Quantum Mechanics

Mathematical Tripos Part IB

Prof. Maria Ubiali

DAMTP, University of Cambridge, Office: B1.16
email: M.Ubiali@damtp.cam.ac.uk

Last update: October 5, 2022

Recommended Books:

These notes follow the notation and take inspiration from the books by S. Gasiorowicz [2], D. J. Griffiths [3] and A. I. M. Rae + J. Napolitano [6].

If you are interested in more advanced references, I suggest you look at the book by Weinberg [7], the one by Sakurai and Napolitano [4] and the ever-green book by Landau and Lifshitz [5].

For an alternative perspective see Chapters 1-3 of the Feynman's lectures on physics [1].

Acknowledgements:

I am grateful to the previous lecturer of this course, Adrian Kent, who shared his lecture notes and suggestions on the material. These notes have also been influenced by the notes by previous lecturers, including Eugene Lim, Jonathan Evans, Nick Dorey. I thank Giovanni Ridolfi for his precious tips and my PhD student James Moore for his extremely useful comments. Any feedback – whether noting typos or other errors or more general comments – is very welcome. The completed lecture notes are intended to be a reasonably complete summary of the course. However, material not covered in the printed notes will be added during lectures, for instance in answer to questions (which are encouraged!) or whenever there is time for further discussion that could be helpful. Non-examinable sections of the lecture notes are marked by asterisks. Further course material will be added from time to time on the course Moodle page.

Contents

Chapter 1

Historical introduction

The development of quantum mechanics in the 1920s was probably the greatest advance in physical science. It was not easy. The principles of quantum mechanics were so contrary to the human intuition built upon classical mechanics that, as S. Weinberg writes, "quantum mechanics has won acceptance through its success" [7]. Indeed, as we shall see along this course, quantum mechanics is essential to modern atomic, molecular, nuclear, and elementary particle physics, and to a great deal of chemistry and condensed matter physics as well. Many of the revolutionary technological developments of the last hundred years rely on quantum mechanics, from semiconductor physics (transistors, diodes, integrated circuits, and hence the computing and IT industries) to all applications of quantum information, including quantum cryptography and quantum authentication.

In this chapter we will consider the problems confronted by physicists in the first years of the 20$^{\text{th}}$ century that ultimately led to modern quantum mechanics. We can summarise them shortly below:
(*)1801-03: Interference/diffraction experiments by Young show that light is a wave
(*)1862-4: Maxwell identifies light as an EM (electromagnetic) wave
(*)1897: Thompson discovers the electron, the first elementary particle
(*)1900: Planck introduces the energy-frequency relation, with h as a new physical constant, and derives the *black body spectrum*
(*)1905: Einstein imparts clearer physical meaning to photons, using them to explain the *photoelectric effect*, and other experimental results
(*) 1909: In a version of the double slit experiment, G.I. Taylor demonstrates that light produces a wave-like interference pattern on photographic film even when the light source is filtered so that only one photon at a time is recorded by the film
(*) 1911: Based on scattering experiments, Rutherford proposes a model of the atom with most of its mass concentrated in a small, compact *nucleus*

(*) 1913: Bohr proposes an atomic model with electrons orbiting a nucleus and with quantisation of their angular momentum, using this to derive observed line spectra

(*) 1923: Compton scattering of X-rays on electrons confirms that photons are relativistic particles of zero rest mass

(*) 1923-24: de Broglie proposes *wave-particle duality* for matter, as for radiation

(*) 1925-30: The emergence of *Quantum Mechanics*, through work of Heisenberg, Born, Jordan, Dirac, Pauli, Schrödinger, and others.

(*) 1927-28: Diffraction experiments of Davisson, Germer and Thompson confirm that electrons behave as waves as well as particles .

