^2} \;=\; \sum_{q=0}^{\infty} \frac{1}{q!}\, \xi^{2q} \;=\; \sum_{k=0}^{\infty} b_k\, \xi^k \; , \tag{4.22}$$

where we have defined the coefficients

$$b_k = \begin{cases} \frac{1}{q!} , & k = 2q , \\ 0 , & k = 2q+1 . \end{cases} \tag{4.23}$$

For $k = 2q$ even we then have

$$\frac{b_{k+2}}{b_k} = \frac{\left(\frac{k}{2}\right)!}{\left(\frac{k+2}{2}\right)!} = \frac{2}{k+2} \sim \frac{2}{k} , \qquad \text{as } k \to \infty . \tag{4.24}$$

Comparing to (4.21), we thus see that $f_{\text{even}}(\xi)$ and e^{ξ^2} have the same asymptotic expansion. Hence the total stationary state wave function is

$$\chi(\xi) = f(\xi)\,\mathrm{e}^{-\xi^2/2} \sim \mathrm{e}^{\xi^2/2} , \tag{4.25}$$

and thus $\psi(x) = \chi(\xi)$ is not normalizable.[10] A similar argument applies to $f_{\text{odd}}(\xi)$ and the function $\xi\,\mathrm{e}^{\xi^2}$. Thus for a normalizable solution the series in (4.20) must terminate.

If $n \geq 0$ is the least integer for which $a_{n+2} = 0$ then the recurrence relation (4.19) implies

$$2n + 1 - \epsilon = 0 . \tag{4.26}$$

Recalling the definition $\epsilon = 2E/\hbar\omega$ in terms of the energy E in (4.6) then gives

$$E = E_n = \left(n + \tfrac{1}{2}\right)\hbar\omega . \tag{4.27}$$

The quantum harmonic oscillator energies are hence labelled by the quantum number $n \in \mathbb{Z}_{\geq 0}$; compare to (2.23) for the particle in a box (and notice there that we instead defined n so that $n \in \mathbb{Z}_{>0}$). Also note that our initial solution with $f \equiv 1$ and $\epsilon = +1$ is in fact the *ground state*, with $n = 0$. Reverting back to the original spatial coordinate x via (4.6) and (4.8), the ground state wave function is hence

$$\psi_0(x) = a_0\,\mathrm{e}^{-m\omega x^2/2\hbar} , \tag{4.28}$$

where a_0 is a normalization constant, with corresponding *ground state energy* $E_0 = \frac{1}{2}\hbar\omega$. The normalized ground state wave function may be obtained by imposing

$$1 = \int_{-\infty}^{\infty} |\psi_0(x)|^2 \mathrm{d}x = |a_0|^2 \int_{-\infty}^{\infty} \mathrm{e}^{-m\omega x^2/\hbar}\,\mathrm{d}x . \tag{4.29}$$

The Gaussian integral is that for a normal distribution of variance $\sigma^2 = \hbar/2m\omega$, and thus via the standard result for this integral we have

$$|a_0|^2 = \sqrt{\frac{m\omega}{\pi\hbar}} , \tag{4.30}$$

leading to the full, normalized time-dependent ground state wave function

$$\Psi_0(x,t) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \mathrm{e}^{-(m\omega x^2 + \mathrm{i}\,\hbar\omega t)/2\hbar} . \tag{4.31}$$

---

[10] It is the last step in (4.25) that needs a little more work to make rigorous.

More generally the stationary state wave functions are $\chi_n(\xi) = f_n(\xi)\,\mathrm{e}^{-\xi^2/2}$, where $f_n(\xi)$ is an even/odd polynomial in $\xi = \sqrt{\frac{m\omega}{\hbar}}x$ of degree n, for n even/odd, respectively. The polynomials f_n may be determined explicitly by setting $\epsilon - 1 = 2n$ in the recurrence relation (4.19). Appropriately normalized, $f_n(\xi) \equiv H_n(\xi)$ is called the *nth Hermite polynomial.* The first few polynomials are given in Table 1, with the corresponding wave functions shown in Figure 9.

n	$H_n(\xi)$
0	1
1	2ξ
2	$4\xi^2 - 2$
3	$8\xi^3 - 12\xi$

Table 1: The first four Hermite polynomials.

We may summarize by stating

Theorem 4.1 *The energies of the one-dimensional quantum harmonic oscillator of angular frequency ω are*

$$E = E_n = \left(n + \tfrac{1}{2}\right)\hbar\omega \ , \tag{4.32}$$

for $n \in \mathbb{Z}_{\geq 0}$ a non-negative integer. The corresponding normalized *stationary state wave functions are*

$$\psi_n(x) = \frac{1}{\sqrt{2^n n!}}\left(\frac{m\omega}{\pi\hbar}\right)^{1/4} H_n(\xi)\,\mathrm{e}^{-\xi^2/2} \ , \tag{4.33}$$

where $\xi = \sqrt{m\omega/\hbar}\,x$ and H_n is the nth *Hermite polynomial, which may be written in closed form as*

$$H_n(\xi) = \mathrm{e}^{\xi^2/2}\left(\xi - \frac{\mathrm{d}}{\mathrm{d}\xi}\right)^n \mathrm{e}^{-\xi^2/2} \, . \tag{4.34}$$

Note that we have not yet derived the normalization factor given in (4.33) (apart from for the ground state $n = 0$ in (4.31)), nor the formula (4.34) for the Hermite polynomials $H_n(\xi)$, although it is straightforward to check the first few polynomials in (4.34) agree with Table 1. Rather than prove these results here, we will instead derive them in section 7 when we solve the quantum harmonic oscillator using algebraic methods.

As stressed at the beginning of this section, the harmonic oscillator is ubiquitous in physics, and it is therefore straightforward to test these results experimentally. For example, in a diatomic molecule the two atoms perform small oscillations of some frequency ω. One can then measure the frequency of absorbed/emitted photons, finding agreement with (4.32).

As for the particle in a box, it is interesting to compare the classical and quantum harmonic oscillators. We focus here on the following new feature. In classical mechanics a particle of

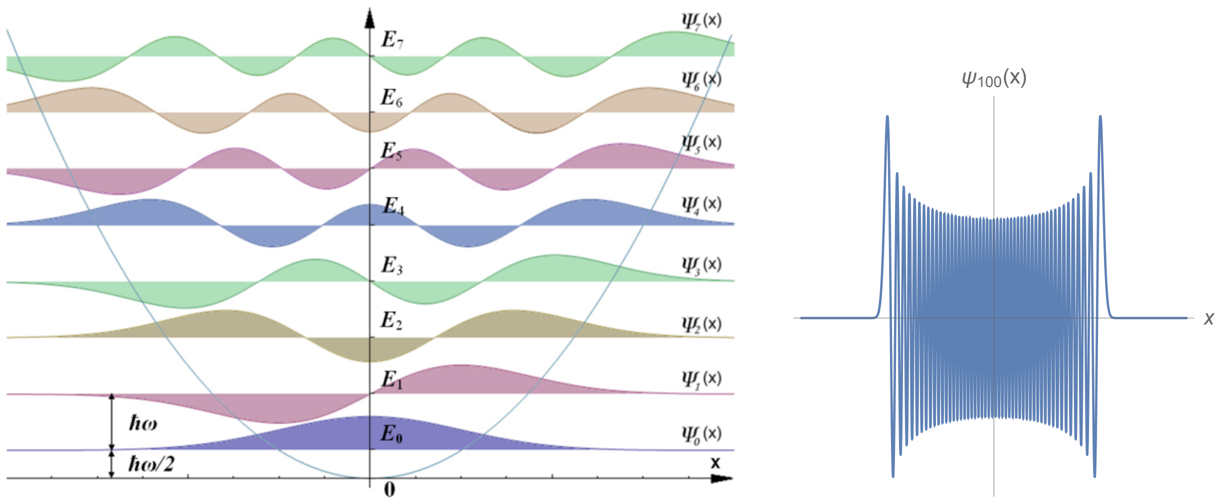

Figure 9: The ground state, the first seven *(credit: AllenMcC. wikipedia)* and the 100th excited state wave functions of the harmonic oscillator.

conserved energy E cannot enter a region of space where $E < V(\mathbf{x})$, as follows from the simple observation that

$$E = \frac{\mathbf{p}^2}{2m} + V(\mathbf{x}) \geq V(\mathbf{x}) . \tag{4.35}$$

The set $\{\mathbf{x} \in \mathbb{R}^3 \mid E < V(\mathbf{x})\}$ is called the *classically forbidden region*, for fixed particle energy E. For example, the ground state of the quantum harmonic oscillator has energy $E_0 = \frac{1}{2}\hbar\omega$, so the corresponding classically forbidden region for this energy is $\{\frac{1}{2}\hbar\omega < \frac{1}{2}m\omega^2 x^2\}$, which is the same as the set $\{|\xi| > 1\}$. The *quantum* probability of finding the particle in this region is

$$\int_{|x|>\sqrt{\hbar/m\omega}} |\Psi_0(x,t)|^2 \, \mathrm{d}x = \frac{1}{\sqrt{\pi}} \int_{|\xi|>1} \mathrm{e}^{-\xi^2} \, \mathrm{d}\xi \simeq 0.157 . \tag{4.36}$$

Here we have used (4.31) the the harmonic oscillator ground state wave function. We thus see that there is a non-zero probability of finding the particle in a region where *classically* it is *impossible* to find the particle! This is what underlies the phenomenon of *quantum tunnelling*, although we shall not pursue this further in this course.[11] We remark that we did not see the particle leaking out of the box in section 2.4, however we would encounter it if instead we put a particle in a potential well with finite height V_0 (with the box corresponding to $V_0 \to \infty$).

4.2 Higher-dimensional oscillators

Having discussed the one-dimensional oscillator, it is straightforward to extend this to oscillators in higher dimensions.

Let us begin by considering a quantum harmonic oscillator in two dimensions, with potential

$$V(x,y) = \frac{1}{2}m\left(\omega_1^2 x^2 + \omega_2^2 y^2\right) . \tag{4.37}$$

---

[11] ***** In keeping with the correspondence principle, for the analogous computation with energy E_n one finds the probability of finding the particle in the classically forbidden region tends to zero as $n \to \infty$, as demonstrated by the right hand side of Figure 9.

The corresponding stationary state Schrödinger equation (2.12) is

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}\right)+V(x,y)\psi \;=\; E\psi\ . \tag{4.38}$$

As for the particle in a box, this may be solved by separation of variables. One writes $\psi(x,y) = X(x)Y(y)$, so that (4.38) separates into

$$\begin{aligned}
-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2X}{\mathrm{d}x^2}+\frac{1}{2}m\omega_1^2x^2X &= E_1\,X\ ,\\
-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2Y}{\mathrm{d}y^2}+\frac{1}{2}m\omega_2^2y^2Y &= E_2\,Y\ ,
\end{aligned}\tag{4.39}$$

where $E_1 + E_2 = E$. We thus have two decoupled one-dimensional oscillators, and it follows from the previous section that

$$E \;=\; E_{n_1,n_2} \;=\; \left(n_1+\tfrac{1}{2}\right)\hbar\omega_1+\left(n_2+\tfrac{1}{2}\right)\hbar\omega_2\ , \tag{4.40}$$

where the quantum numbers $n_1, n_2 \in \mathbb{Z}_{\geq 0}$. The corresponding normalized stationary state wave functions are products

$$\begin{aligned}
\psi_{n_1,n_2}(x,y) &= \psi_{n_1}(x)\psi_{n_2}(y) \qquad\qquad (4.41)\\
&= \frac{1}{\sqrt{2^{n_1+n_2}n_1!n_2!}}\left(\frac{m^2\omega_1\omega_2}{\pi^2\hbar^2}\right)^{1/4} H_{n_1}\left(\sqrt{\frac{m\omega_1}{\hbar}}x\right) H_{n_2}\left(\sqrt{\frac{m\omega_2}{\hbar}}y\right)\mathrm{e}^{-m(\omega_1x^2+\omega_2y^2)/2\hbar}\ ,
\end{aligned}$$

where ψ_n denotes the normalized stationary state wave function for a one-dimensional harmonic oscillator.

Exercise Check that the above statements follow from Theorem 4.1. In particular, how is the normalizable/normalized condition related to the condition in one dimension?

Of course the quadratic form appearing in (4.37) is rather special, in that it is diagonal. One can treat more general quadratic forms by first changing to *normal coordinates*, *i.e.* one first diagonalizes V by an *orthogonal transformation*. Let us illustrate with an example:

Example Consider the oscillator potential

$$V(x,y) \;=\; m\omega^2(x^2+xy+y^2) \;=\; \frac{1}{2}m\omega^2\begin{pmatrix} x & y\end{pmatrix}\begin{pmatrix} 2 & 1\\ 1 & 2\end{pmatrix}\begin{pmatrix} x\\ y\end{pmatrix}\ . \tag{4.42}$$

The matrix here has eigenvalues 3 and 1, so there exists an orthogonal transformation to coordinates u, v with corresponding new potential

$$\begin{aligned}
\widetilde{V}(u,v) &= \frac{1}{2}m\omega^2\begin{pmatrix} u & v\end{pmatrix}\begin{pmatrix} 3 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix} u\\ v\end{pmatrix}\\
&= \frac{1}{2}m\left(3\omega^2u^2+\omega^2v^2\right)\ ,
\end{aligned}\tag{4.43}$$

which is now of the form (4.37).[12] The orthogonality of the transformation ensures that the Laplacian/kinetic term in the Schrödinger equation is invariant, so

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \;=\; \frac{\partial^2}{\partial u^2} + \frac{\partial^2}{\partial v^2} \; . \tag{4.44}$$

Since any quadratic form for V is described by a *symmetric matrix*, we may always diagonalize the potential by an orthogonal transformation, thus reducing the problem to a diagonal form for V, as in (4.37). These remarks apply in any dimension: first change to normal coordinates, then separate variables one at a time. The Schrödinger equation for *any* quadratic potential then reduces to a decoupled set of one-dimensional oscillators, so that the total energy = sum of one-dimensional energies, and the total wave functions = product of one-dimensional wave functions.

We conclude this section with another example of degenerate energy levels:

Example (Degeneracy) Consider the original two-dimensional oscillator potential (4.37) with $\omega_1 = \omega_2 = \omega$. Then

$$E \;=\; E_n \;=\; (n+1)\hbar\omega \; , \tag{4.45}$$

where $n = n_1 + n_2$. The ground state is $n_1 = n_2 = 0$, which is the unique state of energy $\hbar\omega$. However, more generally at level n there are $n+1$ linearly independent wave functions with energy E_n, given by taking (n_1, n_2) to be $(n, 0)$, $(n-1, 1)$, …, $(1, n-1)$, $(0, n)$. Thus E_n has degeneracy $n + 1$.

* The harmonic oscillator is ubiquitous in many applications. While we emphasized that the the photon is massless and therefore not governed by the Schrödinger equation, it turns out that in "quantizing" the classical electromagnetic field one effectively finds an infinite set of quantum harmonic oscillators! This is another reason why the quantum harmonic oscillator is so important.

5 The mathematical structure of quantum theory

The aim in this section is to present a general mathematical formalism for quantum theory, along with *postulates* for interpreting this formalism physically. We have already discussed the measurement of position and energy, and in order to generalize this we first distill some of the mathematical ingredients we have seen so far into a more abstract language. Up to this point you might think that quantum theory is about solving differential equations. Fundamentally though it is not: quantum mechanics is really linear algebra, and this point of view is also crucial for giving a general account of physical observables and measurement. We shall also see that this more abstract formalism leads to elegant derivations of some of the results derived earlier.

---

[12] Explicitly, $u = \frac{1}{\sqrt{2}}(x+y)$, $v = \frac{1}{\sqrt{2}}(-x+y)$, or more geometrically $\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \cos\frac{\pi}{4} & \sin\frac{\pi}{4} \\ -\sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$, so that the orthogonal transformation is a rotation by $\pi/4$.

5.1 States

In section 2.1 we said that a particle is described by a wave function $\Psi(\mathbf{x},t)$ satisfying the Schrödinger equation, with a given solution describing the *state* of the particle. We have also used the fact that the Schrödinger equation is linear, so that solutions can be thought of as living in a (often infinite-dimensional) vector space. The probabilistic interpretation of the wave function requires it to be normalized (3.4). We begin by rephrasing all this more abstractly.

Henceforth we will use the notation ψ for a wave function, suppressing the dependence on space, and also often on time t. Focusing on wave functions on the x-axis for simplicity (with obvious generalizations to $\mathbb{R}^3$), suppose that ψ and ϕ are normalizable. Then we may define the *complex inner product*

$$\langle\phi|\psi\rangle \equiv \int_{-\infty}^{\infty} \overline{\phi(x)}\psi(x)\,\mathrm{d}x \in \mathbb{C} . \tag{5.1}$$

The fact that this integral is finite follows from the *Schwarz inequality* for integrals

$$\left|\int_a^b \overline{\phi(x)}\psi(x)\,\mathrm{d}x\right|^2 \leq \int_a^b |\phi(x)|^2\,\mathrm{d}x \int_a^b |\psi(x)|^2\,\mathrm{d}x , \tag{5.2}$$

with normalizability implying that the integrals on the right hand side are finite on taking $a \to -\infty$, $b \to \infty$. In terms of the inner product (5.1) the *square norm* is then

$$\|\psi\|^2 \equiv \langle\psi|\psi\rangle , \tag{5.3}$$

with the norm $\|\psi\| = \sqrt{\|\psi\|^2}$ given by taking the positive square root, and a wave function is then *normalized* if $\|\psi\|^2 = 1$. The following properties are immediate from (5.1):

$$\begin{aligned}
&\text{(i)} \quad \overline{\langle\phi|\psi\rangle} = \langle\psi|\phi\rangle , \\
&\text{(ii)} \quad \langle\phi|\alpha_1\psi_1 + \alpha_2\psi_2\rangle = \alpha_1\langle\phi|\psi_1\rangle + \alpha_2\langle\phi|\psi_2\rangle , \quad \text{for all } \alpha_1,\alpha_2 \in \mathbb{C} , \\
&\text{(iii)} \quad \langle\beta_1\phi_1 + \beta_2\phi_2|\psi\rangle = \bar{\beta}_1\langle\phi_1|\psi\rangle + \bar{\beta}_2\langle\phi_2|\psi\rangle , \quad \text{for all } \beta_1,\beta_2 \in \mathbb{C} , \\
&\text{(iv)} \quad \|\psi\|^2 \geq 0 .
\end{aligned} \tag{5.4}$$

The inner product is thus linear in the second entry, but *conjugate*-linear in the first entry. Notice that (iii) follows from (i) and (ii). Moreover, recall that we required wave functions to be *continuous* functions, and a standard argument shows that in this case $\|\psi\|^2 = 0$ if and only if $\psi = 0$ is identically zero:[13]

$$\text{(v)} \quad \|\psi\|^2 = 0 \text{ if and only if } \psi = 0 . \tag{5.5}$$

Definition A complex vector space $\mathcal{H}$, equipped with a complex inner product $\langle\phi|\psi\rangle \in \mathbb{C}$ between vectors $\phi,\psi \in \mathcal{H}$ satisfying the above properties (i)–(v), is called a *complex inner product space*.

---

[13] The integral of a non-negative continuous function (such as $|\psi(x)|^2$) is zero iff the function is identically zero.

Definition / postulate The *states* of a quantum system are elements of a complex inner product space $\mathcal{H}$. Proportional vectors represent the same state.

We have sketched above that the space of continuous complex-valued normalizable functions on $\mathbb{R}$ (normalizable wave functions) is an inner product space. But there are also interesting quantum systems where $\mathcal{H}$ is finite-dimensional.[14] Notice that for proportional vectors we have $\phi = \alpha\psi$, with $\alpha \in \mathbb{C}$. On the other hand, for applications to physics we must normalize the states, and if ψ is normalized then $\|\phi\|^2 = |\alpha|^2$, and so ϕ is then normalized if and only if $\alpha = \mathrm{e}^{\mathrm{i}\varphi}$ is a constant phase. We already asserted in section 3.1 that wave functions differing by a constant phase describe physically equivalent states.

Making everything in this section rigorous requires a significant amount of analysis, some of which can be found in the Part B courses on functional analysis. We will largely gloss over many of the subtleties, only making the occasional comment. However, to make sense of certain operations it is worth pointing out that one needs the following:

Definition A *complete* complex inner product space is called a *Hilbert space*. More precisely, the states of a quantum system are then elements of a Hilbert space.

* This is required for certain limits to exist, for example infinite sums like (2.13) that we have already encountered. Here *completeness* is as defined in the Part A Metric Spaces and Complex Analysis course: every Cauchy sequence converges, so that if $\{\psi_n\}$ is a sequence such that for any $\varepsilon > 0$ there is $N \in \mathbb{N}$ with $\|\psi_m - \psi_n\| < \varepsilon$ for all $m, n > N$, then there is a limit $\psi_\infty \in \mathcal{H}$ with $\|\psi_n - \psi_\infty\| \to 0$ as $n \to \infty$. Every finite-dimensional complex inner product space is complete, but this is not true more generally. However, an incomplete inner product space admits a unique "completion", and this includes the space of complex-valued functions we started with above. We will not dwell on any of these details in what follows.

5.2 Observables

We begin with some more mathematical definitions:

Definition A map $A : \mathcal{H} \to \mathcal{H}$ on a complex vector space $\mathcal{H}$ satisfying

$$A(\alpha_1\psi_1 + \alpha_2\psi_2) \;=\; \alpha_1 A\psi_1 + \alpha_2 A\psi_2 \;, \tag{5.6}$$

for all $\alpha_i \in \mathbb{C}$, $\psi_i \in \mathcal{H}$, $i = 1, 2$, is called a *linear operator*. If $\mathcal{H}$ is a complex inner product space then the *adjoint* A^* of A by definition satisfies[15]

$$\langle A^*\phi|\psi\rangle \;=\; \langle\phi|A\psi\rangle \;, \tag{5.7}$$

for all $\phi, \psi \in \mathcal{H}$. If $A^* = A$, then A is called *self-adjoint*.

---

[14] Finite-dimensional complex inner product spaces are studied in the Part A Linear Algebra course.

[15] A^* is often denoted by $A^\dagger$ in the physics literature.

We then have the following remarkable postulate of quantum theory:

Definition / postulate The *observables* of a quantum system are given by self-adjoint linear operators A on the space of states $\mathcal{H}$.