From this quick historical overview, we can see that everything started because of a number of impressive improvements in technology and the discovery of radioactivity that enabled physicists to study in some detail the internal structure of atoms, the mass and charge of the electron, and the interaction of atoms with light. Certain aspects of atomic physics which emerged from these early investigations were puzzling and even paradoxical, in the sense that the observed behaviour of electrons, atoms, and light seemed in contradiction to the known laws of mechanics and electromagnetism. These surprising aspects fell roughly into three categories: (i) the particle-like behaviour of light waves, which includes blackbody radiation, the photoelectric effect, the Compton effect, (ii) the puzzling stability of the atom and (iii) the wave-like behaviour of particles.

In this chapter we will first recap the basic concepts of classical mechanics and wave mechanics that are needed in order to be surprised by the evidences that led to quantum mechanics. We then explore in turn each of the three aspects that we just mentioned.

1.1   Particles and waves in classical mechanics

In order to start our journey in understanding quantum mechanics, it is useful to quickly recap what are the basic concepts of classical mechanics. Roughly speaking, in classical mechanics, one can describe motion in terms of either particles or waves. Classically, they are distinctly different things. In our every day life, we intuitively think of some things as particles (like bullets, cars, cats etc) while some other things as waves (sound waves, water waves, radio etc), because they seemingly behave very differently.

Particles are objects that carry energy and momentum in infinitesimally localised points of space, hence one often calls them "point particles". In the classical world, the state of a particle is determined by 3-dimensional vectors indicating its position $\mathbf{x}$ and its velocity $\mathbf{v} = \dot{\mathbf{x}} \equiv d\mathbf{x}/dt$. If you specify both bits of information at some time t_0 then you can use the equation of motion

$$\mathbf{F}(\mathbf{x}(t), \dot{\mathbf{x}}(t)) = m\ddot{\mathbf{x}}(t) \tag{1.1}$$

to determine $\mathbf{x}(t)$ and $\dot{\mathbf{x}}(t)$ for all time. Importantly, it's not enough to just know only, say, the position of the particle at $t = t_0$. You need both $\mathbf{x}(t_0)$ and $\dot{\mathbf{x}}(t_0)$. Mathematically,

this is because the equation of motion is a second order differential equation and so you need to specify two integration constants to get a unique solution.

Waves, on the other hand, describe motion of entities which are spread out and are not localised. For example, sound is carried by the compression and decompression of air, and the displacement of some parcel of air molecules from its original undisturbed position obeys the wave equation. Also, classical electric fields and magnetic fields are described by waves (you will study this in great detail in the Part IB Electromagnetism class).
From a mathematical point of view, we will refer to waves as any real or complex-valued function with periodicity in time and/or space.

- A function of time t obeying $f(t+T) = f(t)$ has *period T*, *frequency $\nu = 1/T$*, and *angular frequency*

$$\omega = 2\pi\nu = 2\pi/T . \tag{1.2}$$

  Familiar examples are $f(t) = \cos\omega t$, $\sin\omega t$ or $\exp(\pm i\omega t)$.

- A function of position x (in one dimension) obeying $f(x+\lambda) = f(x)$ has *wavelength λ* and *wavenumber*

$$k = 2\pi/\lambda \tag{1.3}$$

  Examples are $f(x) = \cos kx$, $\sin kx$ or $\exp(\pm ikx)$.

- The analogous functions of a position vector $\mathbf{x}$ with periodicity in three dimensions are $f(\mathbf{x}) = \exp(i\mathbf{k} \cdot \mathbf{x})$ where $\mathbf{k}$ is the *wave vector*, and the wavelength is then $\lambda = 2\pi/|\mathbf{k}|$. We shall refer to such functions as *plane waves*.

The wave equation in one dimension for a function $f(x,t)$ is

$$\frac{\partial^2 f(x,t)}{\partial t^2} - c^2 \frac{\partial^2 f(x,t)}{\partial x^2} = 0 \tag{1.4}$$

where c is some constant. This has solutions which are periodic in both position and time:

$$f_\pm(x,t) = A_\pm \exp(\pm ikx - i\omega t) , \; A_\pm \in \mathbb{C}. \tag{1.5}$$

provided that the wavelength and frequency are related by

$$\omega = ck \quad \text{or} \quad \lambda\nu = c . \tag{1.6}$$

$$\text{Dispersion relation}$$

As in Newton's law of motion, Eq. (1.1), the wave equation is a 2nd order in time derivative: this means that we need to specify two initial conditions $f(x,0)$ and $f'(x,0)$ to get a unique solution, as it occurs in particle dynamics.
From a physics point of view, such solutions represent waves which move or *propagate* with speed c to the right or left, according to the sign in tEq. (1.5) (assuming ω, $k > 0$). The

constant $A_\pm$ is the *amplitude* of the wave. The three-dimensional wave equation is obtained by replacing $\partial^2/\partial x^2$ by ∇^2 in Eq. (1.4). This has solutions of an analogous form

$$f(\mathbf{x}, t) \,=\, A \exp(\,i\mathbf{k}\cdot\mathbf{x} \,-\, i\omega t\,) \qquad \text{with} \qquad \omega = c|\mathbf{k}|\ . \tag{1.7}$$

Such a wave propagates in the direction of $\mathbf{k}$, with speed c. For electromagnetic waves $c \approx 3 \cdot 10^8$ m/s is the speed of light.

There is an important remark: other kinds of waves arise as solutions of other governing equations which may differ significantly from the standard wave equation. A function does not have to satisfy the standard wave equation in order to be usefully thought of as a wave! The Schrödinger Equation is one example of an alternative governing equation; it is the central equation in QM and we will study it in some depth in the next chapter. Different governing equations give rise to propagating solutions, provided the frequency is chosen to be a suitable function of the wavenumber, $\omega(k)$. Moreover, if the governing equation is *linear* in f, then any solutions f_1 and f_2 can be combined to give a new solution:

$$f = f_1 + f_2 \tag{1.8}$$

This is the *Principle of Superposition* and it is responsible for much behaviour we tend to think of as *wave-like*.

Interference (or *diffraction*) occurs when waves from different sources merge, or when parts of a wave recombine after passing around or through some obstacle. When a number of such waves are superposed, they may interfere *constructively*, increasing the size of the amplitude, or *destructively*, diminishing the amplitude. The result is an interference or diffraction pattern which depends on the sources or the obstacles. When light is passed through a number of narrow slits, the resulting diffraction pattern provides conclusive evidence that light is a wave. Passing higher energy waves, such as X-rays, through matter gives a way of determining the crystalline arrangement of atoms from the resulting diffraction patterns.

1.2   The particle-like behaviour of light waves

In this section we briefly discuss the first set of puzzling evidences about the behaviour of light, that seemed in contradiction to the known laws of mechanics and electromagnetism, in particular we discuss the evidences that displayed a particle-like behaviour of light waves, which includes black-body radiation, the photoelectric effect and the Compton effect.

1.2.1   Black-body radiation

The blackbody radiation measurements, pioneered by G. Kirchhoff in 1859, have shown that in the thermal equilibrium, the power of electromagnetic radiation by a fully absorbing ("black") surface, per unit frequency interval, drops exponentially at high frequencies.

This is not what could be expected from the combination of classical electrodynamics and statistics, which predicted an infinite growth of the radiation density with frequency. Indeed, classical electrodynamics shows that electromagnetic field modes evolve in time just as harmonic oscillators, and that the number dN of these modes in a large free-space volume $V \gg \lambda^3$, within a small frequency interval $d\omega \ll \omega$ near some ω is

$$dN = 2V \frac{d^3 \mathbf{k}}{(2\pi)^3} = 2V \frac{4\pi k^2 \, dk}{(2\pi)^3} = V \frac{\omega^2}{\pi^2 c^3} d\omega \tag{1.9}$$

where the factor of 2 accounts for the two possible polarisation states of the electromagnetic waves, c is the speed of light, $k = |\mathbf{k}| = \omega/c$ is the modulus of the wave number, and $\lambda = 2\pi/k$ is the radiation wavelength. On the other hand, classical statistics predicts that in the thermal equilibrium at temperature T, the average energy E of each 1D harmonic oscillator should be equal to $k_B T$, where k_B is the Boltzmann constant, $k_B \sim 1.4 10^{-16} \mathrm{erg}/K$. Combining these two results, we readily get the so-called Rayleigh-Jeans formula[1] for the average electromagnetic wave energy per unit volume:

$$\rho(\omega) \equiv \frac{1}{V} \frac{dE}{d\omega} = \frac{k_B T}{V} \frac{dN}{d\omega} = \frac{\omega^2}{\pi^2 c^3} k_B T. \tag{1.10}$$