By *observable*, we mean something that one can *measure*. This is a pretty abstract definition of something one of your physics friends might measure in their lab! As this section progresses we shall see in more detail what this means, and why it makes sense. Let's give the most important examples, namely the position and momentum operators, acting on wave functions:

Definition The *position operator* $\mathbf{X}$ in three dimensions has components X_i, $i = 1, 2, 3$, that are defined on wave functions $\psi \in \mathcal{H}$ via

$$(X_i\psi)(\mathbf{x}) \;=\; x_i\psi(\mathbf{x}) \;. \tag{5.8}$$

In the notation here the operator X_i maps the function ψ to the function $X_i\psi$, where the latter function of $\mathbf{x}$ is given by the right hand side of (5.8). In this case the ith component of the position operator X_i simply multiplies the wave function by x_i. The *momentum operator* $\mathbf{P}$ has components P_i, $i = 1, 2, 3$, that are defined on differentiable wave functions ψ by

$$(P_i\psi)(\mathbf{x}) \;=\; -\mathrm{i}\hbar\frac{\partial\psi}{\partial x_i}(\mathbf{x}) \;. \tag{5.9}$$

We may write these more briefly in vector notation as

$$\mathbf{X}\psi \;=\; \mathbf{x}\psi \;, \qquad \mathbf{P}\psi \;=\; -\mathrm{i}\hbar\nabla\psi \;. \tag{5.10}$$

There are corresponding definitions for motion in one dimension, acting on wave functions $\psi(x)$:

$$X\psi \;=\; x\psi \;, \qquad P\psi \;=\; -\mathrm{i}\hbar\psi' \;, \tag{5.11}$$

where $\psi' \equiv \mathrm{d}\psi/\mathrm{d}x$.

Why is the momentum operator given by $\mathbf{P} = -\mathrm{i}\hbar\nabla$? Ultimately this is a postulate, confirmed by experiments, but we may give some justification by going back to the de Broglie relations (1.4). In particular, we began in section 2 by interpreting a plane wave as the wave function for a free particle, where recall that the second equation in (2.2) reads $\mathbf{p}\Psi = -\mathrm{i}\hbar\nabla\Psi$. We have then promoted $-\mathrm{i}\hbar\nabla = \mathbf{P}$ to a differential operator acting on wave functions, and identified this with the observable for momentum. Notice that in one dimension

$$\|X\psi\|^2 \;=\; \int_{-\infty}^{\infty} x^2|\psi(x)|^2\,\mathrm{d}x \;, \qquad \|P\psi\|^2 \;=\; \hbar^2\int_{-\infty}^{\infty} |\psi'(x)|^2\,\mathrm{d}x \;, \tag{5.12}$$

so that these operators can only be defined on wave functions $\psi \in \mathcal{H}$ for which the expressions in (5.12) are well-defined. Because of this, operators are typically defined only on certain subspaces of $\mathcal{H}$, where they make sense, but as mentioned earlier we won't worry about such technical details.[16]

---

[16] * The interested reader might consult the book by Hannabuss.

It is clear that the operators defined in (5.10) are linear, but we also need to show they are self-adjoint. Focusing again on one dimension, this is immediate for the position operator. Indeed any *real-valued* function $f(x)$ leads to a self-adjoint operator that we may call $f(X)$, where by definition $(f(X)\psi)(x) = f(x)\psi(x)$:

$$\langle \phi | f(X)\psi \rangle \;=\; \int_{-\infty}^{\infty} \overline{\phi(x)} f(x)\psi(x)\,\mathrm{d}x \;=\; \int_{-\infty}^{\infty} \overline{f(x)\phi(x)}\psi(x)\,\mathrm{d}x \;=\; \langle f(X)\phi|\psi\rangle \ . \tag{5.13}$$

For momentum we compute

$$\begin{aligned}\langle \phi | P\psi \rangle \;&=\; \int_{-\infty}^{\infty} \overline{\phi(x)}(-\mathrm{i}\hbar\psi'(x))\,\mathrm{d}x \\ &=\; \left[-\mathrm{i}\hbar\overline{\phi(x)}\psi(x)\right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \mathrm{i}\hbar\overline{\phi'(x)}\psi(x)\,\mathrm{d}x \\ &=\; \int_{-\infty}^{\infty} \overline{(-\mathrm{i}\hbar\phi'(x))}\psi(x)\,\mathrm{d}x \;=\; \langle P\phi|\psi\rangle \ .\end{aligned} \tag{5.14}$$

Here in the second line we have integrated by parts, and in the last line assumed that the wave functions $\phi(x)$, $\psi(x)$ tend to zero at infinity. Notice that in proving self-adjointness of $P = -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}$ the minus sign one encounters in integrating by parts is effectively cancelled by the minus sign that arises by taking the complex conjugate of the factor of $\mathrm{i} = \sqrt{-1}$.

In three dimensions the inner product is

$$\langle \phi|\psi\rangle \;\equiv\; \iiint_{\mathbb{R}^3} \overline{\phi(\mathbf{x})}\psi(\mathbf{x})\,\mathrm{d}^3x \ , \tag{5.15}$$

and one can similarly check that $P_i = -\mathrm{i}\hbar\frac{\partial}{\partial x_i}$ is self-adjoint for each $i = 1, 2, 3$, integrating by parts in either the $x_1 = x$, $x_2 = y$ or $x_3 = z$ integrals, respectively.

We have defined the position and momentum operators, but what about other dynamical quantities? All quantities of interest in classical mechanics are simple functions of position $\mathbf{x}$ and momentum $\mathbf{p}$, and we may build new self-adjoint linear operators using the following:

Lemma 5.1 *Let A, B be self-adjoint linear operators. Then so is*

(i) $A + B$, where $(A + B)\psi(\mathbf{x}) = A\psi(\mathbf{x}) + B\psi(\mathbf{x})$,

(ii) αA, where $\alpha \in \mathbb{R}$ is a real *constant,*

(iii) the composite operator AB, provided A and B commute*; that is $AB = BA$.*

Proof Here (i) and (ii) follow almost immediately from the definitions, while for (iii) we compute

$$\langle \phi|AB\psi\rangle \;=\; \langle A\phi|B\psi\rangle \;=\; \langle BA\phi|\psi\rangle \ . \tag{5.16}$$

Note here that we use self-adjointness for A in the first equality, and then for B in the second equality. This shows that the adjoint $(AB)^* = BA$, and so self-adjointness means $AB = BA$. ■

An important observable in classical mechanics is the energy (2.6)

$$E = \frac{|\mathbf{p}|^2}{2m} + V(\mathbf{x}) , \tag{5.17}$$

and a candidate for the corresponding *quantum observable* is then simply to replace $\mathbf{x} \to \mathbf{X}$ and $\mathbf{p} \to \mathbf{P} = -\mathrm{i}\hbar\nabla$. Using Lemma 5.1 the resulting operator will be self-adjoint, in particular since $|\mathbf{P}|^2 = P_1^2 + P_2^2 + P_3^2$, and P_i^2 is self-adjoint by (iii):

Definition The *Hamiltonian operator* for a particle of mass m moving in a potential V is

$$H \equiv \frac{|\mathbf{P}|^2}{2m} + V(\mathbf{X}) , \tag{5.18}$$

acting on suitably differentiable wave functions via

$$(H\psi)(\mathbf{x}) = -\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{x}) + V(\mathbf{x})\psi(\mathbf{x}) . \tag{5.19}$$

For motion in one dimension along the x-axis we similarly have

$$(H\psi)(x) = \left(\frac{P^2}{2m} + V(X)\right)\psi(x) = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2}(x) + V(x)\psi(x) . \tag{5.20}$$

Definition The *Schrödinger equation* (2.7) for time-dependent states $\psi(t)$ is

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(t) = H\psi(t) , \tag{5.21}$$

while a *stationary state* $\psi(t) = \psi\,\mathrm{e}^{-\mathrm{i}Et/\hbar}$ of energy E satisfies

$$H\psi = E\psi . \tag{5.22}$$

Notice here that $\psi(t) \in \mathcal{H}$ for each time t, and for wave functions we have $(\psi(t))(\mathbf{x}) = \Psi(\mathbf{x}, t)$, in our notation earlier. It is interesting to prove the analogous result to Proposition 3.2 in this new formalism. We compute

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi(t)|\psi(t)\rangle &= \left\langle \frac{\partial\psi(t)}{\partial t}\middle|\,\psi(t)\right\rangle + \left\langle\psi(t)\middle|\frac{\partial\psi(t)}{\partial t}\right\rangle = \left\langle -\frac{\mathrm{i}}{\hbar}H\psi(t)\middle|\,\psi(t)\right\rangle + \left\langle\psi(t)\middle| -\frac{\mathrm{i}}{\hbar}H\psi(t)\right\rangle \\ &= \frac{\mathrm{i}}{\hbar}\langle(H - H^*)\psi(t)|\psi(t)\rangle = 0 . \end{aligned} \tag{5.23}$$

Here in the first equality we have used the product rule, and in the second the Schrödinger equation (5.21). On the second line we have used the definition of the adjoint operator to H and linearity, and the last step uses self-adjointness $H = H^*$. Notice that we proved the latter by appealing to self-adjointness of the momentum operator, and the proof of this in (5.14) assumed the wave functions tend to zero at infinity, as in the proof of Proposition 3.2. However, once we have established self-adjointness, the proof in (5.23) is much more clean and elegant.[17]

---

[17] * Sometimes the self-adjointness of the Hamiltonian is a subtle question even for Hamiltonians of the form (5.18). The most famous example is the case with $V(\mathbf{x}) = \alpha/x^2$. A delightful treatment is provided in section 2 of Kaplan et al.: Conformality Lost, arXiv:0905.4752.

It is no accident that the Hamiltonian operator appears on the right hand side of the Schrödinger equation (5.21): if you look back at the discussion in section 2.1, we precisely made the replacement $\mathbf{p} \to \mathbf{P}$ in the classical energy (2.6) when motivating (2.7). The other dynamical quantity of interest in this course is *angular momentum* $\mathbf{x} \wedge \mathbf{p}$, that we study in detail in section 8. The process of replacing $\mathbf{x} \to \mathbf{X}$, $\mathbf{p} \to \mathbf{P} = -\mathrm{i}\hbar\nabla$ in classical expressions to obtain a corresponding quantum observable is generally called *canonical quantization*, but there is an important caveat. Consider the classical expression xp for motion in one dimension. Then notice that

$$(XP)\psi \; = \; X(-\mathrm{i}\hbar\psi') \; = \; -\mathrm{i}\hbar x\psi' \; , \tag{5.24}$$

but instead

$$(PX)\psi \; = \; P(x\psi) \; = \; -\mathrm{i}\hbar(x\psi)' \; = \; -\mathrm{i}\hbar x\psi' - \mathrm{i}\hbar\psi \; . \tag{5.25}$$

It follows that the relation

$$(XP - PX)\psi \; = \; \mathrm{i}\hbar\psi \tag{5.26}$$

holds for all (differentiable) wave functions $\psi = \psi(x)$, and in particular X and P *do not commute*. We shall come back to this in section 6.2, but for now note from Lemma 5.1 that neither XP nor PX is self-adjoint. However, there is a natural way to fix this: we may replace $xp \to \frac{1}{2}(XP+PX)$, which is self-adjoint. We give a (starred) example of this at the end of section 9.3.

5.3 Measurement

We have claimed that observables in quantum theory correspond to self-adjoint linear operators on a complex inner product space $\mathcal{H}$, and have given some examples, but what does this actually *mean*? We begin with the following important mathematical preliminaries:

Definition Given an *observable* A (self-adjoint linear operator on $\mathcal{H}$), then a state $\psi \in \mathcal{H}$ satisfying

$$A\psi \; = \; \alpha\psi \; , \tag{5.27}$$

where $\alpha \in \mathbb{C}$ is a constant, is called an *eigenstate*[18] of A, with *eigenvalue* α. The set of all eigenvalues of A is called its *spectrum*.

Proposition 5.2 *The eigenvalues of an observable A are* real*, and if ψ_1, ψ_2 are eigenstates with* distinct *eigenvalues $\alpha_1 \neq \alpha_2$, then*

$$\langle\psi_2|\psi_1\rangle \; = \; 0 \; . \tag{5.28}$$

That is, ψ_1 and ψ_2 are orthogonal*.*

---

[18] In mathematics one would usually called this an *eigenvector*, but in quantum theory we use the terminology *eigenstate*, since vectors $\psi \in \mathcal{H}$ are quantum states.

Proof To prove that the eigenvalue α in (5.27) is real, we compute

$$\alpha\langle\psi|\psi\rangle = \langle\psi|\alpha\psi\rangle = \langle\psi|A\psi\rangle = \langle A\psi|\psi\rangle = \langle\alpha\psi|\psi\rangle = \bar{\alpha}\langle\psi|\psi\rangle . \tag{5.29}$$

Here in the first equality we have used linearity in the second entry of the inner product, the middle equality uses self-adjointness of A, and the last equality uses conjugate-linearity in the first entry of the inner product. Since quantum states have $\langle\psi|\psi\rangle \neq 0$, we deduce that $\bar{\alpha} = \alpha$ is real. Similarly if $A\psi_i = \alpha_i\psi_i$, $i = 1, 2$, then

$$\alpha_1\langle\psi_2|\psi_1\rangle = \langle\psi_2|A\psi_1\rangle = \langle A\psi_2|\psi_1\rangle = \bar{\alpha}_2\langle\psi_2|\psi_1\rangle = \alpha_2\langle\psi_2|\psi_1\rangle . \tag{5.30}$$

In the last step we have used the fact that α_2 is real. Since $\alpha_1 \neq \alpha_2$ we thus deduce (5.28). ■

In Prelims Linear Algebra II you proved the Spectral Theorem for self-adjoint operators on *finite-dimensional real* inner product spaces. The generalization to finite-dimensional complex inner product spaces is straightforward, and is in Part A Linear Algebra:

Theorem 5.3 (*Spectral Theorem*) *If $A : \mathcal{H} \to \mathcal{H}$ is self-adjoint and $\mathcal{H}$ is a* finite-dimensional *complex inner product space, then there exists an orthonormal basis of eigenvectors for A.*

In quantum theory we are often interested in infinite-dimensional inner product spaces of wave functions. Proposition 5.2 is valid whether $\mathcal{H}$ is finite-dimensional or infinite-dimensional. Distinct eigenvalues automatically have orthogonal eigenstates, and we may always normalize an eigenstate to have unit norm. When a given eigenvalue has two or more linearly independent eigenstates (so that the eigenvalue is *degenerate*, as in section 2.6), one can use the *Gram-Schmidt procedure* to construct a set of *orthogonal* eigenstates within that eigenspace. We refer to the linear algebra courses for details in finite dimensions, although we won't need them in what follows. Unfortunately the proof of the Spectral Theorem you have seen does not generalize to infinite-dimensional spaces, although with appropriate conditions on A a version of the Spectral Theorem is true. The details are far too involved for us to go into here, so following Dirac we will instead state this as an axiom/postulate that further restricts the class of self-adjoint linear operators that can represent observables in quantum theory:

Axiom / postulate An observable A in quantum theory is required to have a *complete* set of eigenstates. That is, there is a set of orthonormal eigenstates $\{\psi_n\}$ of A, meaning $A\psi_n = \alpha_n\psi_n$, $\alpha_n \in \mathbb{R}$, and

$$\langle\psi_m|\psi_n\rangle = \delta_{mn} , \tag{5.31}$$

such that any $\psi \in \mathcal{H}$ can be written as a linear combination

$$\psi = \sum_n c_n\psi_n , \tag{5.32}$$

where $c_n \in \mathbb{C}$. Notice that when $\psi = \psi(t)$ depends on time t, $c_n = c_n(t)$ in the expansion (5.32).

* We are assuming that the set of eigenstates of A is *countable* above, so can be labelled by an integer n. A Hilbert space with a countable orthonormal basis is called *separable*, and it then follows that all orthonormal bases are countable. Completeness of $\mathcal{H}$ (in the sense of Cauchy sequences being convergent) implies that infinite sums such as (5.32) with $\sum_n |c_n| < \infty$ converge to an element of $\mathcal{H}$. This is why $\mathcal{H}$ should be a Hilbert space.

Notice that the coefficients c_n in the expansion (5.32) are given by

$$\langle \psi_m | \psi \rangle \;=\; \sum_n \langle \psi_m | c_n \psi_n \rangle \;=\; \sum_n c_n \langle \psi_m | \psi_n \rangle \;=\; c_m \;, \tag{5.33}$$

where in the first equality we have substituted (5.32), the second equality uses linearity of the inner product, and the final equality uses orthonormality (5.31).

Definition The dimension d of the eigenspace for a given eigenvalue is called the *degeneracy* of that eigenvalue. An eigenvalue is called *non-degenerate* if the degeneracy $d = 1$.

If we label the orthonormal basis of eigenvectors as $\{\psi_n\}$ and there are degenerate eigenvalues then the corresponding α_n are not all distinct, *i.e.* $\alpha_{n_1} = \alpha_{n_2}$ for some $n_1 \neq n_2$. In this case it is sometimes more convenient to label the *distinct eigenvalues* by α_n, and then (5.32) reads

$$\psi \;=\; \sum_n \sum_{i=1}^{d_n} c_{n,i} \psi_{n,i} \;, \tag{5.34}$$

where d_n = degeneracy of the eigenvalue α_n, precisely as in (2.32), $c_{n,i} \in \mathbb{C}$ and $\psi_{n,i}$ form an orthonormal basis for the eigenspace with eigenvalue α_n.

We are now in a position to state:

Postulate (Quantum measurement postulate) The possible outcomes of a measurement of an observable A are given by the *eigenvalues* $\{\alpha_n\}$ of A. If the system is in a *normalized* quantum state ψ given by (5.32), then the *probability* of obtaining the value α_n in a measurement, for a *non-degenerate* eigenvalue α_n, is

$$\mathbb{P}_\psi(\text{obtaining } \alpha_n) \;=\; |c_n|^2 \;=\; |\langle \psi_n | \psi \rangle|^2 \;\; . \tag{5.35}$$

If α_n is degenerate we instead use the expansion (5.34), and

$$\mathbb{P}_\psi(\text{obtaining } \alpha_n) \;=\; \sum_{i=1}^{d_n} |c_{n,i}|^2 \;=\; \sum_{i=1}^{d_n} |\langle \psi_{n,i} | \psi \rangle|^2 \;\; . \tag{5.36}$$

Obviously (5.35) is a special case of (5.36), but for simplicity we'll focus on non-degenerate eigenvalues in much of what follows. This postulate makes sense, since for a normalized state

$$1 \;=\; \langle \psi | \psi \rangle \;=\; \sum_{m,n} \bar{c}_m c_n \langle \psi_m | \psi_n \rangle \;=\; \sum_{m,n} \bar{c}_m c_n \delta_{mn} \;=\; \sum_n |c_n|^2 \;, \tag{5.37}$$

where we have used (5.32) and (5.31).

Although this section has been very abstract, you will probably have noticed that we have seen much of this before in our discussion of the measurement of energy in section 3.3. Let us spell this out in more detail. The stationary state Schrödinger equation (5.22) says that a stationary state is an eigenstate of the Hamiltonian operator $A = H$, with the energy E being the corresponding eigenvalue. When we solved the stationary state Schrödinger equation for the particle in a box in sections 2.4 and 2.5, and the harmonic oscillator in section 4, we were precisely finding the spectrum of the Hamiltonian operator for those quantum systems! In both cases these give a countably infinite set of energies $\{E_n\}$, labelled (naturally) by an integer n, and in one-dimensional motion on the x-axis all the eigenvalues were non-degenerate. We also found the corresponding normalized eigenstates. For example, for the particle in a box these are given by $\psi_n = \psi_n(x) = \sqrt{\frac{2}{a}}\sin\frac{n\pi x}{a}$ for $x \in [0,a]$ inside the box, and moreover since these are Fourier modes we know these are orthonormal (5.31) – see (3.26). Moreover, the statement of completeness (5.32) here is simply the statement that normalizable functions $f : [0,a] \to \mathbb{C}$ can be expanded as a Fourier series (3.25).

Similar remarks apply to the harmonic oscillator in section 4. In particular, we learn that the normalized wave functions (4.33), proportional to $H_n(\xi)\,\mathrm{e}^{-\xi^2/2}$ with $H_n(\xi)$ the Hermite polynomials, are in fact *orthonormal*, so that

$$\int_{-\infty}^{\infty} H_n(\xi)H_m(\xi)\,\mathrm{e}^{-\xi^2}\,\mathrm{d}\xi \;=\; 0\;, \qquad \text{for } m \neq n\;. \tag{5.38}$$

You can check this by brute force for the Hermite polynomials given in Table 1, but our discussion in this section guarantees (5.38) to be true! The Hamiltonian operator is also complete in this case: that is, any normalizable function ψ on $\mathbb{R}$ can be written as an expansion in the eigenstates of H:

$$\psi(\xi) \;=\; \sum_{n=0}^{\infty} c_n H_n(\xi)\,\mathrm{e}^{-\xi^2/2}\;. \tag{5.39}$$

In practice we will use the formalism in this section for energy and angular momentum in the remainder of the course, but what about position and linear momentum that we started with? The above formalism needs to be slightly modified in these cases. While readers may safely skip to section 5.4 for the purposes of exam preparation, this material is fundamental to quantum theory. In particular, we will see that the Born interpretation of section 3.1 that interprets $|\Psi(\mathbf{x},t)|^2$ as a probability density function for measurement of position is indeed a special case of the *quantum measurement postulate* given in this section. One technical reason why this is not in our syllabus is because we need the *Dirac delta function* in what follows. You met this very briefly at the end of the Prelims Multivariable Calculus course, but a full treatment will appear only in the Part A Integral Transforms option.

* Position and momentum states

For simplicity we will work in one dimension on the x-axis, although the generalization to dynamics in $\mathbb{R}^3$ follows straightforwardly. Consider first the position operator X. Formally, an eigenstate of this operator is a wave function $\psi(x)$ that satisfies

$$x\psi(x) = x_0\psi(x) . \tag{5.40}$$

Here x_0 is the eigenvalue, but the eigenstate equation (5.40) holds for all $x \in \mathbb{R}$. This immediately implies that $\psi(x) = 0$ for all $x \neq x_0$. If $\psi(x)$ is continuous this forces $\psi \equiv 0$, and the position operator doesn't have any eigenstates! Technically, this is true, and in fact there are good *physical* reasons for this that we discuss further in section 6.3. In particular, in an eigenstate of the position operator the particle is located *precisely* at the point x_0, and the Heisenberg uncertainty principle says this is not possible *in principle* (no matter how good your measuring equipment).

Mathematically, (5.40) is solved by the *Dirac delta function* $\psi(x) = \delta(x_0 - x)$.[19] Formally this is zero for $x \neq x_0$, but infinite at the point $x = x_0$. This is not a function in the usual sense, but can be understood rigorously as a *generalized function* or *distribution*.[20] The key property of the Dirac delta function that we need is

$$\int_{-\infty}^{\infty} f(x)\delta(x_0 - x)\,\mathrm{d}x = f(x_0) , \tag{5.41}$$

for any continuous function $f(x)$. We may then write any wave function as

$$\psi(x) = \int_{-\infty}^{\infty} \psi(y)\delta(x - y)\,\mathrm{d}y = \int_{-\infty}^{\infty} \psi(y)\psi_y(x)\,\mathrm{d}y . \tag{5.42}$$

This may be read as saying that $\psi_y(x) \equiv \delta(x - y)$ form a *complete* set of eigenstates of the position operator, where the states are here labelled by the *continuous* variable $y \in \mathbb{R}$. That is, (5.42) is analogous to (5.32), with the sum replaced by an integral, $\sum_n \to \int \mathrm{d}y$, the coefficients $c_n \to \psi(y)$, and the basis states $\psi_n \to \psi_y$. An important caveat is that the Dirac delta function is *not normalizable*, and as such ψ_y are not elements of the Hilbert space of normalizable functions. However, they do satisfy an analogous property:

$$\int_{-\infty}^{\infty} \overline{\psi_{y'}(x)}\psi_y(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} \delta(x - y')\delta(x - y)\,\mathrm{d}x = \delta(y - y') , \tag{5.43}$$

where we note that the Dirac delta function is real. Rather than the Kronecker delta symbol δ_{mn} on the right hand side of (5.31), we instead have the Dirac delta $\delta(y - y')$ on the right hand side of (5.43), which is the analogous mathematical object for continuous variables. The ψ_y are thus in a sense complete and orthonormal, and one might call them *generalized eigenstates*. However, we stress again that it doesn't make sense to set $y = y'$ in (5.43), and ψ_y is not normalizable. Finally, note that the quantum measurement postulate says that the probability of measuring the eigenvalue α_n for an observable A is given by $|c_n|^2$, where the wave function is expanded in the basis of eigenstates of A, $\psi = \sum_n c_n\psi_n$. The analogous expansion for the position operator

---

[19] Dirac introduced his delta function in his 1930 book *Principles of Quantum Mechanics*. The rigorous theory was developed later in the 1940s.

[20] See Part A Integral Transforms.

is (5.42), where the coefficient $c_n \to \psi(y)$, so that the probability *density* is $|\psi(y)|^2$. The Born interpretation given in (3.1) is thus a special case of the formalism in this section.

An eigenstate of the momentum operator $P = -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}$ satisfies

$$-\mathrm{i}\hbar\frac{\mathrm{d}\psi}{\mathrm{d}x} = p\psi \ , \tag{5.44}$$

where the eigenvalue p on the right hand side is the momentum. The solution is of course

$$\psi(x) = \psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}}\mathrm{e}^{\mathrm{i}px/\hbar} \ , \tag{5.45}$$

where we have set the overall multiplicative integration constant to $1/\sqrt{2\pi\hbar}$ (so that these will satisfy (5.47) below), and thus eigenstates ψ_p are again labelled by a continuous variable $p \in \mathbb{R}$.[21] If you take the Part A Integral Transforms option you will learn that any function $\psi(x)$ can be written as

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}\tilde{\psi}(p)\,\mathrm{e}^{\mathrm{i}px/\hbar}\,\mathrm{d}p = \int_{-\infty}^{\infty}\tilde{\psi}(p)\psi_p(x)\,\mathrm{d}p \ , \tag{5.46}$$

which is analogous to (5.32) and (5.42), and says the generalized eigenstates of the momentum operator $\psi_p(x) = (2\pi\hbar)^{-1/2}\mathrm{e}^{\mathrm{i}px/\hbar}$ are complete. They are also orthonormal in the sense of (5.43):

$$\int_{-\infty}^{\infty}\overline{\psi_{p'}(x)}\psi_p(x)\,\mathrm{d}x = \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}x(p-p')/\hbar}\,\mathrm{d}x = \delta(p-p') \ . \tag{5.47}$$

Again, one cannot set $p = p'$ in this formula, and we have already remarked that the plane wave is not normalizable, since $|\mathrm{e}^{\mathrm{i}xp/\hbar}| = 1$. The "coefficient" $\tilde{\psi}(p)$ in the expansion (5.46) is called the *Fourier transform* of $\psi(x)$, and the quantum measurement postulate thus says that $|\tilde{\psi}(p)|^2$ is the *probability density function* for measuring momentum. That is, the probability of measuring the momentum in some interval $[p_1, p_2]$ in a state ψ is

$$\mathbb{P}_\psi([p_1,p_2]) = \int_{p_1}^{p_2}|\tilde{\psi}(p)|^2\,\mathrm{d}p \ , \tag{5.48}$$

where (5.46) inverts to give

$$\tilde{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}\psi(x)\,\mathrm{e}^{-\mathrm{i}px/\hbar}\,\mathrm{d}x \ . \tag{5.49}$$

The definition (5.48) makes sense due to *Plancherel's Theorem*, which follows from (5.46), (5.47) and says $\int_{-\infty}^{\infty}|\psi(x)|^2\,\mathrm{d}x = \int_{-\infty}^{\infty}|\tilde{\psi}(p)|^2\,\mathrm{d}p$. There is much more to say, but this is starred material and we should move on.

We note that in the case of the particle on the circle problem P has a discrete spectrum and normalisable eigenstates. Indeed, because of the double degeneracy of energy eigenvalues, we can choose a new basis for stationary states given in (2.37) as follows

$$\psi_n(x) = \frac{1}{\sqrt{2\pi R}}\,e^{\mathrm{i}nx/R}, \qquad E = E_n = \frac{n^2\hbar^2}{2mR^2} \qquad (n \in \mathbb{Z})\,, \tag{5.50}$$

---

[21] In fact potentially here $p \in \mathbb{C}$, although what follows requires $p \in \mathbb{R}$, and physically momentum should be real!

where note that we get the same energy for n and $-n$. This basis of states is also the eigenbasis of P, as

$$P\psi_n(x) \;=\; p_n\psi_n(x)\,, \qquad p_n \;=\; \frac{\hbar n}{R}\,. \tag{5.51}$$

5.4 Collapse of the wave function

The quantum measurement postulate says that if we measure some physical quantity, represented by a self-adjoint linear operator A on $\mathcal{H}$, then we obtain one of the eigenvalues α_n of A. If we then make another measurement, *immediately* after the first, it's physically reasonable to suppose that we will obtain the value α_n again. But this implies that, after the first measurement, the wave function is in an eigenstate with the measured eigenvalue. One says the wave function has *collapsed onto the eigenspace*, or more simply *collapsed*:

Postulate (Collapse of the wave function) If one measures the observable A and obtains the *non-degenerate* eigenvalue α_n, then immediately after the measurement the quantum state of the system is the eigenstate $\psi_n \in \mathcal{H}$ with $A\psi_n = \alpha_n\psi_n$.

The wave function ψ before the measurement takes the general form (5.32), but after the measurement notice that all the information contained in this quantum state (via the coefficients c_n) is discarded (and indeed lost). This postulate feels somewhat unsatisfactory; we will comment on this in the Epilogue.

In the degenerate eigenvalue case this generalizes to:

Postulate (Collapse of the wave function, degenerate case) If the spectrum of an observable A has degenerate eigenvalues, then one should first expand the quantum state ψ as in (5.34). Immediately after measuring the eigenvalue α_n the quantum state of the system is then

$$\psi \;\to\; \sum_{i=1}^{d_n} c_{n,i}\psi_{n,i}\;. \tag{5.52}$$

Notice that this is *orthogonal projection* of ψ onto the eigenspace for α_n.[22]

5.5 Summary

We may now summarize time evolution in quantum theory.[23] It is a two-step process:

(i) The normalized quantum state of the system $\psi(t)$ evolves in time according to the Schrödinger equation (5.21). This is a *deterministic* process. Given *any* observable A we may write

$$\psi(t) \;=\; \sum_n c_n(t)\,\psi_n\;, \tag{5.53}$$

---

[22] This more general statement of the collapse postulate is sometimes known as *Lüder's postulate.*

[23] For simplicity we focus on the non-degenerate eigenvalue case.

where $\{\psi_n\}$ are the complete orthonormal eigenstates of A, and $c_n(t) \in \mathbb{C}$. It is often most convenient here to choose $A = H$, so that ψ_n are stationary states of energy E_n and $c_n(t) = c_n \mathrm{e}^{-\mathrm{i}E_n t/\hbar}$, with $c_n \in \mathbb{C}$ now *constant* in time.

(ii) Suppose at some time $t = t_0$ we measure an observable A. Then

$$\mathbb{P}_\psi(\text{obtaining eigenvalue } \alpha_n \text{ of } A) \;=\; |\langle \psi_n | \psi(t_0) \rangle|^2 \;, \tag{5.54}$$

where ψ_n is the eigenstate of A with eigenvalue α_n. Immediately after obtaining the value α_n the wave function collapses to $\psi(t_0) \to \psi_n$. This is evidently a *probabilistic* process. One now starts time evolution again via the Schrödinger equation, with the initial condition $\psi(t_0) = \psi_n$.

Notice that the probabilistic interpretation of the wave function requires us to normalize ψ immediately after the collapse. This is particularly relevant for the degenerate case formula (5.52), where the right hand side will not be normalized.

Example (Two-state system) Let us give an example of this process of time evolution and measurement for a simple two-state system in which $\mathcal{H} = \mathbb{C}^2$ with the usual complex inner product. This might seem like a somewhat artificial toy model, but in fact the example we give describes real physics.[24] We choose the Hamiltonian to be

$$H \;=\; \hbar\omega \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \;. \tag{5.55}$$

Notice that this is indeed a self-adjoint operator, as the matrix is *Hermitian*. The eigenvalues are $E = E_\pm = \pm\hbar\omega$, with corresponding normalized stationary states

$$\psi_+ \;=\; \begin{pmatrix} 1 \\ 0 \end{pmatrix} \;, \qquad \psi_- \;=\; \begin{pmatrix} 0 \\ 1 \end{pmatrix} \;, \tag{5.56}$$

so that $H\psi_\pm = E_\pm \psi_\pm$. The general solution (5.53) to the Schrödinger equation $\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\psi = H\psi$ is

$$\psi(t) \;=\; c_+\psi_+\,\mathrm{e}^{-\mathrm{i}\omega t} + c_-\psi_-\,\mathrm{e}^{\mathrm{i}\omega t} \;, \tag{5.57}$$

where $c_\pm \in \mathbb{C}$ are constants.

Suppose that the observable we want to measure is the self-adjoint operator (Hermitian matrix)

$$A \;=\; \frac{1}{2}\hbar \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \;. \tag{5.58}$$

---

[24] * The Hamiltonian (5.55) describes the electron spin in a magnetic field along the z-axis, and the observable A we measure in (5.58) is the spin angular momentum along the x-axis. Take a look at this example again after reading section 8.3.

Two-state systems can approximate the dynamics of quantum systems under certain circumstances: e.g. if the system has a large gap above the ground and first excited states and we are studying low energy configurations.

It is easy to check that this has eigenvalues $\pm\frac{1}{2}\hbar$, with corresponding normalized eigenstates

$$\phi_\pm = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ \pm 1 \end{pmatrix} . \tag{5.59}$$

That is, $\|\phi_\pm\|^2 = \overline{\phi_\pm}\cdot\phi_\pm = 1$, and $A\phi_\pm = \pm\frac{1}{2}\hbar\,\phi_\pm$. Suppose that at time $t = 0$ the eigenvalue $+\frac{1}{2}\hbar$ for A has just been measured. The wave function has then collapsed onto the corresponding eigenstate, so $\psi(0) = \phi_+$. Substituting this initial condition into (5.57) with $t = 0$ allows us to compute $c_+ = c_- = \frac{1}{\sqrt{2}}$, and hence

$$\psi(t) = \frac{1}{\sqrt{2}}\psi_+\,\mathrm{e}^{-\mathrm{i}\omega t} + \frac{1}{\sqrt{2}}\psi_-\,\mathrm{e}^{\mathrm{i}\omega t} = \frac{1}{\sqrt{2}}\begin{pmatrix} \mathrm{e}^{-\mathrm{i}\omega t} \\ \mathrm{e}^{\mathrm{i}\omega t} \end{pmatrix} . \tag{5.60}$$

Suppose that at time $t_0 > 0$ the observable A is again measured. The quantum measurement postulate then says that

$$\begin{aligned} \mathbb{P}_\psi(\text{obtaining eigenvalue } +\tfrac{1}{2}\hbar) &= \left|\overline{\phi_+}\cdot\psi(t_0)\right|^2 = \left|\frac{1}{2}\left(\mathrm{e}^{-\mathrm{i}\omega t_0} + \mathrm{e}^{\mathrm{i}\omega t_0}\right)\right|^2 = \cos^2\omega t_0 , \\ \mathbb{P}_\psi(\text{obtaining eigenvalue } -\tfrac{1}{2}\hbar) &= \left|\overline{\phi_-}\cdot\psi(t_0)\right|^2 = \left|\frac{1}{2}\left(\mathrm{e}^{-\mathrm{i}\omega t_0} - \mathrm{e}^{\mathrm{i}\omega t_0}\right)\right|^2 = \sin^2\omega t_0 . \end{aligned} \tag{5.61}$$

Notice that the probabilities sum to 1, as they should. If the eigenvalue $\pm\frac{1}{2}\hbar$ is measured, the wave function instantaneously collapses onto $\psi(t_0) \to \phi_\pm$, respectively, and one uses this as an initial condition at $t = t_0$ in the general solution (5.57) to the Schrödinger equation.

6 Statistical aspects of quantum theory

6.1 Expectation and dispersion

Having presented the general probabilistic interpretation of quantum mechanics in the previous section, it is natural to introduce some further notions from probability theory. At the end of section 3.1 we defined the *expectation value* of a function of position, and this generalizes to:

Definition The *expectation value of an operator A in a normalized quantum state ψ* is

$$\mathbb{E}_\psi(A) \equiv \langle\psi|A\psi\rangle . \tag{6.1}$$

We will usually be interested in applying this to *self-adjoint* linear operators, *i.e.* observables, but it is convenient to define (6.1) more generally.

When $A = f(X)$ is a function of the position operator X, for a wave function $\psi = \Psi(x,t)$ in one dimension the definition (6.1) reads

$$\mathbb{E}_\psi(f(X)) = \int_{-\infty}^{\infty} \overline{\Psi(x,t)}f(x)\Psi(x,t)\,\mathrm{d}x = \int_{-\infty}^{\infty} f(x)|\Psi(x,t)|^2\,\mathrm{d}x , \tag{6.2}$$

which coincides with the definition (3.15). More generally suppose that the self-adjoint linear operator A has a complete orthonormal basis of eigenstates ψ_n, with $A\psi_n = \alpha_n\psi_n$, and expand $\psi = \sum_n c_n\psi_n$ as in (5.32). Then we calculate

$$\mathbb{E}_\psi(A) \;=\; \langle\psi|A\psi\rangle \;=\; \sum_n c_n\langle\psi|A\psi_n\rangle \;=\; \sum_n \alpha_n c_n\langle\psi|\psi_n\rangle \;=\; \sum_n \alpha_n|c_n|^2 \; . \tag{6.3}$$

Here in the last equality we have used (5.33), so that $\langle\psi|\psi_n\rangle = \overline{\langle\psi_n|\psi\rangle} = \bar{c}_n$. The expression on the right hand side of (6.3) is precisely how one defines expected value in probability, summing the possible values α_n, weighted by the probability of obtaining that value.

Notice that physically $\mathbb{E}_\psi(A)$ will be the average (mean) value of the measurement of the observable A, taken over a set of systems that are prepared in the *same* quantum state ψ prior to the measurement. This is physically different from repeatedly measuring the observable A on a single system, because the wave function collapse in section 5.4 changes the state ψ after the measurement. In order to state some further properties of expectation value, we first introduce:

Definition

(i) The *identity operator* $\mathbb{1}$ satisfies $\mathbb{1}\psi = \psi$, for all $\psi \in \mathcal{H}$.

(ii) An operator A is said to be *non-negative* if $\langle\psi|A\psi\rangle \geq 0$, for all $\psi \in \mathcal{H}$.

The properties in the next Proposition then all follow immediately from the definitions above:

Proposition 6.1

(*i*) *Linearity:* $\mathbb{E}_\psi(\alpha A + \beta B) = \alpha\mathbb{E}_\psi(A) + \beta\mathbb{E}_\psi(B)$, *for all* $\alpha, \beta \in \mathbb{C}$.

(*ii*) $\mathbb{E}_\psi(\mathbb{1}) = 1$.

(*iii*) *For self-adjoint* A, $\mathbb{E}_\psi(A)$ *is* real.

(*iv*) *For a* non-negative *operator* A, $\mathbb{E}_\psi(A) \geq 0$.

To see (*iii*) note that

$$\overline{\mathbb{E}_\psi(A)} \;=\; \langle A\psi|\psi\rangle \;=\; \langle\psi|A\psi\rangle \;=\; \mathbb{E}_\psi(A) \; , \tag{6.4}$$

where the second equality uses self-adjointness of A.

Notice also that

$$\mathbb{E}_\psi\left((A - \mathbb{E}_\psi(A)\mathbb{1})^2\right) \;=\; \mathbb{E}_\psi(A^2) - 2(\mathbb{E}_\psi(A))^2 + (\mathbb{E}_\psi(A))^2 \;=\; \mathbb{E}_\psi(A^2) - (\mathbb{E}_\psi(A))^2 \; . \tag{6.5}$$

In quantum theory the *standard deviation* is known instead as *dispersion*:

Definition The *dispersion* of an observable A in a normalized quantum state ψ is

$$\Delta_\psi(A) \equiv \sqrt{\mathbb{E}_\psi\left((A-\mathbb{E}_\psi(A)\mathbb{1})^2\right)} = \sqrt{\mathbb{E}_\psi(A^2)-(\mathbb{E}_\psi(A))^2} \; . \tag{6.6}$$

We then have the following result:

Proposition 6.2 *The dispersion of an observable A in the normalized state ψ is zero if and only if ψ is an eigenstate of A. Moreover, in this case the associated eigenvalue is $\mathbb{E}_\psi(A)$.*

Proof Suppose that ψ is an eigenstate of A with eigenvalue α, so $A\psi = \alpha\psi$. Then

$$\mathbb{E}_\psi(A) = \langle\psi|A\psi\rangle = \langle\psi|\alpha\psi\rangle = \alpha \; . \tag{6.7}$$

We also have the identity

$$\begin{aligned} \|(A-\mathbb{E}_\psi(A)\mathbb{1})\psi\|^2 &= \langle(A-\mathbb{E}_\psi(A)\mathbb{1})\psi|(A-\mathbb{E}_\psi(A)\mathbb{1})\psi\rangle \\ &= \langle\psi|(A-\mathbb{E}_\psi(A)\mathbb{1})^2\psi\rangle \\ &= \Delta_\psi(A)^2 \; . \end{aligned} \tag{6.8}$$

Here the second line follows from self-adjointness of $A-\mathbb{E}_\psi(A)\mathbb{1}$, and the last line is the definition of dispersion squared (6.6). Thus $\Delta_\psi(A)=0$ if and only if $A\psi = \mathbb{E}_\psi(A)\psi$, that is if and only if ψ is an eigenstate of A with eigenvalue $\mathbb{E}_\psi(A)$. ■

From a probabilistic point of view, having zero dispersion/standard deviation means that one is *certain* to measure the value $\mathbb{E}_\psi(A)$. Proposition 6.2 is hence another way to see that eigenstates of A are states with a definite value of this observable.

Example Consider the particle in a box $[0,a]$ where the Hamiltonian is $H = P^2/2m$ inside the box. Recall this has normalized stationary state wave functions $\psi_n(x) = \sqrt{\frac{2}{a}}\sin\frac{n\pi x}{a}$ inside the box. We compute the expectation value of the momentum operator $P = -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}$ in a stationary state:

$$\mathbb{E}_{\psi_n}(P) = \int_0^a \overline{\psi_n(x)}(-\mathrm{i}\hbar)\psi_n'(x)\,\mathrm{d}x = -\mathrm{i}\hbar\frac{2}{a}\frac{n\pi}{a}\int_0^a \sin\frac{n\pi x}{a}\cos\frac{n\pi x}{a}\,\mathrm{d}x = 0 \; . \tag{6.9}$$

The expected value of the momentum is thus zero. However,

$$\mathbb{E}_{\psi_n}(P^2) = 2m\mathbb{E}_{\psi_n}(H) = 2mE_n \; , \tag{6.10}$$

where in the last step we have used that ψ_n satisfies $H\psi_n = E_n\psi_n$, by definition. Thus in this case the dispersion is

$$\Delta_{\psi_n}(P) = \sqrt{\mathbb{E}_{\psi_n}(P^2)} = \sqrt{2mE_n} = \frac{n\pi\hbar}{a} \; . \tag{6.11}$$

where in the last equality we used the energies E_n given by (2.23).

6.2 Commutation relations

At the end of section 5.2 recall that we showed $(XP - PX)\psi = \mathrm{i}\hbar\psi$ for any quantum state $\psi \in \mathcal{H}$, where X and P are the position and momentum observables for one-dimensional motion – see equation (5.26). Similarly, in three dimensions where $P_i = -\mathrm{i}\hbar\frac{\partial}{\partial x_i}$ we compute

$$(X_i P_j)\psi \; = \; -\mathrm{i}\hbar x_i \frac{\partial \psi}{\partial x_j} \; , \tag{6.12}$$

and

$$(P_j X_i)\psi \; = \; P_j(x_i\psi) \; = \; -\mathrm{i}\hbar\frac{\partial}{\partial x_j}(x_i\psi) \; = \; -\mathrm{i}\hbar x_i \frac{\partial \psi}{\partial x_j} - \mathrm{i}\hbar\delta_{ij}\psi \; , \tag{6.13}$$

where we have used $\partial x_i/\partial x_j = \delta_{ij}$. It follows that $(X_iP_j - P_jX_i)\psi = \mathrm{i}\hbar\delta_{ij}\psi$ holds for all $\psi \in \mathcal{H}$, and we may hence identify the operators

$$(X_iP_j - P_jX_i) \; = \; \mathrm{i}\hbar\delta_{ij}\mathbb{1} \; . \tag{6.14}$$

Definition The *commutator* of two operators A and B is

$$[A, B] \; \equiv \; AB - BA \; . \tag{6.15}$$

We have thus shown

Proposition 6.3 (*The* canonical commutation relations) *In one dimension the position and momentum operators satisfy*

$$[X, P] \; = \; \mathrm{i}\hbar\mathbb{1} \; . \tag{6.16}$$

In three dimensions

$$[X_i, P_j] \; = \; \mathrm{i}\hbar\delta_{ij}\mathbb{1} \; , \qquad [X_i, X_j] \; = \; 0 \; = \; [P_i, P_j] \; . \tag{6.17}$$

Here $[X_i, X_j] = 0$ follows since multiplying by coordinates commute with each other, and similarly $[P_i, P_j] = 0$ follows since partial differentiations with respect to different coordinates also commute (acting on sufficiently well-behaved functions).