The prediction that $\rho(\omega)$ is proportional to ω^2 was actually in agreement with observation for small values of ω, but it fails badly for larger values, see Fig. 1.1. Indeed, it diverges at $\omega \to \infty$ – the so-called ultraviolet catastrophe. On the other hand, the blackbody radiation measurements were compatible with the phenomenological law suggested in 1900 by Max Planck:

$$\rho(\omega) = \frac{\omega^2}{\pi^2 c^3} \frac{\hbar \omega}{\exp(\hbar \omega / k_B T) - 1}, \tag{1.11}$$

where $\hbar = h/(2\pi)$ was a constant, which was fitted to give

$$h \sim 6.626 \cdot 10^{-34} \mathrm{Joule} \times \sec \qquad \hbar = \frac{h}{2\pi} \sim 1.055 \cdot 10^{-34} \, \mathrm{Joule} \times \sec, \tag{1.12}$$

where h is the Planck constant, while $\hbar$ is often referred to as *reduced Planck constant*[2]. The dimension of the new constant h, are

$$[h] = M \, L^2 \, T^{-1} = [\text{energy}] \times [\text{time}] = [\text{position}] \times [\text{momentum}]. \tag{1.13}$$

This is the same dimension as angular momentum and of the action in classical mechanics

Although at the beginning it appeared simply as a constant in a fit of a set of data, we now think of this constant as representing the "strength" of quantum effects. Despite having

---

[1] A numerical error in Rayleigh's derivation was corrected by his student J. Jeans in 1905.

[2] Older books on quantum mechanics often like to use h, but in these lectures we will use $\hbar$, as most modern physicists now do.
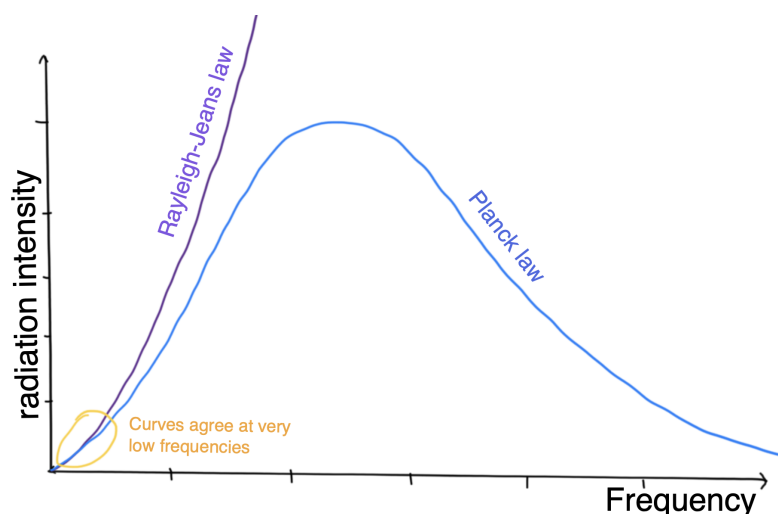
Figure 1.1:   Schematic illustration of the Rayleigh-Jeans prediction versus the Planck prediction for the intensity of radiation emitted by a black body at equilibrium at a given temperature T as a function of the emitted light angular frequency ω.

these new profound features, early 20th century physicists were guided by the expectation to recover classical physics in limit $\hbar \to 0$. Indeed, in this limit, the Planck law reduces exactly to the Rayleigh-Jeans law.