The non-commutativity of X and P has profound consequences in quantum mechanics. A lot of the interesting structure of quantum theory hinges on the canonical commutation relation (6.16). The latter is sometimes introduced as an *axiom*, with the form of the position operator $P = -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}$ then deduced from this.

Proposition 6.4 *For all operators A, B and C the commutator satisfies the following identities:*

(*i*) *Anti-symmetry:* $[A, B] = -[B, A]$.

(*ii*) *Linearity:* $[A, \beta B + \gamma C] = \beta[A, B] + \gamma[A, C]$, *for all* $\beta, \gamma \in \mathbb{C}$.

(*iii*) *The* Leibniz rule*:* $[A, BC] = B[A, C] + [A, B]C$*.* (*Similarly,* $[AB, C] = A[B, C] + [A, C]B$*.*)

(*iv*) *The* Jacobi identity*:* $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$*.*

(*v*) *If* $[A, B] = \mathrm{i}C$ *with* A*,* B *both self-adjoint, then* C *is also self-adjoint.*

Proof The first two are particularly simple consequences of the definition (6.15). We show (*iii*):

$$[A, BC] \;=\; ABC - BCA \;=\; (AB - BA)C + B(AC - CA) \;=\; B[A, C] + [A, B]C \;, \tag{6.18}$$

and similarly (*iv*):

$$\begin{aligned} [A, [B, C]] \;&=\; [A, BC] - [A, CB] \;=\; B[A, C] + [A, B]C - C[A, B] - [A, C]B \\ &=\; [B, [A, C]] + [[A, B], C] \;=\; -[B, [C, A]] - [C, [A, B]] \;. \end{aligned} \tag{6.19}$$

Here in the second equality we have used the Leibniz rule, and in the last step we have used anti-symmetry. Finally for (*v*) recall that in the proof of Lemma 5.1 we showed that for A, B self-adjoint operators we have $(AB)^* = BA$. Thus $[A, B]^* = (AB - BA)^* = BA - AB = -[A, B]$. The factor of i in $C = -\mathrm{i}[A, B]$ then ensures self-adjointness of C, noting that $(\mathrm{i}\mathbb{1})^* = -\mathrm{i}\mathbb{1}$. ■

Example Consider the Hamiltonian $H = P^2/2m + V(X)$ for motion in one dimension. Then

(i) $[X, H] = \mathrm{i}\hbar P/m$,

(ii) $[P, H] = -\mathrm{i}\hbar V'(X)$.

On Problem Sheet 3 you will explore the implications of these results for the correspondence principle.

Proof For (i) note that multiplying wave functions by functions of x commute, so $[X, V(X)] = 0$ and hence using linearity

$$[X, H] \;=\; [X, P^2/2m] \;=\; \frac{1}{2m}\left(P[X, P] + [X, P]P\right) \;=\; \frac{\mathrm{i}\hbar}{m}P \;, \tag{6.20}$$

where in the second equality we have used linearity (*ii*) and the Leibniz rule (*iii*) from Proposition 6.4, and the last equality uses the canonical commutation relation (6.16). For (ii) linearity and $[P, P^2] = 0$ implies that $[P, H] = [P, V(X)]$. The simplest way to compute this commutator is to substitute $P = -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}$ and act on a wave function $\psi(x)$:

$$\begin{aligned} [P, V(X)]\psi(x) \;&=\; -\mathrm{i}\hbar\left(\frac{\mathrm{d}}{\mathrm{d}x}(V(x)\psi(x)) - V(x)\frac{\mathrm{d}\psi}{\mathrm{d}x}\right) \;=\; -\mathrm{i}\hbar V'(x)\psi(x) \\ &=\; (-\mathrm{i}\hbar V'(X)\psi)(x) \;. \qquad \blacksquare \end{aligned} \tag{6.21}$$

6.3 Heisenberg's uncertainty principle

We begin this section by proving:

Proposition 6.5 *Let* A, B *be self-adjoint operators with* $[A, B] = \mathrm{i}C$*. Then for any* $s \in \mathbb{R}$ *and normalized* $\psi \in \mathcal{H}$ *we have*

$$\begin{aligned} &(i) \quad (A - \mathrm{i}sB)^*(A - \mathrm{i}sB) \;=\; A^2 + sC + s^2B^2 \ , \\ &(ii) \quad \|(A - \mathrm{i}sB)\psi\|^2 \;=\; \mathbb{E}_\psi(A^2) + s\mathbb{E}_\psi(C) + s^2\mathbb{E}_\psi(B^2) \ , \\ &(iii) \quad \mathbb{E}_\psi(A^2)\mathbb{E}_\psi(B^2) \;\geq\; \frac{1}{4}(\mathbb{E}_\psi(C))^2 \ , \end{aligned}$$

with equality in the last statement if and only if for some real s_0, $(A - \mathrm{i}s_0B)\psi = 0$.

Proof For (i) we compute

$$\begin{aligned} (A - \mathrm{i}sB)^*(A - \mathrm{i}sB) \;&=\; (A + \mathrm{i}sB)(A - \mathrm{i}sB) \;=\; A^2 - \mathrm{i}s(AB - BA) + s^2B^2 \\ &=\; A^2 + sC + s^2B^2 \ . \end{aligned} \tag{6.22}$$

For (ii) we have

$$\begin{aligned} \|(A - \mathrm{i}sB)\psi\|^2 \;&=\; \langle (A - \mathrm{i}sB)\psi | (A - \mathrm{i}sB)\psi \rangle \;=\; \langle \psi | (A - \mathrm{i}sB)^*(A - \mathrm{i}sB)\psi \rangle \\ &=\; \langle \psi | (A^2 + sC + s^2B^2)\psi \rangle \\ &=\; \mathbb{E}_\psi(A^2) + s\mathbb{E}_\psi(C) + s^2\mathbb{E}_\psi(B^2) \ . \end{aligned} \tag{6.23}$$

Finally for (iii) notice that the left hand side of (ii) is manifestly greater than or equal to zero. Thus $\mathbb{E}_\psi(A^2) + s\mathbb{E}_\psi(C) + s^2\mathbb{E}_\psi(B^2) \geq 0$ and this quadratic in s must then have discriminant

$$(\mathbb{E}_\psi(C))^2 - 4\mathbb{E}_\psi(A^2)\mathbb{E}_\psi(B^2) \;\leq\; 0 \ . \tag{6.24}$$

This is precisely the inequality in (iii). Finally, equality in (iii) means that the quadratic has a repeated real root $s = s_0$. But since this quadratic equation is

$$0 \;=\; \mathbb{E}_\psi(A^2) + s_0\mathbb{E}_\psi(C) + s_0^2\mathbb{E}_\psi(B^2) \;=\; \|(A - \mathrm{i}s_0B)\psi\|^2 \ , \tag{6.25}$$

from (ii), we see that s_0 is a root of the quadratic if and only if $(A - \mathrm{i}s_0B)\psi = 0$. ■

Corollary 6.6 (*Heisenberg's uncertainty principle*) *For normalized* ψ *we have*

$$\Delta_\psi(X)\Delta_\psi(P) \;\geq\; \frac{1}{2}\hbar \ . \tag{6.26}$$

Equality holds in (6.26) if and only if

$$\psi(x) \;=\; \exp\left[\frac{1}{2s_0\hbar}(x - \mu)^2 + \gamma\right] \ , \tag{6.27}$$

for some negative constant s_0*, and complex constants* $\mu, \gamma \in \mathbb{C}$*. In this case* ψ *is said to describe a* minimum uncertainty state*.*

Proof We set

$$A \equiv X - \mathbb{E}_\psi(X)\mathbb{1} \ , \qquad B \equiv P - \mathbb{E}_\psi(P)\mathbb{1} \ , \tag{6.28}$$

so that using linearity of the commutator

$$[A, B] = [X, P] - \mathbb{E}_\psi(P)[X, \mathbb{1}] - \mathbb{E}_\psi(X)([\mathbb{1}, P - \mathbb{E}_\psi(P)\mathbb{1}]) = [X, P] = \mathrm{i}\hbar\mathbb{1} \ , \tag{6.29}$$

and hence $C = -\mathrm{i}[A, B] = \hbar\mathbb{1}$. We then have

$$\mathbb{E}_\psi(A^2) = (\Delta_\psi(X))^2 \ , \qquad \mathbb{E}_\psi(B^2) = (\Delta_\psi(P))^2 \ , \tag{6.30}$$

and since $\mathbb{E}_\psi(C) = \hbar$ the inequality (*iii*) in Proposition 6.5 gives (6.26).

From Proposition 6.5 the case with equality means there is a real s_0 with $(A - \mathrm{i}s_0 B)\psi = 0$. Setting $\mu = \mathbb{E}_\psi(X) - \mathrm{i}s_0\mathbb{E}_\psi(P)$ and substituting from (6.28) with $P = -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}$ we have

$$0 = (A - \mathrm{i}s_0 B)\psi = \left(x - \mathrm{i}s_0(-\mathrm{i}\hbar)\frac{\mathrm{d}}{\mathrm{d}x} - \mu\right)\psi \ , \tag{6.31}$$

which rearranges to

$$\psi' = \frac{1}{s_0\hbar}(x - \mu)\psi \ . \tag{6.32}$$

We thus have a first order differential equation for $\psi(x)$, which immediately integrates to

$$\log\psi(x) = \frac{1}{2s_0\hbar}(x - \mu)^2 + \gamma \ , \tag{6.33}$$

where $\gamma \in \mathbb{C}$ is an integration constant. Exponentiating then proves (6.27). Notice that for this to be *normalizable* we need the real constant $s_0 < 0$, and (6.27) is then a Gaussian. ■

Notice that the proof of the Heisenberg uncertainty principle (6.26) follows directly from the canonical commutation relation $[X, P] = \mathrm{i}\hbar$. The principle says that in any state ψ there is a positive lower bound on the product of dispersions for position X and momentum P. The more precisely that the position is known, the smaller the value of $\Delta_\psi(X)$; but as the latter is decreased, $\Delta_\psi(P)$ must increase so that (6.26) is obeyed, and the momentum is known with less precision. Of course the same statements hold with X and P interchanged. Thus position and momentum cannot be simultaneously measured to arbitrary accuracy – this is the uncertainty.

In particular, a state ψ with definite position would be an eigenstate of the position operator, and Proposition 6.2 then says that $\Delta_\psi(X) = 0$. But then the uncertainty relation (6.26) formally says that $\Delta_\psi(P) = \infty$, and the momentum in this state is completely uncertain! This is why the position and momentum eigenstates discussed at the end of section 5.3 are not genuine states, and not normalizable. For the momentum eigenstate $\psi_p(x) \propto \mathrm{e}^{\mathrm{i}px/\hbar}$, which is an eigenstate of the momentum operator $P = -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}$ with eigenvalue p, we already commented that $|\mathrm{e}^{\mathrm{i}px/\hbar}|^2 = 1$ which is not integrable over the real line. Indeed, $|\psi_p(x)|^2$ being constant means that the particle is equally likely to be found anywhere, or in other words has completely uncertain position. If you knew the momentum of a particle *exactly*, then you would be equally likely to find it in your lab as you would be to find it down the back of your sofa, at the bottom of the Mariana Trench, or in the middle of the Sun.

Here we derived the uncertainty principle from quantum theory. Historically, this principle was used in combination with classical physics very creatively to explain many phenomena of quantum origin, including the stability and size of atoms. It guided the development of quantum theory and it has been satisfactorily embedded into its more complete framework.

Example (Minimum uncertainty state) Corollary 6.6 says that a minimum uncertainty state, by definition with $\Delta_\psi(X)\Delta_\psi(P) = \frac{1}{2}\hbar$, is described by a Gaussian wave function. We have seen an example of this already: recall that the ground state of the harmonic oscillator (4.28) has stationary state wave function

$$\psi_0(x) \;=\; a_0\,\mathrm{e}^{-m\omega x^2/2\hbar} \;. \tag{6.34}$$

In particular we identify $\mu = 0$ and $s_0 = -1/(m\omega)$.

Proposition 6.5 says that similar remarks apply to *any* pair of observables which do not commute. In fact we can easily prove

Corollary 6.7 (*Generalized uncertainty principle*) *Let A_1, A_2 be self-adjoint operators. Then for normalized $\psi \in \mathcal{H}$ we have*

$$\Delta_\psi(A_1)\Delta_\psi(A_2) \;\geq\; \frac{1}{2}\left|\mathbb{E}_\psi(-\mathrm{i}[A_1, A_2])\right| \;. \tag{6.35}$$

Proof The proof is similar to the proof of the Heisenberg uncertainty principle. We set

$$A \;\equiv\; A_1 - \mathbb{E}_\psi(A_1)\mathbb{1} \;, \qquad B \;\equiv\; A_2 - \mathbb{E}_\psi(A_2)\mathbb{1} \;, \tag{6.36}$$

so that $[A, B] = [A_1, A_2]$. The definition of dispersion in (6.6) then implies $\mathbb{E}_\psi(A^2) = (\Delta_\psi(A_1))^2$, $\mathbb{E}_\psi(B^2) = (\Delta_\psi(A_2))^2$, and taking the square root of the inequality (*iii*) in Proposition 6.5 gives (6.35). ■

7 The harmonic oscillator revisited

In this section we revisit the quantum harmonic oscillator of section 4, but now armed with our new, more powerful, formalism.

7.1 Raising and lowering operators

The Hamiltonian operator (5.20) for the one-dimensional quantum harmonic oscillator reads

$$H \;=\; \frac{P^2}{2m} + V(X) \;=\; \frac{1}{2m}P^2 + \frac{1}{2}m\omega^2 X^2 \;. \tag{7.1}$$

This leads to the stationary state Schrödinger equation (4.5) that we previously solved as a differential equation using a power series method. Instead we begin this section by introducing:

Definition The *raising* and *lowering operators* are defined by

$$a_\pm \; \equiv \; \frac{1}{\sqrt{2m\omega\hbar}} \left(\mp \mathrm{i} P + m\omega X \right) \; . \tag{7.2}$$

Since X and P are self-adjoint, we note that $a_+ = (a_-)^*$.

Lemma 7.1 *The Hamiltonian (7.1) may be written as*

$$H \; = \; \left(N + \tfrac{1}{2}\mathbb{1} \right) \hbar\omega \; , \tag{7.3}$$

where we have defined the self-adjoint number operator

$$N \; \equiv \; a_+ a_- \; . \tag{7.4}$$

Proof We compute

$$\begin{aligned} N \; = \; a_+ a_- \; &= \; \frac{1}{2m\omega\hbar} \left(-\mathrm{i}P + m\omega X \right) \left(\mathrm{i}P + m\omega X \right) \\ &= \; \frac{1}{2m\omega\hbar} \left(P^2 + m^2\omega^2 X^2 + \mathrm{i}m\omega [X, P] \right) \\ &= \; \frac{1}{2m\omega\hbar} \left(P^2 + m^2\omega^2 X^2 - m\omega\hbar\mathbb{1} \right) \; = \; \frac{1}{\hbar\omega} \left(\frac{1}{2m} P^2 + \frac{1}{2} m\omega^2 X^2 \right) - \frac{1}{2}\mathbb{1} \; , \end{aligned} \tag{7.5}$$

where we have used the canonical commutation relation (6.16). Rearranging gives (7.3). Finally,

$$N^* \; = \; (a_+ a_-)^* \; = \; a_-^* a_+^* \; = \; a_+ a_- \; = \; N \; . \blacksquare \tag{7.6}$$

Notice that we have essentially "factorized" the quadratic Hamiltonian (7.1) into a product of the two linear operators a_+, a_-, and that the extra factor of $\frac{1}{2}$ at the end of (7.5) comes from the fact that X and P (and hence a_+ and a_-) do not commute.

Proposition 7.2 *The raising and lowering operators and the number operator N satisfy*

(i) $[a_-, a_+] = \mathbb{1}$.

(ii) $[N, a_\pm] = \pm a_\pm$.

(iii) $\|a_- \psi\|^2 = \langle \psi | N \psi \rangle$.

(iv) $\|a_+ \psi\|^2 = \langle \psi | (N + \mathbb{1}) \psi \rangle$.

Here $\psi \in \mathcal{H}$ is any state.

Proof From the definition and using linearity of the commutator, for (i) we compute

$$[a_-, a_+] \;=\; \frac{1}{2m\omega\hbar}[\mathrm{i}P + m\omega X, -\mathrm{i}P + m\omega X] \;=\; \frac{1}{2m\omega\hbar} m\omega\left(-\mathrm{i}[X,P] + \mathrm{i}[P,X]\right) \;=\; \mathbb{1} \;, \tag{7.7}$$

the last equality using the canonical commutation relation (6.16). For (ii) we have

$$[N, a_+] \;=\; [a_+a_-, a_+] \;=\; a_+[a_-, a_+] + [a_+, a_+]a_- \;=\; a_+ \;, \tag{7.8}$$

where we have used the Leibniz rule (iii) in Proposition 6.4, and the last step uses (7.7). Similarly

$$[N, a_-] \;=\; [a_+a_-, a_-] \;=\; a_+[a_-, a_-] + [a_+, a_-]a_- \;=\; -a_- \;. \tag{7.9}$$

Next for (iii) we compute

$$\|a_-\psi\|^2 \;=\; \langle a_-\psi | a_-\psi\rangle \;=\; \langle \psi | a_+a_-\psi\rangle \;=\; \langle \psi | N\psi\rangle \;. \tag{7.10}$$

Finally (iv) is similar:

$$\|a_+\psi\|^2 \;=\; \langle a_+\psi | a_+\psi\rangle \;\;=\;\; \langle \psi | a_-a_+\psi\rangle \;=\; \langle \psi | ([a_-, a_+] + a_+a_-)\psi\rangle \;=\; \langle \psi | (\mathbb{1} + N)\psi\rangle \;, \tag{7.11}$$

using (i) in the last step. ■

Next we look at the eigenstates of the number operator N, which will be eigenstates of the Hamiltonian due to Lemma 7.1:

Lemma 7.3 *Suppose that* $N\psi = \lambda\psi$, *with* $\lambda \in \mathbb{R}$ *and* $\psi \neq 0 \in \mathcal{H}$. *Then*

(a) $a_\pm\psi$ *are also eigenstates of* N *(provided they are non-zero), with eigenvalues* $\lambda \pm 1$.

(b) $\lambda \geq 0$, *with* $\lambda = 0$ *if and only if* $a_-\psi = 0$ *(giving a* ground state *of* N*).*

Proof For (a) we compute

$$N(a_\pm\psi) \;=\; ([N, a_\pm] + a_\pm N)\psi \;=\; (\pm a_\pm + \lambda a_\pm)\psi \;=\; (\lambda \pm 1)a_\pm\psi \;, \tag{7.12}$$

where in the second equality we used (ii) from Proposition 7.2. Next for (b) for note that (iii) of Proposition 7.2 says that if $N\psi = \lambda\psi$ then $\lambda\|\psi\|^2 = \langle\psi|N\psi\rangle = \|a_-\psi\|^2 \geq 0$. This implies that $\lambda \geq 0$, with equality if and only if $a_-\psi = 0$. ■

Note that the raising and lowering operators $a_\pm$ precisely raise and lower the eigenvalue of an eigenstate of N by ±1, hence the name. By repeatedly acting with $a_\pm$ we deduce inductively that $a_\pm^n\psi$ will be eigenstates of N with eigenvalues $\lambda \pm n$, *provided* none of the states are zero. We examine this point more closely in the proof of the following Proposition:

Proposition 7.4 *The eigenvalues of* N *are* $n \in \mathbb{Z}_{\geq 0}$. *Moreover, if there is a* unique *ground state* ψ_0 *of* N, *up to normalization, then an eigenstate with eigenvalue* n *is proportional to* $a_+^n\psi_0$.

Proof Let ψ be an eigenstate of N with eigenvalue μ. Suppose that for all $n \geq 0$ we have $a_-^n\psi \neq 0 \in \mathcal{H}$, so that inductively $a_-^n\psi$ are eigenstates of N with eigenvalues $\mu - n$, respectively. For sufficiently large n we will have eigenvalue $\lambda = \mu - n < 0$, contradicting part (b) of Lemma 7.3. Hence there is a smallest $n \geq 0$ for which $a_-^n\psi \neq 0$, but $a_-^{n+1}\psi = 0$. Then $a_-^n\psi$ has eigenvalue $\lambda = \mu - n$, with $a_-(a_-^n\psi) = 0$, and part (b) of Lemma 7.3 implies that $\mu = n$. This proves that the spectrum of N is a subset of $\mathbb{Z}_{\geq 0}$.

Suppose now that ψ_0 is a ground state of N. Inductively we have that $a_+^n\psi_0$ are eigenstates of N with eigenvalues $n \in \mathbb{Z}_{\geq 0}$. These states are indeed all non-zero since from (iv) of Proposition 7.2

$$\|a_+^{n+1}\psi_0\|^2 \;=\; \langle a_+^n\psi_0|(N+\mathbb{1})a_+^n\psi_0\rangle \;=\; (n+1)\|a_+^n\psi_0\|^2 \;>\; 0\;, \tag{7.13}$$

where $\|\psi_0\|^2 > 0$ and then the inequality in (7.13) follows inductively on n. This proves that the spectrum of N is $\mathbb{Z}_{\geq 0}$ (we have found an eigenstate $a_+^n\psi_0$ with eigenvalue n, for all $n \geq 0$).

Finally, suppose that ψ is *any* state with eigenvalue $n > 1$. Then $a_-^n\psi$ has eigenvalue zero under N, but so does $a_-^n(a_+^n\psi_0)$. Both of these are hence ground states, and so if the ground state is unique up to normalization it follows that $a_-^n\psi = \kappa a_-^n a_+^n\psi_0$ for some $\kappa \in \mathbb{C}$. Defining $\tilde{\psi} \equiv \psi - \kappa a_+^n\psi_0$, it follows that $N\tilde{\psi} = n\tilde{\psi}$ has eigenvalue n, but $a_-^n\tilde{\psi} = 0$. We may then use property (iii) from Proposition 7.2 inductively to deduce

$$\begin{aligned}\|\tilde{\psi}\|^2 \;&=\; \frac{1}{n}\langle\tilde{\psi}|N\tilde{\psi}\rangle \;=\; \frac{1}{n}\|a_-\tilde{\psi}\|^2 \;=\; \frac{1}{n(n-1)}\langle a_-\tilde{\psi}|Na_-\tilde{\psi}\rangle \;=\; \frac{1}{n(n-1)}\|a_-^2\tilde{\psi}\|^2\\ &=\; \cdots \;=\; \frac{1}{n!}\|a_-^n\tilde{\psi}\|^2 \;=\; 0\;.\end{aligned} \tag{7.14}$$

So $\tilde{\psi} = 0$ and the eigenstates with eigenvalue n are unique up to normalization. ■

Notice that the proof of this result used only the *algebraic properties* of the raising and lowering operators, and *not* their original definition (7.2) in terms of position and momentum operators. However, we may use the latter to write the ground state condition as

$$0 \;=\; a_-\psi_0 \quad\Rightarrow\quad (\mathrm{i}P + m\omega X)\psi_0 \;=\; 0\;, \tag{7.15}$$

which substituting $P = -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}$ leads to the first order ODE

$$\hbar\psi_0' + m\omega x\psi_0 \;=\; 0\;. \tag{7.16}$$

This integrates to

$$\psi_0(x) \;=\; a_0\,\mathrm{e}^{-m\omega x^2/2\hbar}\;, \tag{7.17}$$

with $a_0 \in \mathbb{C}$ a complex integration constant, which is indeed the ground state wave function derived earlier in (4.28). Combining Lemma 7.1 with Proposition 7.4, we have thus proven

Corollary 7.5 *The energies of the one-dimensional quantum harmonic oscillator of frequency* ω *are* $E_n = \left(n + \frac{1}{2}\right)\hbar\omega$, *with corresponding stationary states* $a_+^n\psi_0$, *where* ψ_0 *is the ground state (7.17).*

7.2 Normalized states and wave functions

Suppose that the ground state ψ_0 is normalized, so $\|\psi_0\| = 1$. Recall this fixes the integration constant a_0 in (7.17) via (4.30), so that $a_0 = (m\omega/\pi\hbar)^{1/4}$. Then using (7.13) we inductively see that

$$\|a_+^n\psi_0\|^2 \;=\; n!\|\psi_0\|^2 \;=\; n!\ . \tag{7.18}$$

This proves

Proposition 7.6 *The* normalized *stationary states of the quantum harmonic oscillator are*

$$\psi_n \;\equiv\; \frac{1}{\sqrt{n!}}a_+^n\psi_0\ . \tag{7.19}$$

Moreover, these are then orthonormal

$$\langle\psi_m|\psi_n\rangle \;=\; \delta_{mn}\ . \tag{7.20}$$

Notice that eigenstates with different eigenvalues are orthogonal from Proposition 5.2, since H is self-adjoint.

* Hermite polynomials

Recall that in section 4.1 we introduced the variable

$$\xi \;\equiv\; \sqrt{\frac{m\omega}{\hbar}}x\ , \tag{7.21}$$

in terms of which the raising and lowering operators (7.2) are (substituting $P = -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}$)

$$a_\pm \;=\; \frac{1}{\sqrt{2m\omega\hbar}}\left(\mp\hbar\frac{\mathrm{d}}{\mathrm{d}x} + m\omega x\right) \;=\; \frac{1}{\sqrt{2}}\left(\mp\frac{\mathrm{d}}{\mathrm{d}\xi} + \xi\right)\ . \tag{7.22}$$

On the other hand the ground state wave function is

$$\psi_0 \;=\; a_0\,\mathrm{e}^{-\xi^2/2}\ , \tag{7.23}$$

so that from Proposition 7.6 the *normalized* stationary state wave functions are

$$\psi_n \;=\; \frac{1}{\sqrt{n!}}a_+^n\psi_0 \;=\; \frac{1}{\sqrt{2^n n!}}a_0\left(\xi - \frac{\mathrm{d}}{\mathrm{d}\xi}\right)^n\mathrm{e}^{-\xi^2/2}\ . \tag{7.24}$$

This is a polynomial in ξ times $\mathrm{e}^{-\xi^2/2}$, and to extract the polynomial piece we may simply multiply by $\mathrm{e}^{\xi^2/2}$, and define

$$H_n(\xi) \;\equiv\; \mathrm{e}^{\xi^2/2}\left(\xi - \frac{\mathrm{d}}{\mathrm{d}\xi}\right)^n\mathrm{e}^{-\xi^2/2}\ . \tag{7.25}$$

This is the expression (4.34) for the Hermite polynomials, and completes the proof of Theorem 4.1.

8 Angular momentum

In classical mechanics the angular momentum of a particle moving in three dimensions is the vector $\mathbf{L} = \mathbf{x} \wedge (m\mathbf{v}) = \mathbf{x} \wedge \mathbf{p}$. In section 9 we will study the hydrogen atom, consisting of an electron orbiting a proton under an inverse square law force of attraction. In the first year Dynamics course you learned that *conservation of angular momentum* plays an important role in solving this inverse square law force problem, and the same will be true in quantum theory. In this section we develop a general theory of angular momentum in quantum mechanics.

8.1 Angular momentum operators

As in section 5.2 we may obtain the quantum operator corresponding to $\mathbf{x} \wedge \mathbf{p}$ via the replacement $\mathbf{x} \to \mathbf{X}$, $\mathbf{p} \to \mathbf{P}$:

Definition The *angular momentum operator* $\mathbf{L} = (L_1, L_2, L_3)$ has components L_i, $i = 1, 2, 3$, defined by

$$L_1 \equiv X_2 P_3 - X_3 P_2 \ , \qquad L_2 \equiv X_3 P_1 - X_1 P_3 \ , \qquad L_3 \equiv X_1 P_2 - X_2 P_1 \ , \tag{8.1}$$

where $\mathbf{X}$ and $\mathbf{P}$ are the position and momentum operators. We may write this more succinctly by introducing the *Levi-Civita alternating symbol* ϵ_{ijk}, defined by

$$\epsilon_{ijk} \equiv \begin{cases} +1 & \text{if } ijk \text{ is an even permutation of } 123 \ , \\ -1 & \text{if } ijk \text{ is an odd permutation of } 123 \ , \\ 0 & \text{otherwise} \ . \end{cases} \tag{8.2}$$

Notice that $\epsilon_{ijk} = \epsilon_{kij} = \epsilon_{jki}$ (cyclically permuting the indices), while $\epsilon_{ijk} = -\epsilon_{ikj}$, and that the cross product of two vectors $\mathbf{A}$, $\mathbf{B}$ then has ith component $(\mathbf{A} \wedge \mathbf{B})_i = \sum_{j,k=1}^{3} \epsilon_{ijk} A_j B_k$. The operators (8.1) are then

$$L_i = \sum_{j,k=1}^{3} \epsilon_{ijk} X_j P_k \ . \tag{8.3}$$

Note that $\mathbf{L}$ is self-adjoint: using self-adjointness of $\mathbf{X}$ and $\mathbf{P}$ and the fact that for $j \neq k$ X_j and P_k commute, we have $(X_j P_k)^* = P_k X_j = X_j P_k$. This implies $L_i^* = L_i$.

Proposition 8.1 *The angular momentum operator satisfies the following commutation relations:*

$$(i) \quad [L_i, X_j] = \mathrm{i}\hbar \sum_{k=1}^{3} \epsilon_{ijk} X_k \ , \qquad (ii) \quad [L_i, P_j] = \mathrm{i}\hbar \sum_{k=1}^{3} \epsilon_{ijk} P_k \ ,$$

$$(iii) \quad [L_i, L_j] = \mathrm{i}\hbar \sum_{k=1}^{3} \epsilon_{ijk} L_k \ . \tag{8.4}$$

Proof Notice that on the left hand side of (i) we have a free index j on X_j, so we need to relabel the indices j, k that are summed over in the definition (8.3). Writing the latter instead as m, n, using linearity of the commutator we compute

$$\begin{aligned}
[L_i, X_j] &= \sum_{m,n=1}^{3} \epsilon_{imn}[X_m P_n, X_j] = \sum_{m,n=1}^{3} \epsilon_{imn} \left(X_m [P_n, X_j] + [X_m, X_j] P_n \right) \\
&= \sum_{m,n=1}^{3} \epsilon_{imn} \left(-\mathrm{i}\hbar X_m \delta_{nj} + 0 \right) = -\mathrm{i}\hbar \sum_{m=1}^{3} \epsilon_{imj} X_m \\
&= \mathrm{i}\hbar \sum_{m=1}^{3} \epsilon_{ijm} X_m \ . \qquad (8.5)
\end{aligned}$$

Here the second equality uses the Leibniz rule from Proposition 6.4, the third equality uses the canonical commutation relations (6.17), and in the final step we have used anti-symmetry of $\epsilon_{imj} = -\epsilon_{ijm}$. The proof of (ii) is very similar. Finally for (iii) we calculate

$$\begin{aligned}
[L_1, L_2] &= [L_1, X_3 P_1 - X_1 P_3] = X_3 [L_1, P_1] + [L_1, X_3] P_1 - X_1 [L_1, P_3] - [L_1, X_1] P_3 \\
&= 0 + \mathrm{i}\hbar (-X_2) P_1 - X_1 \mathrm{i}\hbar (-P_2) - 0 \\
&= \mathrm{i}\hbar (X_1 P_2 - X_2 P_1) = \mathrm{i}\hbar L_3 \ . \qquad (8.6)
\end{aligned}$$

Here the second equality uses the Leibniz rule, and the third equality second uses (i) and (ii) of (8.4). The other components of (iii) follow from cyclically permuting the indices.[25] ■

Definition Any self-adjoint vector operator $\mathbf{J}$, with components J_i, $i = 1, 2, 3$, satisfying

$$[J_i, J_j] = \mathrm{i}\hbar \sum_{k=1}^{3} \epsilon_{ijk} J_k \ , \qquad (8.7)$$

is called an *angular momentum operator*. The particular angular momentum operator $\mathbf{L}$ defined by (8.3) is then more precisely called the *orbital angular momentum*.

Much of what follows relies only on the commutation relations (8.7), so we shall for now work with a general angular momentum operator.

Proposition 8.2 *If* $\mathbf{A}$ *and* $\mathbf{B}$ *are operators satisfying* $[J_i, A_j] = \mathrm{i}\hbar \sum_{k=1}^{3} \epsilon_{ijk} A_k$ *and* $[J_i, B_j] = \mathrm{i}\hbar \sum_{k=1}^{3} \epsilon_{ijk} B_k$*, then*

$$[J_i, \mathbf{A} \cdot \mathbf{B}] = 0 \ , \qquad (8.8)$$

where $\mathbf{A} \cdot \mathbf{B} \equiv A_1 B_1 + A_2 B_2 + A_3 B_3$*.*

---

[25] Cyclically permuting the indices (123) → (231) in the definitions (8.1) permutes $(L_1, L_2, L_3) \to (L_2, L_3, L_1)$.

Proof We compute

$$[J_i, \mathbf{A}\cdot\mathbf{B}] \;=\; \sum_{j=1}^{3}[J_i, A_jB_j] \;=\; \sum_{j=1}^{3}(A_j[J_i,B_j] + [J_i,A_j]B_j)$$
$$= \mathrm{i}\hbar\sum_{j,k=1}^{3}\epsilon_{ijk}(A_jB_k + A_kB_j) \;=\; 0\;, \tag{8.9}$$

with the last equality holding since $\epsilon_{ijk} = -\epsilon_{ikj}$ is anti-symmetric. ■

Corollary 8.3 *It follows that* L_i *commutes with the operators* $\mathbf{X}\cdot\mathbf{X}$*,* $\mathbf{P}\cdot\mathbf{P}$*,* $\mathbf{X}\cdot\mathbf{P}$*, and* $\mathbf{L}\cdot\mathbf{L}$*.*

8.2 Raising and lowering operators

Our aim will be to find the eigenvalues and eigenstates for angular momentum operators, following the algebraic treatment of the harmonic oscillator in section 7. However, the different components J_i of angular momentum do not commute: $[J_1, J_2] = \mathrm{i}\hbar J_3$, and the generalized uncertainty principle (6.35) then says

$$\Delta_\psi(J_1)\Delta_\psi(J_2) \;\geq\; \frac{\hbar}{2}\left|\mathbb{E}_\psi(J_3)\right|\;. \tag{8.10}$$

In an eigenstate ψ of J_3 with non-zero eigenvalue the right hand side of (8.10) is positive, and hence ψ cannot simultaneously be an eigenstate of J_1 and J_2 from Proposition 6.2! We cannot then simultaneously measure the components of angular momentum to arbitrary precision.

In finite dimensions it is a general fact that given two *commuting* self-adjoint operators A, B on $\mathcal{H}$, so $[A,B]=0$, one can simultaneously diagonalize both, *i.e.* find a basis for $\mathcal{H}$ where the basis vectors are eigenstates for both A and B. A proof for general inner product spaces can be found in Appendix A1.3 of the book by Hannabuss, although this is definitely non-examinable. For angular momentum the individual components of J_i do not commute, but Proposition 8.2 suggests introducing the following:

Definition The *total angular momentum* is defined by

$$J^2 \;\equiv\; \mathbf{J}\cdot\mathbf{J} \;=\; J_1^2 + J_2^2 + J_3^2\;. \tag{8.11}$$

Proposition 8.2 implies

$$[J^2, J_i] \;=\; 0\;, \qquad i \;=\; 1,2,3\;, \tag{8.12}$$

and we can hope to look for simultaneous eigenstates of J^2 and *one* of the components J_i. It is conventional to take this to be J_3. In doing so it turns out to be convenient to introduce:

Definition The *raising* and *lowering operators* for angular momentum are

$$J_\pm \;\equiv\; J_1 \pm \mathrm{i}J_2\;. \tag{8.13}$$

Notice that $J_+ = (J_-)^*$, using self-adjointness of $\mathbf{J}$.

Proposition 8.4 *The raising and lowering operators satisfy*

- *(i)* $[J^2, J_\pm] = 0$.
- *(ii)* $J_\pm J_\mp = J^2 - J_3^2 \pm \hbar J_3$.
- *(iii)* $[J_+, J_-] = 2\hbar J_3$.
- *(iv)* $[J_3, J_\pm] = \pm\hbar J_\pm$.

Proof *(i)* is immediate from (8.12) and linearity. Next for *(ii)* we calculate

$$J_\pm J_\mp \;=\; (J_1 \pm \mathrm{i}J_2)(J_1 \mp \mathrm{i}J_2) \;=\; J_1^2 + J_2^2 \mp \mathrm{i}(J_1J_2 - J_2J_1) \;=\; J^2 - J_3^2 \pm \hbar J_3 \ , \tag{8.14}$$

the last equality using the definition of J^2 in (8.11) and the commutator (8.7). Then for *(iii)*

$$[J_+, J_-] \;=\; J_+J_- - J_-J_+ \;=\; 2\hbar J_3 \ , \tag{8.15}$$

follows from subtracting the two expressions in *(ii)*. Finally, *(iv)* is proven via

$$[J_3, J_\pm] \;=\; [J_3, J_1 \pm \mathrm{i}J_2] \;=\; \mathrm{i}\hbar J_2 \pm \hbar J_1 \;=\; \pm\hbar J_\pm \ . \qquad \blacksquare \tag{8.16}$$

8.3 Representations of angular momentum

We give a brief summary of the most important results from the representation theory of angular momentum, mostly without proof. A more complete discussion is found in the non-examinable appendix B. A more complete and satisfying discussion is given in the Part B7.3 course Further Quantum Theory.

There is a whole theory of this subject called *representation theory*, where instead of the angular momentum commutation relations one can study more general structures (called *Lie algebras*). There are Part B and Part C courses on these topics. Angular momentum is in some sense the simplest non-trivial example.

As for the harmonic oscillator in section 7, it is helpful to first establish some key properties of the action of the raising and lowering operators $J_\pm$ on states:

Proposition 8.5 *Let ψ be a common eigenstate of J^2 and J_3, satisfying*[26]

$$J^2\psi \;=\; \lambda\hbar^2\psi \ , \qquad J_3\psi \;=\; m\hbar\psi \ . \tag{8.17}$$

Then

- *(i)* $J^2 J_\pm\psi = \lambda\hbar^2 J_\pm\psi$.
- *(ii)* $J_3 J_\pm\psi = (m \pm 1)\hbar J_\pm\psi$.

---

[26] The factors of $\hbar$ in the eigenvalues $\lambda\hbar^2$ and $m\hbar$ are for later convenience.

Proof Part (*i*) of Proposition 8.4 says that $J^2 J_\pm = J_\pm J^2$, which applying to ψ immediately gives (*i*) of Proposition 8.5. Similarly for (*ii*), part (*iv*) of Proposition 8.4 implies

$$J_3 J_\pm \psi \ = \ (J_\pm J_3 \pm \hbar J_\pm)\,\psi \ = \ J_\pm(m\hbar \pm \hbar)\psi \ = \ (m \pm 1)\hbar J_\pm \psi \ . \qquad \blacksquare \tag{8.18}$$

Using this Proposition we can argue much as we did for the harmonic oscillator to prove:

Theorem 8.6 *The eigenvalues of J^2 have the form $j(j+1)\hbar^2$, where $j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \ldots$ takes non-negative half-integer values. For each choice of j, the eigenvalues of J_3 are $m\hbar$, where $m = -j, -j+1, -j+2, \ldots, j-1, j$. The degeneracy of each eigenvalue is the same as that of $j\hbar$.*

Proof For a proof see appendix B. ■

There we also show that in the absence of degeneracies an eigenspace $\mathcal{H}_j$ on which J^2 takes the eigenvalue $j(j+1)\hbar^2$ has dimension $2j+1$, with an orthonormal basis of the form $\{\psi_m \mid m = -j, -j+1, \ldots, j-1, j\}$ such that

$$\begin{aligned} J_3 \psi_m \ &= \ m\hbar\, \psi_m \ , \\ J_\pm \psi_m \ &= \ \sqrt{(j \mp m)(j \pm m + 1)}\hbar\, \psi_{m\pm 1} \ . \end{aligned} \tag{8.19}$$

Notice that in Proposition 8.5 we started by *assuming* we had a common eigenstate ψ of J^2 and J_3, but in the process of proving Theorem 8.6 and eq. (8.19) we have done much more: for each choice of $j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \ldots$, we have effectively *constructed* finite-dimensional inner product spaces $\mathcal{H}_j$ on which the angular momentum operators J_i act.

Let us look at this in a little more detail. For a fixed non-negative half-integer j, we have a complex vector space $\mathcal{H}_j$ of dimension $2j+1 \in \mathbb{N}$. There is a preferred basis $\{\psi_{-j}, \psi_{-j+1}, \ldots, \psi_{j-1}, \psi_j\}$ for this vector space in which the action of J_3 and $J_\pm \equiv J_1 \pm \mathrm{i} J_2$ on a basis vector ψ_m is defined by (8.19). In particular the basis elements are eigenstates of J_3, with the action of $J_\pm$ effectively moving "up and down" among the basis elements. Notice that $J_+\psi_j = 0$ and $J_-\psi_{-j} = 0$ follow from the second equation in (8.19). The J_i act linearly on $\mathcal{H}_j$ in the obvious way, defining

$$J_i \left(\sum_{m=-j}^{j} c_m\, \psi_m \right) \ \equiv \ \sum_{m=-j}^{j} c_m\, J_i\, \psi_m \ , \tag{8.20}$$

for arbitrary complex coefficients c_m. Moreover, we can check that this action of the J_i obeys the commutation relations (8.7). In the $J_\pm, J_3$ basis the commutation relations are equivalent to $[J_+, J_-] = 2\hbar J_3$, $[J_3, J_\pm] = \pm\hbar J_\pm$, and using (8.19) we can compute

$$\begin{aligned} J_+ J_- \psi_m \ &= \ \sqrt{(j+m)(j-m+1)}\hbar\, J_+\,(\psi_{m-1}) \ = \ (j+m)(j-m+1)\hbar^2\, \psi_m \ , \\ J_- J_+ \psi_m \ &= \ \sqrt{(j-m)(j+m+1)}\hbar\, J_-\,(\psi_{m+1}) \ = \ (j-m)(j+m+1)\hbar^2\, \psi_m \ , \end{aligned} \tag{8.21}$$

so that

$$[J_+, J_-]\psi_m \ = \ [(j+m)(j-m+1) - (j-m)(j+m+1)]\,\hbar^2\, \psi_m \ = \ 2m\hbar^2 \psi_m \ = \ 2\hbar J_3 \psi_m \ . \tag{8.22}$$

The relation $[J_+, J_-] = 2\hbar J_3$ then holds when acting on any state in $\mathcal{H}_j$. Similarly,

$$J_3 J_\pm \psi_m \;=\; (m\pm 1)\hbar\, J_\pm\psi_m \ , \qquad J_\pm J_3\psi_m \;=\; m\hbar\, J_\pm\psi_m \ , \tag{8.23}$$

so that $[J_3, J_\pm] = \pm\hbar J_\pm$ holds acting on any state.

Definition A vector space together with a linear action of the angular momentum operators obeying the commutation relations (8.7) is called a *representation* of angular momentum.

For *every* dimension $2j+1 \in \mathbb{N}$, we have thus constructed a representation of angular momentum. Moreover, notice that the operators J_i are also *self-adjoint* (being observables), where the inner product is such that the basis $\{\psi_m\}$ is orthonormal. When we write out the associated linear maps as matrices, this will mean those matrices are *Hermitian*. Let's look at these representation spaces for small values of j in a bit more detail:

Example ($\mathcal{H}_0$, called the *trivial representation*)

Here $j = 0$ so $\dim \mathcal{H}_0 = 2j+1 = 1$, $\mathcal{H}_0 = \mathbb{C}$, and the space is spanned by a single state ψ_0 with (from (8.19))

$$J_3\psi_0 \;=\; 0 \ , \qquad J_\pm\psi_0 \;=\; 0 \ , \qquad J^2\psi_0 \;=\; 0 \ . \tag{8.24}$$

This is called the *trivial representation*, for obvious reasons.

Example ($\mathcal{H}_{1/2}$, called the *spin representation*)

Here $j = \frac{1}{2}$, so $\dim \mathcal{H}_{1/2} = 2j+1 = 2$, with the allowed values of m being $m = \pm\frac{1}{2}$. We write $\psi_\pm \equiv \psi_{\pm 1/2}$ as the two basis vectors for $\mathcal{H}_{1/2} = \mathbb{C}^2$, so that

$$J_3\psi_\pm \;=\; \pm\frac{1}{2}\hbar\psi_\pm \ , \tag{8.25}$$

and

$$J_+\psi_+ \;=\; 0 \;=\; J_-\psi_- \ , \qquad J_+\psi_- \;=\; \hbar\psi_+ \ , \qquad J_-\psi_+ \;=\; \hbar\psi_- \ , \tag{8.26}$$

where the normalizations in the last two equations are determined using (8.19). Writing the orthonormal basis vectors $\psi_\pm$ as

$$\psi_+ \;=\; \begin{pmatrix} 1 \\ 0 \end{pmatrix} \ , \qquad \psi_- \;=\; \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ , \tag{8.27}$$

from equations (8.25) we can read off the corresponding matrices

$$J_3 \;=\; \frac{1}{2}\hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \ , \quad J_+ \;=\; \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \ , \quad J_- \;=\; \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \ . \tag{8.28}$$

Since from the definitions (8.13) $J_1 = \frac{1}{2}(J_+ + J_-)$ and $J_2 = -\frac{1}{2}\mathrm{i}(J_+ - J_-)$, we can also write down

$$J_1 = \frac{1}{2}\hbar \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \qquad J_2 = \frac{1}{2}\hbar \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} . \tag{8.29}$$

It's simple to check that these 2×2 matrices J_i are indeed self-adjoint (Hermitian), and obey the angular momentum commutation relations (8.7). For example, $J_1 J_2 - J_2 J_1 = \mathrm{i}\hbar J_3$, where the product means matrix multiplication.

* We remark that we have already encountered this structure in the guise of the two-state system in section 5.5. The J_i operators we constructed here together with $\mathbb{1}$ form a basis in the space of Hermitian operators acting on $\mathcal{H}_{1/2}$, hence are very useful in applications.

Definition The matrix representation of J_i given by (8.28), (8.29) is called the *spin representation* of angular momentum. The matrices

$$\sigma_1 \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad \sigma_2 \equiv \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} , \quad \sigma_3 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} , \tag{8.30}$$

are called the *Pauli matrices*. These are traceless and Hermitian, with $J_i = \frac{1}{2}\hbar\sigma_i$, $i = 1, 2, 3$, and correspondingly the Pauli matrices satisfy $[\sigma_i, \sigma_j] = 2\mathrm{i} \sum_{k=1}^{3} \epsilon_{ijk}\sigma_k$.

This spin representation is in fact directly relevant to matter particles such as the electron. In addition to having standard orbital angular momentum $\mathbf{J} = \mathbf{L}$, described in the next section, electrons have a kind of "internal" angular momentum called *spin* (often denoted $\mathbf{J} = \mathbf{S}$), where specifically $j = \frac{1}{2}$. This is a purely quantum mechanical notion, and doesn't have a classical analogue. In practice this means that the wave function $\psi = \Psi(\mathbf{r}, t)$ for an electron doesn't take values in $\mathbb{C}$, but rather in $\mathcal{H}_{1/2} = \mathbb{C}^2$. An electron in an eigenstate of J_3 then has spin state given by either ψ_+ or ψ_- in (8.27), which are called *spin up* and *spin down* states, respectively.

Of course there is nothing special about the z-axis direction in all of this, and J_3. It is straightforward to check that both J_1 and J_2 in (8.29) also have eigenvalues $\pm\frac{1}{2}\hbar$, with corresponding normalized eigenstates

$$\psi_{1,\pm} \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix} , \qquad \psi_{2,\pm} \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} \mp \mathrm{i} \\ 1 \end{pmatrix} . \tag{8.31}$$

The normalization condition here is $\|\psi_{i,\pm}\|^2 = \overline{\psi_{i,\pm}} \cdot \psi_{i,\pm} = 1$. Defining also $\psi_{3,\pm} \equiv \psi_\pm$, we thus have

$$J_i \psi_{i,\pm} = \pm\frac{1}{2}\hbar\, \psi_{i,\pm} , \qquad i = 1, 2, 3 . \tag{8.32}$$

For example, if the spin angular momentum of an electron is measured along the z-axis and the value $\frac{1}{2}\hbar$ is obtained, then the system is in the eigenstate $\psi_+ = \psi_{3,+}$ ("spin up" along the z-axis). If we now measure the spin angular momentum along the x-axis, the probability of obtaining the value $\frac{1}{2}\hbar$ ("spin up" along the x-axis) is by the quantum measurement postulate of section 5.3 given by

$$\mathbb{P} \,=\, \left|\overline{\psi_{3,+}} \cdot \psi_{1,+}\right|^2 \,=\, \left|\frac{1}{\sqrt{2}}\right|^2 \,=\, \frac{1}{2} \ . \tag{8.33}$$

* The *spin representation* won't be immediately relevant to our treatment of the electron in the hydrogen atom in section 9. However, in a background *magnetic field* the spin **S** directly enters the Hamiltonian, affecting the electron's dynamics. Spin is also important for describing *multiple* electrons, for example in all other atoms apart from hydrogen. Spin is studied further in the Part B7.3 course Further Quantum Theory.

Of course, we can go on. For example, for $j = 1$ with respect to the basis $\{\psi_{+1}, \psi_0, \psi_{-1}\}$ of $\mathcal{H}_1 = \mathbb{C}^3$ in eq. (8.19) one finds the 3×3 Hermitian matrices

$$J_1 \,=\, \frac{1}{\sqrt{2}}\hbar \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \ , \quad J_2 \,=\, \frac{1}{\sqrt{2}}\hbar \begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix} \ , \quad J_3 \,=\, \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \ . \tag{8.34}$$

We shall look at the representation with $j = 1$ in a different way in the next subsection. As we said at the start of section 5, quantum theory is really linear algebra, and we can always represent linear maps as matrices by choosing a basis. Up until now though we have mainly been dealing with infinite-dimensional inner product spaces, and writing down linear operators as infinite-dimensional matrices isn't the most convenient way to proceed. We note that Heisenberg, Born and Jordan did invent quantum mechanics as a theory of such infinite dimensional matrices.

Notice that one way to summarize the results of this section is to say that for all $n \in \mathbb{N}$ we have constructed three $n \times n$ Hermitian matrices J_i, $i = 1, 2, 3$, that satisfy the commutation relations (8.7).

8.4 Orbital angular momentum and spherical harmonics

We originally introduced angular momentum via the *orbital* angular momentum operators $\mathbf{J} = \mathbf{L} = \mathbf{X} \wedge \mathbf{P}$ in (8.3). On the other hand, we know that $\mathbf{X}$ and $\mathbf{P}$ act on wave functions $\psi(\mathbf{x})$ as multiplication by $\mathbf{x}$ and $-\mathrm{i}\hbar\nabla$, respectively, which makes the $\mathbf{L}$ *differential operators* acting on such wave functions. For example,

$$L_3 \,=\, X_1 P_2 - X_2 P_1 \,=\, -\mathrm{i}\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) \ , \tag{8.35}$$

writing $\mathbf{x} = (x, y, z)$. As for the harmonic oscillator it must then be possible to find *explicit* eigenfunctions for these operators, representing the eigenstates of angular momentum we found abstractly in the previous subsection.

It is convenient to first write the operators L_i in terms of *spherical polar coordinates* $(x, y, z) = (r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta)$, and then using the chain rule

$$\frac{\partial}{\partial\phi} \,=\, \frac{\partial x}{\partial\phi}\frac{\partial}{\partial x} + \frac{\partial y}{\partial\phi}\frac{\partial}{\partial y} + \frac{\partial z}{\partial\phi}\frac{\partial}{\partial z} \,=\, -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y} \ . \tag{8.36}$$

Thus

$$L_3 \;=\; -\mathrm{i}\hbar\frac{\partial}{\partial\phi} \; . \tag{8.37}$$

Recall that the eigenstates ψ_m defined in the previous subsection had eigenvalue $m\hbar$ under $J_3 = L_3$, so that $L_3\psi_m = m\hbar\psi_m$ reads

$$\frac{\partial\psi_m}{\partial\phi} \;=\; \mathrm{i}m\psi_m \; , \tag{8.38}$$

which integrates to

$$\psi_m(r,\theta,\phi) \;=\; F(r,\theta)\,\mathrm{e}^{\mathrm{i}m\phi} \; . \tag{8.39}$$

The coordinate ϕ has period 2π, so in order for the function (8.39) to be single-valued (well-defined) we must have $m \in \mathbb{Z}$ being an integer. Correspondingly then also $j \in \mathbb{Z}_{\geq 0}$ must be integer in Theorem 8.6. This proves

Proposition 8.7 *For orbital angular momentum the parameters j, m in Theorem 8.6 must both be* integers*. In this context we instead label $j = \ell \in \mathbb{Z}_{\geq 0}$.*

To pursue this further, let us also work out the raising and lowering operators in terms of spherical polars:

Proposition 8.8 *In spherical polar coordinates the raising and lowering operators are*

$$\begin{aligned} L_+ \;&=\; L_1 + \mathrm{i}L_2 \;=\; \hbar\mathrm{e}^{\mathrm{i}\phi}\left(\frac{\partial}{\partial\theta} + \mathrm{i}\cot\theta\frac{\partial}{\partial\phi}\right) \; , \\ L_- \;&=\; L_1 - \mathrm{i}L_2 \;=\; -\hbar\mathrm{e}^{-\mathrm{i}\phi}\left(\frac{\partial}{\partial\theta} - \mathrm{i}\cot\theta\frac{\partial}{\partial\phi}\right) \; , \end{aligned} \tag{8.40}$$

while the total angular momentum operator is

$$L^2 \;=\; L_1^2 + L_2^2 + L_3^2 \;=\; -\hbar^2\left(\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right) \; . \tag{8.41}$$

Proof * The non-examinable proof can be found in appendix C, but is conceptually straightforward and follows similar calculations to (8.36). ■

In section 8.3 recall that we took ψ to satisfy the eigenstate equations (8.17), which read

$$L^2\psi \;=\; \lambda\hbar^2\psi \; , \qquad L_3\psi \;=\; m\hbar\psi \; , \tag{8.42}$$

and recall that $\lambda = j(j+1) = \ell(\ell+1)$ in our notation in this section. In terms of the differential operators (8.41), (8.37), these become the differential equations

$$\left(\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right)Y_{\ell,m} \;=\; -\ell(\ell+1)Y_{\ell,m} \; , \qquad \frac{\partial}{\partial\phi}Y_{\ell,m} \;=\; \mathrm{i}mY_{\ell,m} \; , \tag{8.43}$$

where $Y_{\ell,m} = Y_{\ell,m}(\theta,\phi)$. From Theorem 8.6, we expect to find solutions for each non-negative integer $\ell \in \mathbb{Z}_{\geq 0}$, where for each ℓ in turn there is a solution for each $m = -\ell, -\ell+1, \ldots, \ell$. Of course we have solved the dependence on ϕ already above, and so may immediately write

$$Y_{\ell,m}(\theta,\phi) \;=\; P_{\ell,m}(\theta)\, \mathrm{e}^{\mathrm{i}m\phi} \ . \tag{8.44}$$

On Problem Sheet 4 you will show that the solution with $m = \ell$ is simply

$$P_{\ell,\ell}(\theta) \;=\; a_\ell\, (\sin\theta)^\ell \ , \tag{8.45}$$

where a_ℓ is an appropriate normalization constant. We know from section 8.3 that we may obtain $Y_{\ell,m}(\theta,\phi)$ by applying the lowering differential operator L_- in (8.40) iteratively. That is, from (8.19)

$$Y_{\ell,m-1}(\theta,\phi) \;\equiv\; \frac{1}{\hbar\sqrt{(\ell+m)(\ell-m+1)}} L_- Y_{\ell,m}(\theta,\phi) \ , \tag{8.46}$$

starting from $m = \ell$, and then iteratively $m = \ell-1, \ell-2, \ldots, -\ell+1$, and from this one can extract $P_{\ell,m-1}(\theta)$ from (8.44).

Definition The functions $Y_{\ell,m}(\theta,\phi)$ are called *spherical harmonics*. Appropriately normalizing a_ℓ in (8.45), they satisfy the orthonormality property

$$\int_{\theta=0}^{\pi}\int_{\phi=0}^{2\pi} \overline{Y_{\ell,m}(\theta,\phi)} Y_{\ell',m'}(\theta,\phi)\, \sin\theta\, \mathrm{d}\theta\, \mathrm{d}\phi \;=\; \delta_{\ell,\ell'}\delta_{m,m'} \ . \tag{8.47}$$

* The constant a_ℓ is fixed by normalizing $\int_{\theta=0}^{\pi}\int_{\phi=0}^{2\pi} |a_\ell|^2 (\sin\theta)^{2\ell} \sin\theta\, \mathrm{d}\theta\, \mathrm{d}\phi \;=\; 1$, which gives $a_\ell = \frac{1}{2^\ell \ell!}\sqrt{(2\ell+1)!/4\pi}$.

The fact that these are orthogonal for different values of ℓ and m follows immediately from Proposition 5.2. Notice that the integral in (8.47) includes the area element on the unit sphere $\sin\theta\, \mathrm{d}\theta\, \mathrm{d}\phi$. This comes from the fact that a wave funtion is $\psi = \psi(r,\theta,\phi)$, and in polar coordinates $\mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z = r^2 \sin\theta\, \mathrm{d}r\, \mathrm{d}\theta\, \mathrm{d}\phi$. The angular momentum operators L_i in (8.37), (8.40) act trivially on the radial dependence r, which is why we obtained functions of θ and ϕ, and we're effectively then integrating these functions over the unit sphere $S^2 \subset \mathbb{R}^3$ in (8.47). We shall add radial dependence back to wave functions in the next section. The spherical harmonics are *complete*: that is, any normalizable function $f(\theta,\phi)$ on the sphere can be written as

$$f(\theta,\phi) \;=\; \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} c_{\ell,m} Y_{\ell,m}(\theta,\phi) \ , \tag{8.48}$$

where $c_{\ell,m}$ are constants.

Example (Spherical harmonics for $\ell = 1$) Also putting in the normalization factors, one finds

$$Y_{1,1}(\theta,\phi) \;=\; \sqrt{\frac{3}{8\pi}}\sin\theta\, \mathrm{e}^{\mathrm{i}\phi} \ , \quad Y_{1,0}(\theta,\phi) \;=\; -\sqrt{\frac{3}{4\pi}}\cos\theta \ , \quad Y_{1,-1}(\theta,\phi) \;=\; -\sqrt{\frac{3}{8\pi}}\sin\theta\, \mathrm{e}^{-\mathrm{i}\phi} \ . \tag{8.49}$$

See Problem Sheet 4 (up to fixing the normalization factors, which is a bit tedious). The angular momentum differential operators (8.37), (8.40) act linearly on this space of eigenfunctions $\{Y_{1,1}, Y_{1,0}, Y_{1,-1}\}$, and one can verify that in this basis the operators $J_i = L_i$ precisely take the matrix form given in (8.34).

9 The hydrogen atom

9.1 Atoms

We begin with a discussion of *Coulomb's law.*[27] In general, a point charge e_2 at the origin induces an electrostatic force on another point charge e_1 at position $\mathbf{x}$ given by the inverse square law

$$\mathbf{F} = \frac{1}{4\pi\epsilon_0} \frac{e_1 e_2}{r^2} \frac{\mathbf{x}}{r} , \tag{9.1}$$

where as usual $r = |\mathbf{x}|$. Notice that the Coulomb force (9.1) is proportional to the product of the charges, so that opposite (different sign) charges attract, while like (same sign) charges repel. Electric charge is measured in *Coulombs*, C, and the proportionality constant $\epsilon_0 \simeq 8.85 \times 10^{-12}$ $\mathrm{C}^2\ \mathrm{N}^{-1}\ \mathrm{m}^{-2}$ in (9.1) is called the *permittivity of free space.* The Coulomb force is conservative, of the form $\mathbf{F} = -\nabla V$ where

$$V(\mathbf{x}) = V(r) = \frac{1}{4\pi\epsilon_0} \frac{e_1 e_2}{r} , \tag{9.2}$$

is the *Coulomb potential.*

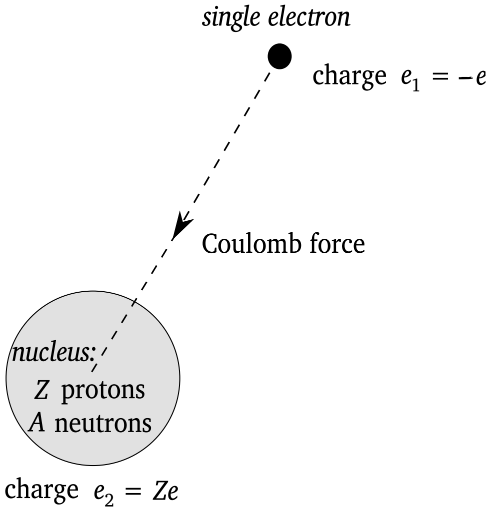

Figure 10: An atom consisting of a nucleus of atomic number Z and a single orbiting electron.

An atom consists of negatively charged electrons e^- orbiting a positively charged nucleus, where the force of attraction is (predominantly) electrostatic. An electron carries electric charge $-e$, where $e \simeq 1.60 \times 10^{-19}$ C. In general, the nucleus of an atom consists of Z positively charged

---

[27] * Coulomb discovered his law in 1783. From a modern point of view, this law is subsumed into Maxwell's classical theory of electromagnetism – see the Part B7.2 course Electromagnetism.

protons, each of charge $+e$, and A neutrons, of charge 0, and these are tightly bound together (by the *strong force*). In the following we consider an atom of *atomic number* Z, with a single orbiting electron. In particular, the hydrogen atom has $Z = 1$. Since the nucleus is vastly heavier than the electron ($m_{\text{proton}} \simeq 1836\, m_{e^-}$), this implies that the centre of mass of the atomic system will always be very close to the nucleus. We thus make the simplifying assumption that the nucleus is fixed, at the origin. The potential for the force acting on the electron is then given by (9.2) with $e_1 = -e$, $e_2 = Ze$ (see Figure 10):

$$V(r) \; = \; -\frac{Z\, e^2}{4\pi\epsilon_0\, r} \; . \tag{9.3}$$

Recall that in section 1.2 we discussed the empirical formula (1.3) for the energy levels of the hydrogen atom. Our task in the remainder of this section is to derive this formula theoretically, using the Schrödinger equation. However, before moving on to this, as usual we pause to make some comments on the corresponding classical problem. The $1/r$ Coulomb potential is formally the same as Newton's gravitational potential that you encountered in Prelims, the only difference being that gravity is always attractive. Thus classically the atomic system in Figure 10 is similar to a planet orbiting the sun under gravity.[28] In particular, the energy levels are continuous, not discrete as in (1.3).

* There is an even more serious problem with the classical picture above. Maxwell's classical theory of electromagnetism predicts that an *accelerating* charge, such as an electron in orbit around a nucleus, emits electromagnetic radiation, and thus continuously loses energy. The electron would quickly spiral in towards the nucleus, and all atoms would hence be unstable! We shall see how quantum theory gets around this problem.

9.2 Central potentials

For hydrogen-like atoms we must solve the three-dimensional stationary state Schrödinger equation (2.12) with potential $V(r)$ given by (9.3):

$$-\frac{\hbar^2}{2M}\nabla^2\psi + V(r)\,\psi \; = \; E\psi \; , \tag{9.4}$$

where $M = m_{e^-} \simeq 9.12 \times 10^{-31}$ kg is the mass of an electron.[29] For the time being it is no more complicated to consider a general central potential $V(r)$, depending only on distance $r = |\mathbf{x}|$ from the origin.

We begin by separating variables in spherical polar coordinates (r, θ, ϕ), writing

$$\psi \; = \; R(r)Y(\theta,\phi) \; . \tag{9.5}$$

---

[28] Of course there is *also* a gravitational attraction between the electron and nucleus, but this is many orders of magnitude smaller.

[29] In this section we denote the mass of the particle by M, rather than m, since the latter will be used to denote the angular momentum quantum number in $Y_{\ell,m}(\theta,\phi)$.

We then need the form of the Laplacian in spherical polars, which from Prelims is

$$\nabla^2\psi \;=\; \frac{1}{r}\frac{\mathrm{d}^2}{\mathrm{d}r^2}\left[rR(r)\right]Y(\theta,\phi)+R(r)\,\frac{1}{r^2}\left(\frac{\partial^2}{\partial\theta^2}+\cot\theta\frac{\partial}{\partial\theta}+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right)Y(\theta,\phi)\;. \tag{9.6}$$

Substituting into (9.4) and multiplying by $2Mr^2/\hbar^2RY$, we see that we can separate all r-dependent terms to one side, giving

$$\frac{1}{Y}\left(\frac{\partial^2}{\partial\theta^2}+\cot\theta\frac{\partial}{\partial\theta}+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right)Y \;=\; \frac{2Mr^2}{\hbar^2}(V(r)-E)-\frac{r(rR)''}{R}\;, \tag{9.7}$$

where a prime will denote differentiation with respect to r. Both sides of (9.7) must then be constant, say $-\lambda$, and in particular we have

$$\left(\frac{\partial^2}{\partial\theta^2}+\cot\theta\frac{\partial}{\partial\theta}+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right)Y(\theta,\phi)\;=\;-\lambda Y(\theta,\phi)\;. \tag{9.8}$$

This is precisely the first equation in (8.43)! Indeed, you might have noticed in section 8.4 that L^2 is minus $r^2\hbar^2$ times the angular part of the Laplace operator in spherical polar coordinates. We may then use our results from section 8 to deduce that the solutions to (9.8) are spherical harmonics $Y=Y_{\ell,m}(\theta,\phi)$ with eigenvalues $\lambda=\ell(\ell+1)$. Here $\ell\in\mathbb{Z}_{\geq 0}$, and for each ℓ we have $m=-\ell,-\ell+1,\ldots,\ell$. Setting the right hand side of (9.7) equal to $-\lambda=-\ell(\ell+1)$ and rearranging then gives

$$\frac{1}{r}(rR)''-\frac{\ell(\ell+1)}{r^2}R-\frac{2M}{\hbar^2}V(r)R\;=\;-\frac{2ME}{\hbar^2}R\;. \tag{9.9}$$

We have thus reduced the problem to solving a second order ODE for a single function $R(r)$.

Notice that we have shown that for any central potential $V(r)$ the stationary states $\psi=R(r)Y_{\ell,m}(\theta,\phi)$ are simultaneously eigenstates of the Hamiltonian operator H *and* also the angular momentum operators L^2 and L_3. This could have been foreseen: for the Hamiltonian

$$H\;=\;\frac{1}{2M}|\mathbf{P}|^2+V(r)\;, \tag{9.10}$$

one can verify that

$$[H,L^2]\;=\;0\;=\;[H,L_3]\;, \tag{9.11}$$

and commuting operators can have simultaneous eigenfunctions.

9.3 The spectrum of the hydrogen atom

For the hydrogen-like atom the central potential is given by (9.3). Recall here that the nucleus has charge Ze, for some positive integer Z, and the electron has charge $-e$ and mass M. The method for solving (9.9) is then very similar to that for the harmonic oscillator in section 4.1, although there is one technical difference, as we shall see. As for the harmonic oscillator, we first change variables to remove the physical constants. We define

$$\kappa^2\;\equiv\;-\frac{2ME}{\hbar^2}\;,\qquad \beta\;\equiv\;\frac{2Z}{a}\;, \tag{9.12}$$

where we have introduced

Definition The *Bohr radius* is

$$a \equiv \frac{4\pi\epsilon_0\hbar^2}{Me^2} \simeq 5.29 \times 10^{-11}\,\mathrm{m} \, , \tag{9.13}$$

where $M = m_{e^-}$ is the mass of the electron.

Of course, *a priori* the constant κ may be imaginary, but we are anticipating that the energy E will be negative, as in (1.3). Equation (9.9) now becomes

$$R'' + \frac{2}{r}R' - \frac{\ell(\ell+1)}{r^2}R + \frac{\beta}{r}R = \kappa^2 R \, . \tag{9.14}$$

It turns out to be convenient to then substitute

$$R(r) \equiv f(r)\,\mathrm{e}^{-\kappa r} \, . \tag{9.15}$$

The motivation is similar to that for the harmonic oscillator, and we are looking for $f(r)$ that is a finite power series. Then at large values of r the term on the right hand side of (9.14) dominates over the second, third and fourth terms on the left hand side. One thus expects the $R(r) = A_\pm\,\mathrm{e}^{\pm\kappa r}$ solutions to the equation $R'' = \kappa^2 R$ to determine the leading asymptotic behaviour of solutions to (9.14). The *normalization condition* for the full wave function $\psi(\mathbf{x}) = R(r)Y_{\ell,m}(\theta,\phi)$ is

$$\begin{aligned} 1 = \iiint_{\mathbb{R}^3} |\psi(\mathbf{x})|^2 \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z &= \int_0^\infty r^2 |R(r)|^2\,\mathrm{d}r \int_{\theta=0}^{\pi}\int_{\phi=0}^{2\pi} |Y_{\ell,m}(\theta,\phi)|^2 \,\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi \, , \\ &= \int_0^\infty r^2|R(r)|^2\,\mathrm{d}r \, , \end{aligned} \tag{9.16}$$

where in the last step we have used the fact that the spherical harmonics $Y_{\ell,m}(\theta,\phi)$ are normalized as in (8.47). Thus one must certainly take the minus sign in $\mathrm{e}^{\pm\kappa r}$ to obtain a normalizable solution near to $r = \infty$. We also see that E must be *negative*; otherwise $|\mathrm{e}^{\pm\kappa r}| = 1$, and this asymptotic solution is not normalizable at infinity.

The substitution (9.15) leads to

$$R' = \left(f' - \kappa f\right)\mathrm{e}^{-\kappa r} \, , \qquad R'' = \left(f'' - 2\kappa f' + \kappa^2 f\right)\mathrm{e}^{-\kappa r} \, , \tag{9.17}$$

so that (9.14) becomes

$$f'' + \left(\frac{2}{r} - 2\kappa\right)f' - \left(\frac{2\kappa - \beta}{r} + \frac{\ell(\ell+1)}{r^2}\right)f = 0 \, . \tag{9.18}$$

We again try a series solution. The main technical difference with the harmonic oscillator equation is that the coefficients in (9.18) are *singular* at $r = 0$.[30] Thus we should write

$$f(r) = \sum_{k=0}^{\infty} a_k r^{k+c} \, , \tag{9.19}$$

---

[30] * More precisely, $r = 0$ is a *regular singular point* of (9.18). Again, Fuchs' theorem applies and determines the form of the series in (9.19). See appendix A.

for some real constant $c \in \mathbb{R}$, and without loss of generality we may assume $a_0 \neq 0$. We then have

$$f' = \sum_{k=0}^{\infty} (k+c) a_k r^{k+c-1} \ , \qquad f'' = \sum_{k=0}^{\infty} (k+c)(k+c-1) a_k r^{k+c-2} \ , \tag{9.20}$$

and (9.18) becomes

$$\sum_{k=0}^{\infty} \left[(k+c)(k+c-1) a_k r^{k+c-2} + \left(\frac{2}{r} - 2\kappa\right)(k+c) a_k r^{k+c-1} \right.$$
$$\left. - \left(2\kappa - \beta + \frac{\ell(\ell+1)}{r}\right) a_k r^{k+c-1} \right] = 0 \ . \tag{9.21}$$

The *lowest* power of r in (9.21) is r^{c-2}, with coefficient

$$a_0 \left[c(c-1) + 2c - \ell(\ell+1) \right] = a_0 (c - \ell)(c + \ell + 1) = 0 \ . \tag{9.22}$$

Since $a_0 \neq 0$ we deduce that $c = -\ell - 1$ or $c = \ell$. However, $c = -\ell - 1$ is forbidden: with this choice of c, the function $R(r) = f(r)\,\mathrm{e}^{-\kappa r} \sim \frac{a_0}{r^{\ell+1}}$ near to $r = 0$. The wave function is *singular* at the origin, which in itself would not be a problem, since we only require normalizability in the sense of (9.16). However for $\ell > 0$, the singularity is not integrable. This leaves the $\ell = 0$, which gives a normalizable function for $c = -1$, which however does not solve the Schrödinger equation at the origin. To see this, recall that $\nabla^2 \frac{1}{r} = -4\pi \delta^3(\mathbf{x})$, giving a term in the Schrödinger equation which does not get cancelled. We thus discard the $c = -\ell - 1$ solution.

Hence $c = \ell$, and setting to zero the coefficient of r^{k+c-1} in (9.21) gives the recurrence relation

$$\left[(k+c+1)(k+c) + 2(k+c+1) - \ell(\ell+1) \right] a_{k+1} = \left[2\kappa(k+c) + 2\kappa - \beta \right] a_k \ . \tag{9.23}$$

Putting $c = \ell$, after some simplifications one finds

$$a_{k+1} = \frac{2\kappa(k+\ell+1) - \beta}{(k+1)(k+2+2\ell)} a_k \ . \tag{9.24}$$

Fixing a_0, this determines the solution completely. Notice that, in contrast to the harmonic oscillator, here we have only a *single* solution – we have already discarded the $c = -\ell - 1$ solution as singular and thus unphysical.

Finally, we must impose that the solution is normalizable. Suppose that the series does *not* terminate. Then from (9.24)

$$\frac{a_{k+1}}{a_k} \sim \frac{2\kappa}{k} \ , \qquad \text{as } k \to \infty \ . \tag{9.25}$$

Compare this to the Taylor expansion of the function $\mathrm{e}^{2\kappa r} = \sum_{k=0}^{\infty} b_k r^k$, where $b_k = (2\kappa)^k / k!$. In this case

$$\frac{b_{k+1}}{b_k} = \frac{(2\kappa)^{k+1} k!}{(2\kappa)^k (k+1)!} \sim \frac{2\kappa}{k} \ . \tag{9.26}$$

If the series (9.19) does not terminate the function $f(r)$ and $\mathrm{e}^{2\kappa r}$ have the same asymptotic expansion, and hence

$$R(r) \,=\, f(r)\,\mathrm{e}^{-\kappa r} \,\sim\, \mathrm{e}^{\kappa r} \tag{9.27}$$

is *not* normalizable in (9.16).

The series must hence terminate. That is, there is a least integer $n \geq 0$ such that $a_{n+1} = 0$. The recurrence relation (9.24) then gives

$$2\kappa(n+\ell+1) \,=\, \beta\ . \tag{9.28}$$

Note that then $a_k = 0$ for all $k \geq n+1$, and this makes $f(r)$ in (9.19) a polynomial of degree $N-1$, where we define

$$N \,\equiv\, n+\ell+1 \,\geq\, \ell+1\ . \tag{9.29}$$

Recalling the definitions (9.12) of κ and β, this becomes

$$E \,=\, E_N \,=\, -\frac{\hbar^2}{2M}\kappa_N^2 \,=\, -\frac{\hbar^2 Z^2}{2Ma^2N^2}\ , \tag{9.30}$$

for a positive integer $N \in \mathbb{N}$. Here κ_N is the solution to (9.28), namely

$$\kappa \,=\, \kappa_N \,\equiv\, \frac{\beta}{2N} \,=\, \frac{Z}{aN}\ . \tag{9.31}$$

The energy levels (9.30) are precisely of the form (1.3), and this was an enormous triumph for quantum theory. Denoting the polynomial $f(r)$ as $f_{N,\ell}(r)$, since both N and ℓ enter its definition, we have thus proven:

Theorem 9.1 *The energy levels for an atom consisting of a single electron orbiting a nucleus of atomic number Z are*

$$E \,=\, E_N \,=\, -\frac{\hbar^2 Z^2}{2Ma^2}\cdot\frac{1}{N^2}\ , \tag{9.32}$$

where the positive integer $N \in \mathbb{N}$ is called the principal quantum number*. The corresponding wave functions are*

$$\psi \,=\, \psi_{N,\ell,m}(r,\theta,\phi) \,=\, f_{N,\ell}(r)\,\mathrm{e}^{-Zr/aN}\,Y_{\ell,m}(\theta,\phi)\ , \tag{9.33}$$

where $f_{N,\ell}(r)$ is a polynomial of degree $N-1$. Appropriately normalized, the latter are known as generalized Laguerre polynomials*.*

The wave functions (9.33) have L^2 eigenvalue $\ell(\ell+1)\hbar^2$ and L_3 eigenvalue $m\hbar$, where ℓ is called the azimuthal quantum number*, and m is called the* magnetic quantum number*. For fixed principal quantum number $N \in \{1,2,3,\ldots\}$, we see from (9.29) that the range of ℓ is $\ell = 0,1,\ldots,N-1$, and for fixed ℓ the range of m is $m = -\ell,-\ell+1,\ldots,\ell$.*

Notice that we can count the degeneracy of the energy eigenstates as follows. The energy E_N depends on $N = n + \ell + 1$, so as stated in the Theorem above $0 \leq \ell \leq N - 1$, and for each ℓ there are $2\ell + 1$ allowed values of m. Thus

$$\text{degeneracy of } E_N \;=\; \sum_{\ell=0}^{N-1} (2\ell + 1) \;=\; N^2 \ . \tag{9.34}$$

In particular the *ground state* has $N = 1$ and is non-degenerate, leading to a *stable atom*.

* In section 8.3 we mentioned that in addition to orbital angular momentum $\mathbf{J} = \mathbf{L}$, represented by the spherical harmonic $Y_{\ell,m}(\theta, \phi)$ in the electron wave function (9.33), the electron also has an "internal" angular momentum called "spin" $\mathbf{J} = \mathbf{S}$ with $j = \frac{1}{2}$. This means that the full electron wave function ψ takes value in $\mathcal{H}_{1/2} = \mathbb{C}^2$, rather than $\mathbb{C}$. In practice this means that (9.33) multiplied by an arbitrary linear combination of the spin up $\psi_+ = (1, 0)^T$ and spin down $\psi_- = (0, 1)^T$ states in (8.27). The *Pauli exclusion principle* says that no two electrons can occupy the same quantum state, including this spin.

These additional remarks in fact allow us to derive the structure of the Periodic Table of elements! We have described the hydrogen atom already. A helium atom has two electrons, and ignoring electron-electron interactions the wave functions of each electron are given by (9.33). Recall that the ground state with $N = 1$ is unique, but taking into account spin and the Pauli exclusion principle, if one electron of helium occupies the $N = 1$ ground state with spin up ψ_+, the other must be in the $N = 1$ ground state with spin down ψ_-. This is the first row of the Periodic Table, with hydrogen and helium, where the ground state is described by electrons in these two $N = 1$ states. Electrons with the same principal quantum number N are said to be in the same *shell*.

The next row of the Periodic Table, lithuim through to neon, has 8 elements, but we can now explain why. By the Pauli exclusion principle, having filled the two $N = 1$ states with electrons (the first electron shell), the additional electrons must occupy the $N = 2$ states in the lowest energy state. We have already explained that the degeneracy of these is $N^2 = 2^2 = 4$, and including the spin up and spin down degrees of freedom, that makes $2 \times 4 = 8$ different electron states in the second shell. The next electron shell similarly has $2 \times 3^2 = 18$ states, which accounts for the third row of the Periodic Table, with 8 elements, plus the 10 elements in the middle of the 4th row (scandium through to zinc).

The structure of the columns of the Periodic Table can be understood in terms of (ℓ, m) values. For example, the 8 $N = 2$ states split as $8 = 2 \times 1 + 2 \times 3$, where there is one $\ell = 0$ state, but three $\ell = 1$ states with $m = 1, 0$ and -1, respectively, which explains the two columns on the left of the periodic table, and the 6 on the right. Similarly, 18 splits via orbital angular momentum states into $18 = 2 \times 1 + 2 \times 3 + 2 \times 5$. The structure of the Periodic Table can thus be understood in terms of total energy, orbital angular momentum, and spin of the electrons in the atoms. An important shortcoming of the single electron model of atomic orbitals that we built is that it predicts degeneracies between levels with different n and ℓ values, if the sum $n + \ell$ is equal. Relativistic effects and electron-electron interactions lift this degeneracy.

Given the new quantum number N, and the simple expression for the energy (9.32) in terms of N, you might wonder if there is an *algebraic* way to derive these energy levels, as for the harmonic oscillator in section 7. In fact there is, as first shown by Pauli in 1926. In fact Pauli's derivation of (9.32) was just *before* Schrödinger deduced the same result from his new equation! Pauli made use of the *Laplace-Runge-Lenz vector* $\mathbf{A} \equiv \mathbf{p} \wedge \mathbf{L} - MK\,\mathbf{x}/r$, which one can show is *classically* conserved for the potential $V(r) = -K/r$. In fact one can deduce that classical solutions are conic sections *algebraically*, using the fact that $\dot{\mathbf{A}} = \mathbf{0}$. The corresponding quantum operator is not immediately clear, because $\mathbf{P}$ and $\mathbf{L}$ do not commute. However, using the comment at the end of section 5.2 one can check that

$$\mathbf{A} \equiv \frac{1}{2}\left(\mathbf{P} \wedge \mathbf{L} - \mathbf{L} \wedge \mathbf{P}\right) - MK\frac{\mathbf{X}}{|\mathbf{X}|} , \tag{9.35}$$

is a self-adjoint operator. Moreover, $[\mathbf{A}, H] = 0$, where H is the Hamiltonian (9.10) with potential $V = -K/|\mathbf{X}|$, and $[L_i, A_j] = \mathrm{i}\hbar \sum_{k=1}^{3} \epsilon_{ijk} A_k$, as in Proposition 8.2. For more details of this fascinating topic, which is definitely well beyond our syllabus, see section 8.7 of the book by Hannabuss.

9.4 Rotationally symmetric solutions

We begin by looking more closely at the ground state. This has $N - 1 = \ell = m = 0$. Since $Y_{0,0}(\theta, \phi) = \frac{1}{\sqrt{4\pi}}$ is a constant, the wave function is hence independent of θ and ϕ, and can be written as

$$\psi_{1,0,0} = \tilde{a}_0\, \mathrm{e}^{-\kappa_1 r} = \tilde{a}_0\, \mathrm{e}^{-Zr/a} , \tag{9.36}$$

where we have used (9.31). It is straightforward to normalize (9.36), using (9.16):

$$\begin{aligned} 1 &= \iiint_{\mathbb{R}^3} |\psi_{1,0,0}|^2 \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = 4\pi|\tilde{a}_0|^2 \int_0^\infty r^2 \mathrm{e}^{-2\kappa_1 r}\,\mathrm{d}r \\ &= \pi|\tilde{a}_0|^2 \frac{\mathrm{d}^2}{\mathrm{d}\kappa_1^2} \int_0^\infty \mathrm{e}^{-2\kappa_1 r}\,\mathrm{d}r = \pi|\tilde{a}_0|^2 \frac{\mathrm{d}^2}{\mathrm{d}\kappa_1^2}\left[\frac{1}{2\kappa_1}\right] \\ &= \frac{\pi|\tilde{a}_0|^2 a^3}{Z^3} . \end{aligned} \tag{9.37}$$

Here in the first line we have integrated over the unit sphere to obtain area 4π. We may hence take $\tilde{a}_0 = \sqrt{Z^3/\pi a^3}$, and then

$$\psi_{1,0,0} = \sqrt{\frac{Z^3}{\pi a^3}}\, \mathrm{e}^{-Zr/a} , \tag{9.38}$$

is the *normalized ground state wave function.*

Example Let us compute the expectation value of r, the distance of the electron from the nucleus, in the ground state. This is given by

$$\begin{aligned} \mathbb{E}_{\psi_{1,0,0}}(r) &= \int_0^\infty r|\psi_{1,0,0}|^2\, 4\pi r^2\,\mathrm{d}r \\ &= -\frac{4\pi|\tilde{a}_0|^2}{8}\frac{\mathrm{d}^3}{\mathrm{d}\kappa_1^3}\left[\frac{1}{2\kappa_1}\right] = \frac{3\pi|\tilde{a}_0|^2}{2\kappa_1^4} \\ &= \frac{3a}{2Z} , \end{aligned} \tag{9.39}$$

where we have used $\tilde{a}_0 = \sqrt{Z^3/\pi a^3}$. Thus for the hydrogen atom, with $Z = 1$, the average distance of the electron from the nucleus, in the ground state, is $\frac{3}{2}$ times the Bohr radius a, which is numerically $\simeq 7.94 \times 10^{-11}$ m.

The ground state is the first in an infinite series of rotationally symmetric (or spherically symmetric) solutions, one at each energy level. One simply puts $\ell = 0 = m$. The wave functions (9.33) are then of the form

$$\psi_{N,0,0} \;=\; \frac{f_{N,0}(r)}{\sqrt{4\pi}}\, \mathrm{e}^{-\kappa_N r} \;, \tag{9.40}$$

where $f_{N,0}$ is a polynomial of degree $N-1$ satisfying (9.18) with $\ell = 0$, and $\kappa_N = \beta/2N$. Using the recurrence relation (9.24) one can write the solution as

$$f_{N,0}(r) \;=\; a_{N,0}\left[1 + \frac{2\kappa_N - \beta}{2} r + \frac{(4\kappa_N - \beta)(2\kappa_N - \beta)}{12} r^2 + \cdots\right] \;. \tag{9.41}$$

Here the coefficient of r^N is zero, since it is proportional to $2N\kappa_N - \beta$, with all higher order terms then also zero. For example, up to normalization, setting $\kappa_2 = \beta/4$, $\kappa_3 = \beta/6$ this gives

$$f_{2,0}(r) \;=\; 1 - \frac{\beta}{4} r \;, \qquad f_{3,0}(r) \;=\; 1 - \frac{\beta}{3} r + \frac{\beta^2}{54} r^2 \;. \tag{9.42}$$

The corresponding spherically symmetric wave functions $\psi_N(r) \equiv \psi_{N,0,0}(r)$ are shown in Figure 11.

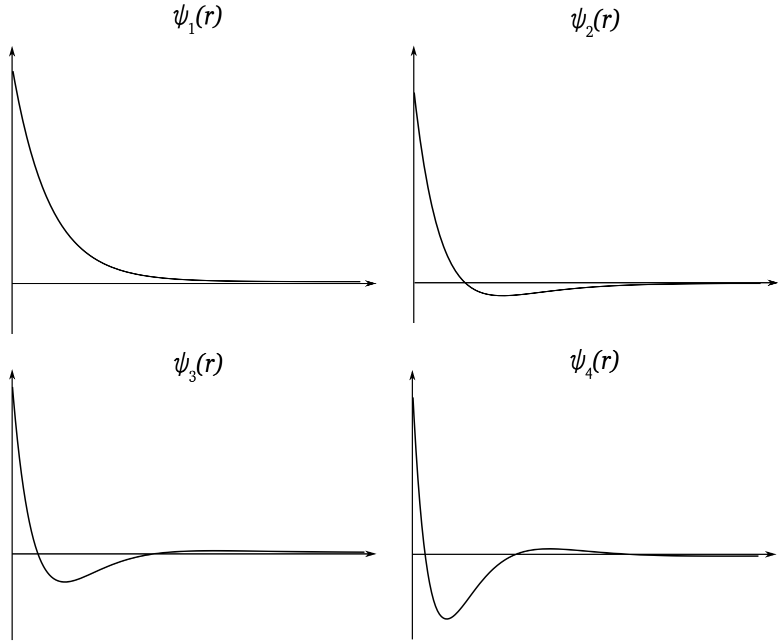

Figure 11: The ground state and first three spherically symmetric excited state wave functions $\psi_N(r) \equiv \psi_{N,0,0}(r)$ of the hydrogen atom.

10 Quantum key distribution

Quantum information theory is a blossoming new field between physics and computer science with extremely promising applications. The holy grail of this field is building a universal quantum computer. This is an enormous challenge that involves engineering, many-body quantum physics, and quantum algorithmic ideas. We can get a glimpse at this new world through the Quantum key distribution (QKD) protocol that enables unbreakable secret communication between two parties. All we need is our knowledge of two-state systems.

Alice and Bob are at different locations and want to establish a secret shared binary key that they can then use to both encrypt and decrypt messages for future classical communication. They have a communication channel, which however can be intercepted by an eavesdropper, Eve. We will see that her mischievous attempt is futile, and is revealed to Alice and Bob.

Alice and Bob agree to use four states in two different bases $\psi_{3,+} = \uparrow$, $\psi_{3,-} = \downarrow$ in the $Z \equiv J_3$ basis and $\psi_{1,+} = \rightarrow$, $\psi_{1,-} = \leftarrow$ in the $X \equiv J_1$ basis, where J_i are the angular momentum operators in the spin-1/2 representation. (In practice, photons with two polarizations, another two-level system, are used.) Alice will send these four states and Bob will measure them in the basis Z or X. Recall from eqs. (8.27) and (8.31) that

$$\begin{aligned}\uparrow &= \frac{1}{\sqrt{2}}\left(\rightarrow + \leftarrow\right) \ , \\ \downarrow &= \frac{1}{\sqrt{2}}\left(\rightarrow - \leftarrow\right) \ , \end{aligned} \tag{10.1}$$

and similarly the inverse relations

$$\begin{aligned}\rightarrow &= \frac{1}{\sqrt{2}}\left(\uparrow + \downarrow\right) \ , \\ \leftarrow &= \frac{1}{\sqrt{2}}\left(\uparrow - \downarrow\right) \ . \end{aligned} \tag{10.2}$$

These relations tell us that if Bob measures e.g. ↑ in the X basis, he gets → and ← with equal, 1/2 probability.

The protocol then consists of four steps. Alice and Bob can establish these rules on the phone and Eve can listen in.

1. Alice sends a series of electrons in states of either of ↑, ↓, →, ← randomly selected by her.
2. Bob measures the photons in either the Z or X basis chosen randomly by him and records the binary eigenvalue of $(\mathbb{1} + Z)/2$ or $(\mathbb{1} + X)/2$.
3. Over their nonsecure phone line, Bob tells Alice what basis he used at each step, but he keeps silent on the outcomes. Alice then tells him "Yes" if he used an aligned basis to measure and "No" if the basis was misaligned.
4. They throw away their records for each No. For each Yes, both Alice and Bob know the state of the electron with full certainty and can now write down their binary key. This is trivially true for Alice, while Bob learns that he conducted a measurement on an eigenstate.

Alice sends	→	↓	↓	→	→	↑	↓	←	↓	↑
Bob uses	Z	Z	X	Z	X	Z	X	Z	Z	Z
Bob finds	↑	↓	→	↓	→	↑	←	↓	↓	↑
Alice says	No	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Key	-	1	-	-	0	0	-	-	1	0

In the table below we illustrate the protocol for 10 electrons that resulted in the shared key 10010.

Let us now analyze what happens if Eve tries to intercept the QKD protocol. If she just listens to the phone call, but doesn't touch the electrons, she clearly learns nothing. We will assume that she eavesdrops on every electron sent by Alice, by measuring them in a randomly chosen basis and then transmitting the resulting state to Bob. (Other eavesdropping scenarios can also be shown to be ineffective.)

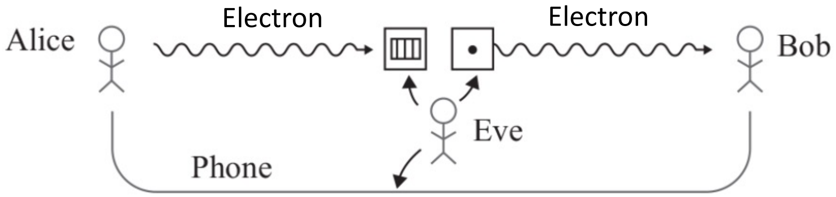

Figure 12: The QKD setup: Alice sends a sequence of electrons in randomly chosen states to Bob. These electrons are intercepted by Eve, who measures them with her left box and then generates replacement electrons for Bob with her right box. Eve also eavesdrop on Alice and Bob's phone conversation. *Figure adapted from Zwiebach: Mastering Quantum Mechanics with the author's permission.*

Let us focus on the electrons that end up contributing to the key because the phone call tagged them as Y. This means that Bob's basis is aligned with the electron state Alice sent. Eve's detector is also aligned with the state with probability $p_{\text{align}} = \frac{1}{2}$ and misaligned with the same probability. Now we have to compute the probability of three possible outcomes:

- **(a)** Eve's detector is also **aligned**, all three put the same mark in their notebooks. This outcome has probability $p_{\mathbf{a}} = p_{\text{align}} = \frac{1}{2}$.
- **(m)** Eve's detector is **misaligned**, the outgoing electron is an equal weight superposition in Bob's basis. [E.g. $\uparrow \xrightarrow[X_{\text{Eve}}]{} \rightarrow = \frac{1}{\sqrt{2}}(\uparrow + \downarrow)$.]
  - **(m+o)** After Bob makes his measurement, the electron is back to its **original** state. This outcome has probability $p_{\mathbf{m+o}} = p_{\text{align}} \cdot p_{\text{measure}} = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.
  - **(m+f)** After Bob makes his measurement, the electron has **flipped** compared to its original state and Alice's and Bob's notebooks disagree. This outcome has probability $p_{\mathbf{m+f}} = p_{\text{align}} \cdot p_{\text{measure}} = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.

With Eve's interception, the probability that Alice's and Bob's n-digit keys will agree is

$$p_{\text{agree},n} = (p_{\mathbf{a}} + p_{\mathbf{m+o}})^n = \left(\frac{3}{4}\right)^n . \tag{10.3}$$

Then over the phone Alice can pick $n/2$ digits randomly and compare with Bob. Choosing $n = 140$, the probability that the compared portion of the key agrees with Eve's interception is extremely small, $p_{\text{agree},70} \approx 10^{-9}$. If the shared portion of the keys agree, they can safely assume that Eve was not eavesdropping and they can use the remaining 70 digits of their key to safely communicate. We remark that this last step can be dramatically improved for efficiency.

An alternative QKD protocol was proposed by Artur Ekert from the Mathematical Institute that builds on entanglement, a key concept in modern quantum theory. You can learn about entanglement in the Part B7.3 course Further Quantum Theory and in the Part C7.4 course Introduction to Quantum Information.

11 Epilogue

Quantum mechanics stands as one of the great achievements of the 20th century, not just because of the radically different view of the physical world it provides, confirmed by every conceivable experiment, but also because of what this in turn has led to. Much of our modern day technology was developed from ideas born from quantum theory. But quantum mechanics still leaves us with some fundamental questions that, despite the best efforts of a century of theorists and experimentalists, still have no satisfactory answers. In this final section we briefly mention some of the issues.

Time evolution in quantum mechanics is described by the two-step process summarized in section 5.5. The first, where wave functions evolve in time according to the Schrödinger equation, is *deterministic*. Most of the issues in quantum theory involve the second of these two steps: measurement. This is *fundamentally probabilistic* – something that many founders of quantum theory, notably Einstein, thought anathema. Probability enters into classical mechanics when we are ignorant of some information required to fully specify the state of the system. We have seen examples of this earlier in the course. For example, if you are ignorant of the initial position of a classical particle in a box, then when you look to see where it is it is equally likely to be anywhere in the box, *i.e.* its position is given by the uniform distribution. In 1935 Einstein, Podolsky and Rosen famously argued that quantum theory is "incomplete" in this sense, so that more information is required to fully specify the state of particles such as electrons, and with this added information quantum mechanics would become deterministic. Whatever that information might be goes under the name of *hidden variables*. However, in 1964 John Bell gave a very general proof that this cannot be true, called *Bell's theorem*. It is based on a beautiful mathematical argument, but more importantly it is also something that can be *experimentally confirmed*. You can learn about Bell's theorem in the Part C7.4 course Introduction to Quantum Information.

Bell's theorem of course has some assumptions that go into it, that one might debate, but even if we accept that quantum theory is fundamentally probabilistic, there are still issues. In particular, you might well have asked: what constitutes a "measurement" exactly? Do wave functions collapse only when someone in a lab is "measuring" something, or does it happen all the time as part of quantum dynamics? If so, *when* does this happen exactly, and *how/why*? A famous thought experiment that exemplifies some of these questions is *Schrödinger's cat*.[31] In its original form, one imagines a cat in a sealed box containing a vial of poison gas, and an apparatus for releasing the gas at the moment a radioactive atom decays. The radioactive decay is a quantum process, so that the radioactive atom is in a *quantum superposition* of decayed and non-decayed states which changes with time. What is the state of the cat? Is it also in a "macroscopic superposition" of alive and dead? The state of the atom will collapse into an eigenstate when observed. Suppose this requires opening the box. Does this mean that the cat will only be in an eigenstate (definitely alive or dead) when we open the box and observe it? In a variant of this thought experiment, due to Eugene Wigner in 1961, one imagines the decay of the atom simply turning on a light in the box, and replaces the cat by a person, "Wigner's friend". Does the wave function of the atom collapse when Wigner's friend sees the light go on, or only when Wigner opens the box? This gets to the heart of the problem: who or what is allowed to be an observer? If I make an observation and don't tell you the answer, does the wave function collapse for me but not for you? The latter would be a very solipsistic view of quantum theory and the nature of reality.

Of course, both theorists and experimentalists have tried to address these philosophical questions concretely, and there is a lot more one could say. It would be far more conservative to say that "measurement" is a result of a quantum particle interacting with a very large macroscopic system, irrespective of whether this involves an "observer". After all, any measuring apparatus, including any observer, are in principle described by a wave function, but one describing an enormously large number of quantum particles. One can imagine that "measurement" of the quantum particle is an approximation of some well-defined but complex dynamical interaction of the combined systems. This can be made more precise, and goes under the general name of *decoherence*, but it still doesn't really solve the measurement problem. Historically, conceptual problems that persisted in physics despite decades (or even centuries) of work were eventually solved by taking a completely different point of view. Something for the 21st century to resolve, perhaps.

A * Fuchs' theorem and Frobenius series

This appendix is included only for interest and completeness. It is certainly not intended to be part of the course.

In the main text we solved both the harmonic oscillator (4.14) and the radial part of the hydrogen atom (9.18) using a (generalized) power series expansion. Here we discuss this problem

---

[31] Schrödinger was a Fellow of Magdalen College, Oxford in 1935 when he came up with this thought experiment.

more generally. Consider the second order ODE

$$\frac{\mathrm{d}^2 f}{\mathrm{d}x^2} + p(x)\,\frac{\mathrm{d}f}{\mathrm{d}x} + q(x)\,f \;=\; 0 \;, \tag{A.1}$$

for the function $f(x)$. A point x_0 is called an *ordinary point* of (A.1) if $p(x)$, $q(x)$ have Taylor expansions about x_0, valid in some positive radius of convergence; otherwise x_0 is called a *singular point.* If x_0 is a singular point, but $(x-x_0)p(x)$, $(x-x_0)^2q(x)$ both have Taylor expansions around x_0, then x_0 is called a *regular singular point.*

By a *generalized power series*, or *Frobenius series*, about x_0 we mean

$$f(x) \;=\; \sum_{k=0}^{\infty} a_k\,(x-x_0)^{k+c} \;, \tag{A.2}$$

for some real number $c \in \mathbb{R}$. Without loss of generality, $a_0 \neq 0$. Notice that for c a non-negative integer this reduces to a normal power series. Without loss of generality, we now set $x_0 = 0$. The equation obtained from the ODE (A.1) by setting the *lowest power* of x to zero in the expansion (A.2) is called the *indicial equation.* Since (A.1) is second order, this is a quadratic equation for c. Fuchs and Frobenius proved the following result:

Theorem A.1

- *If $x = 0$ is an ordinary point, then (A.1) possesses two distinct power series solutions, so that one can effectively set $c = 0$ in (A.2). These series converge for $|x| < r$, where r is the minimum radius of convergence of $p(x)$ and $q(x)$.*

- *If $x = 0$ is a regular singular point, then there is* at least one *solution of (A.1) of the form (A.2). Again, this series solution converges for $|x| < r$, where r is the minimum radius of convergence of $xp(x)$ and $x^2q(x)$.*

The precise behaviour for regular singular points is a little involved. If the indicial equation for c has *distinct* roots that differ by a *non-integer*, then there are *two* series solutions of the form (A.2), with the corresponding values of c solving the indicial equation. Otherwise there is not *necessarily* a second series solution; if there is not one can nevertheless say more about the second solution, but we refer the interested reader to the literature for details.

B * Representations of angular momentum

In this appendix we give a proof of some of the results stated in section 8.3. To keep the appendix self-contained we repeat some of the material from section 8.3.

As for the harmonic oscillator in section 7, it is helpful to first establish some key properties of the action of the raising and lowering operators $J_\pm$ on states:

Proposition B.1 *Let* ψ *be a common eigenstate of* J^2 *and* J_3*, satisfying*[32]

$$J^2\psi \;=\; \lambda\hbar^2\psi\ , \qquad J_3\psi \;=\; m\hbar\psi\ . \tag{B.1}$$

Then

(*i*) $J^2 J_\pm\psi = \lambda\hbar^2 J_\pm\psi$.

(*ii*) $J_3 J_\pm\psi = (m\pm 1)\hbar J_\pm\psi$.

(*iii*) $\|J_\pm\psi\|^2 = [\lambda - m(m\pm 1)]\hbar^2\|\psi\|^2$.

(*iv*) $\lambda \geq m(m\pm 1)$*, and* $\lambda = m(m\pm 1)$ *if and only if* $J_\pm\psi = 0$.

Proof Part (*i*) of Proposition 8.4 says that $J^2 J_\pm = J_\pm J^2$, which applying to ψ immediately gives (*i*) of Proposition B.1. Similarly for (*ii*), part (*iv*) of Proposition 8.4 implies

$$J_3 J_\pm\psi \;=\; (J_\pm J_3 \pm \hbar J_\pm)\,\psi \;=\; J_\pm(m\hbar \pm \hbar)\psi \;=\; (m\pm 1)\hbar J_\pm\psi\ . \tag{B.2}$$

Next $(J_\pm)^* = J_\mp$, so for (*iii*)

$$\begin{aligned} \|J_\pm\psi\|^2 \;&=\; \langle J_\pm\psi|J_\pm\psi\rangle \;=\; \langle\psi|J_\mp J_\pm\psi\rangle \;=\; \langle\psi|(J^2 - J_3^2 \mp \hbar J_3)\psi\rangle \\ &=\; \left(\lambda - m^2 \mp m\right)\hbar^2\langle\psi|\psi\rangle\ , \end{aligned} \tag{B.3}$$

where we have used (*ii*) from Proposition 8.4. Finally, part (*iii*) that we have just proven immediately implies $\lambda \geq m(m\pm 1)$, with equality if and only if $J_\pm\psi = 0$. ■

Using this Proposition we can argue much as we did for the harmonic oscillator to prove:

Theorem B.2 *The eigenvalues of* J^2 *have the form* $j(j+1)\hbar^2$*, where* $j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \ldots$ *takes non-negative half-integer values. For each choice of* j*, the eigenvalues of* J_3 *are* $m\hbar$*, where* $m = -j, -j+1, -j+2, \ldots, j-1, j$.

Proof Suppose that ψ is a simultaneous eigenstate of both J^2 and J_3, with eigenvalues $\lambda\hbar^2$ and $m_0\hbar$, respectively. Applying the raising and lowering operators $J_\pm^n$ inductively to ψ, parts (*i*) and (*ii*) from Proposition B.1 generalize to

$$J^2(J_\pm^n\psi) \;=\; \lambda\hbar^2(J_\pm^n\psi)\ , \qquad J_3(J_\pm^n\psi) \;=\; (m_0 \pm n)\hbar(J_\pm^n\psi)\ . \tag{B.4}$$

The states $\{J_\pm^n\psi\}$ then all have the same J^2 eigenvalue $\lambda\hbar^2$, but different J_3 eigenvalues $m\hbar$, where $m = m_0 \pm n$. Assuming that *all* of these states are non-zero, part (*iv*) of Proposition B.1 says that $\lambda \geq m(m\pm 1)$, which may equivalently be written as

$$\lambda + \frac{1}{4} \;\geq\; \left(m \pm \frac{1}{2}\right)^2 \;=\; \left(m_0 \pm n \pm \frac{1}{2}\right)^2\ . \tag{B.5}$$

---

[32] The factors of $\hbar$ in the eigenvalues $\lambda\hbar^2$ and $m\hbar$ are for later convenience.

For fixed λ and m_0, both upper sign and lower sign inequalities in (B.5) will be violated for large enough $n \in \mathbb{N}$.

We look at the raised states first. The above implies there must be an eigenstate with $J_+^n \psi \neq 0$, but $J_+^{n+1}\psi = 0$. Let us call the J_3 eigenvalue for this state $j\hbar$ (where $j = m_0 + n$), and relabel the eigenstate as $\psi_j \equiv J_+^n \psi$. Then by definition $J_+\psi_j = 0$ and (iv) of Proposition B.1 gives

$$\lambda = j(j+1) \ . \tag{B.6}$$

Now act on ψ_j with the lowering operators J_-^k. Again, if all these states are non-zero we will violate the lower sign inequality in (B.5) for sufficiently large k, so there is some state $J_-^k\psi_j \neq 0$ with $J_-^{k+1}\psi_j = 0$. Calling the J_3 eigenvalue of this state $m\hbar$, the lower sign for (iv) gives

$$j(j+1) = \lambda = m(m-1) \qquad \Rightarrow \qquad (m+j)(m-j-1) = 0 \ . \tag{B.7}$$

We cannot have $m = j+1$, because we have applied lowering operators on ψ_j that decrease the J_3 eigenvalue, starting with eigenvalue $j\hbar$. So $m = -j$. Since we applied J_-^k on ψ_j to get from J_3 eigenvalue $j\hbar$ to $-j\hbar$, it follows that $k\hbar = j\hbar - (-j\hbar) = 2j\hbar$, so that $2j = k \in \mathbb{Z}_{\geq 0}$. Thus the possible values of j are $j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \ldots$. The states $J_-^{j-m}\psi_j$, where $m = -j, -j+1, \ldots, j-1, j$, then have J_3 eigenvalue $m\hbar$. ■

* One can also prove that the degeneracy of each eigenvalue is the same as that of $j\hbar$.

Corollary B.3 *If there are no degeneracies then an eigenspace $\mathcal{H}_j$ on which J^2 takes the eigenvalue $j(j+1)\hbar^2$ has dimension $2j+1$, with an orthonormal basis of the form $\{\psi_m \mid m = -j, -j+1, \ldots, j-1, j\}$ such that*

$$\begin{aligned} J_3\psi_m &= m\hbar\,\psi_m \ , \\ J_\pm\psi_m &= \sqrt{(j\mp m)(j\pm m+1)}\hbar\,\psi_{m\pm 1} \ . \end{aligned} \tag{B.8}$$

Proof The only thing left to prove here is the normalization factor in the second equation of (B.8). Suppose that ψ_j is normalized, and define

$$\psi_{m-1} \equiv \frac{1}{\hbar\sqrt{(j+m)(j-m+1)}} J_-\psi_m \ , \tag{B.9}$$

starting from $m = j$, and then iteratively $m = j-1, j-2, \ldots, -j+1$. This defines the basis, and by construction the lower sign in the second equation in (B.8) holds. Each of these is normalized since

$$\begin{aligned} \|\psi_{m-1}\|^2 &= \frac{1}{\hbar^2(j+m)(j-m+1)}\|J_-\psi_m\|^2 \\ &= \frac{1}{\hbar^2(j+m)(j-m+1)}\left[(j(j+1) - m(m-1)\right]\hbar^2\|\psi_m\|^2 \\ &= \|\psi_m\|^2 \ , \end{aligned} \tag{B.10}$$

where we have used (*iii*) from Proposition B.1 with $\lambda = j(j+1)$. Moreover, the different states ψ_m, $\psi_{m'}$ for $m \neq m'$ are orthogonal because they have distinct eigenvalues under J_3. Finally, using (*ii*) from Proposition 8.4 we may apply J_+ to the left hand side of (B.9) to obtain

$$\begin{aligned} J_+\psi_{m-1} &= \frac{1}{\hbar\sqrt{(j+m)(j-m+1)}}\left(J^2 - J_3^2 + \hbar J_3\right)\psi_m \\ &= \frac{1}{\hbar\sqrt{(j+m)(j-m+1)}}\left[j(j+1) - m^2 + m\right]\hbar^2\psi_m \\ &= \sqrt{(j-m+1)(j+m)}\hbar\,\psi_m \ , \end{aligned} \tag{B.11}$$

which is the upper sign in the second equation of (B.8), with m replaced by $m-1$. ■

Notice that in Proposition B.1 we started by *assuming* we had a common eigenstate ψ of J^2 and J_3, but in the process of proving Theorem B.2 and Corollary B.3 we have done much more: for each choice of $j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \ldots$, we have effectively *constructed* finite-dimensional inner product spaces $\mathcal{H}_j$ on which the angular momentum operators J_i act.

C * Angular momentum operators in spherical polar coordinates

In this appendix we prove Proposition 8.8, which is an exercise in using the chain rule. We write $\mathbf{x} = (x, y, z)$ in *spherical polar coordinates* $(x, y, z) = (r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta)$. Calculating as in (8.36), we find

$$\frac{\partial}{\partial\theta} = \frac{\partial x}{\partial\theta}\frac{\partial}{\partial x} + \frac{\partial y}{\partial\theta}\frac{\partial}{\partial y} + \frac{\partial z}{\partial\theta}\frac{\partial}{\partial z} = z\cos\phi\frac{\partial}{\partial x} + z\sin\phi\frac{\partial}{\partial y} - z\tan\theta\frac{\partial}{\partial z} \ , \tag{C.1}$$

where we used $\partial z/\partial\theta = -z\tan\theta$. Similarly we may write (8.36) differently as

$$\frac{\partial}{\partial\phi} = \frac{\partial x}{\partial\phi}\frac{\partial}{\partial x} + \frac{\partial y}{\partial\phi}\frac{\partial}{\partial y} + \frac{\partial z}{\partial\phi}\frac{\partial}{\partial z} = \tan\theta\left(-z\sin\phi\frac{\partial}{\partial x} + z\cos\phi\frac{\partial}{\partial y}\right) \ . \tag{C.2}$$

We then combine these to calculate

$$\frac{\partial}{\partial\theta} \pm \mathrm{i}\cot\theta\frac{\partial}{\partial\phi} = z\,\mathrm{e}^{\mp\mathrm{i}\phi}\frac{\partial}{\partial x} \pm \mathrm{i}z\,\mathrm{e}^{\mp\mathrm{i}\phi}\frac{\partial}{\partial y} - z\tan\theta\frac{\partial}{\partial z} \ , \tag{C.3}$$

where we have used $\mathrm{e}^{\pm\mathrm{i}\phi} = \cos\phi \pm \mathrm{i}\sin\phi$. Thus

$$\pm\mathrm{e}^{\pm\mathrm{i}\phi}\left(\frac{\partial}{\partial\theta} \pm \mathrm{i}\cot\theta\frac{\partial}{\partial\phi}\right) = \mathrm{i}z\frac{\partial}{\partial y} \pm z\frac{\partial}{\partial x} \mp z\tan\theta\,\mathrm{e}^{\pm\mathrm{i}\phi}\frac{\partial}{\partial z} \ . \tag{C.4}$$

On the other hand

$$\begin{aligned} \frac{1}{\hbar}(L_1 \pm \mathrm{i}L_2) &= -\mathrm{i}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) \pm \left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) \\ &= \mathrm{i}z\frac{\partial}{\partial y} \pm z\frac{\partial}{\partial x} \mp (x \pm \mathrm{i}y)\frac{\partial}{\partial z} \ . \end{aligned} \tag{C.5}$$

But $x \pm \mathrm{i}y = r\sin\theta\,\mathrm{e}^{\pm\mathrm{i}\phi} = z\tan\theta\,\mathrm{e}^{\pm\mathrm{i}\phi}$, which shows that (C.4) and (C.5) are the same, thus proving (8.40).

Finally, using (*ii*) from Proposition 8.4 and (8.40), (8.37) we compute

$$\begin{aligned}
L^2 &= L_+L_- + L_3^2 - \hbar L_3 \\
&= \hbar^2\left\{ \mathrm{e}^{\mathrm{i}\phi}\left(\frac{\partial}{\partial\theta} + \mathrm{i}\cot\theta\frac{\partial}{\partial\phi}\right)\left[-\mathrm{e}^{-\mathrm{i}\phi}\left(\frac{\partial}{\partial\theta} - \mathrm{i}\cot\theta\frac{\partial}{\partial\phi}\right)\right] - \frac{\partial^2}{\partial\phi^2} + \mathrm{i}\frac{\partial}{\partial\phi}\right\} \\
&= \hbar^2\left[-\frac{\partial^2}{\partial\theta^2} + \mathrm{i}\frac{\partial\cot\theta}{\partial\theta}\frac{\partial}{\partial\phi} - \cot\theta\left(\frac{\partial}{\partial\theta} - \mathrm{i}\cot\theta\frac{\partial}{\partial\phi}\right) - \cot^2\theta\frac{\partial^2}{\partial\phi^2} - \frac{\partial^2}{\partial\phi^2} + \mathrm{i}\frac{\partial}{\partial\phi}\right] \\
&= -\hbar^2\left(\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right) \ , \qquad\qquad \text{(C.6)}
\end{aligned}$$

where in the last step we used $\mathrm{d}\cot\theta/\mathrm{d}\theta = -1/\sin^2\theta$, and the identity $1/\sin^2\theta = \cot^2\theta + 1$.