After the initial formula to fit the data, Planck later gave a derivation of the above formula, based on the assumption that the radiation inside the back-body was quantised, *i.e.* the energy of the radiation comes in integer multiples of $\hbar\omega$. A simpler derivation compared to the original by Planck is given by Weinberg in Ref. [?], following the derivation of the black-body formula by Hendrik Lorentz in 1910. If you are curious, you can read about it in the Appendix of these notes.

Despite the success of Planck's description of the data, there remained considerable skepticism about the reality of light quanta. It could have just been a mathematical trick to derive the desired result. It was only thanks to further experimental evidences including the photoelectric effect and the Compton scattering that the existence of these light quanta was accepted.

1.2.2   The photoelectric effect

In the late 19th century, experiments took place, in which a metal surface was hit in the vacuum with monochromatic light. Changing the frequency and the intensity of the incident light, it was observed that at certain frequencies, electrons were emitted from the metal surface, as in Fig. 1.2. This effect is known as the *photoelectric effect*. Surprisingly,

whether the electrons were emitted or not depended only on the frequency, and not on the intensity of the incident light, while the rate of emitted electrons depended on its intensity!

The observed effects were hard to explain in terms of the electromagnetic wave model of light. If we think of an incident wave transmitting energy to the electrons and knocking them out of the metal, we would expect the rate of electron emission to be constant, and we would also expect that light of any frequency would eventually transmit enough energy to electrons to cause some of them to be emitted, as the intensity of light $I = |A|^2$ increases. But this was not what was observed. On the contrary, to emit any electrons, the incident light needed to have an angular frequency ω (with $\omega = 2\pi c/\lambda$, where λ is the wavelength of the incident light) above a given threshold, depending on the particular metal. When the angular frequency was large enough to emit electrons from the metal surface, the maximum energy of the emitted electrons, $E_{\max}$, was found to be proportional to ω and to obey

$$E_{\max} = \hbar\omega - \Phi, \tag{1.14}$$

where Φ is the so-called work function of the metal and the constant of proportionality $\hbar$ is the reduced Planck constant that we encountered in the previous section.

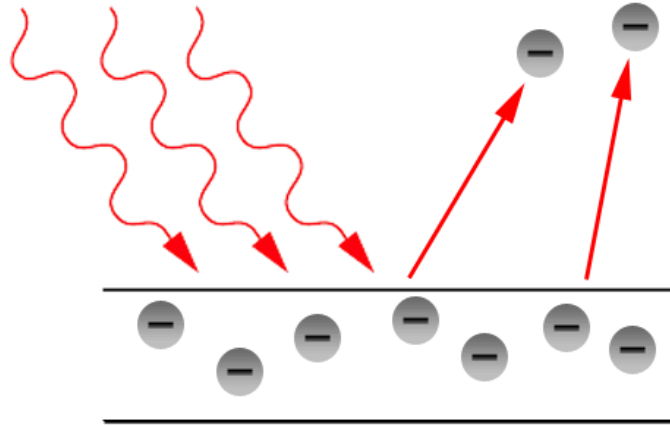Figure 1.2: Schematic illustration of the photoelectric effect. [*Licensed under Wikimedia Commons*]

To explain the photoelectric effect, Albert Einstein (in 1905) was led to postulate that light is quantised in small light quanta called *photons*, and that a photon of angular frequency ω has an energy proportional to ω, $E = \hbar\omega$. The photoelectric effect can be explained as the

result of single photons hitting single electrons near the metal surface, if one assumes that an electron needs to acquire a kinetic energy $\geq \Phi$ to overcome the binding energy of the metal. An electron which acquires an energy $\hbar\omega$ thus carries away an energy that is less or equal to $\hbar\omega - \Phi = E_{\max}$. This allows one to explain the emission rate observations: the average rate of photon arrival is proportional to the intensity of the light, and the rate of emission of electrons is proportional to the rate of scattering with photons. Therefore, the number of electrons emitted is proportional to the intensity of the incident light I.

1.2.3   The Compton scattering

A further evidence of the quantised nature of light was obtained nearly twenty years later, in 1923, when Arthur H. Compton observed the scattering of X-rays by electrons associated with atoms in a crystal, the *Compton scattering*. Because the X-ray energies were much
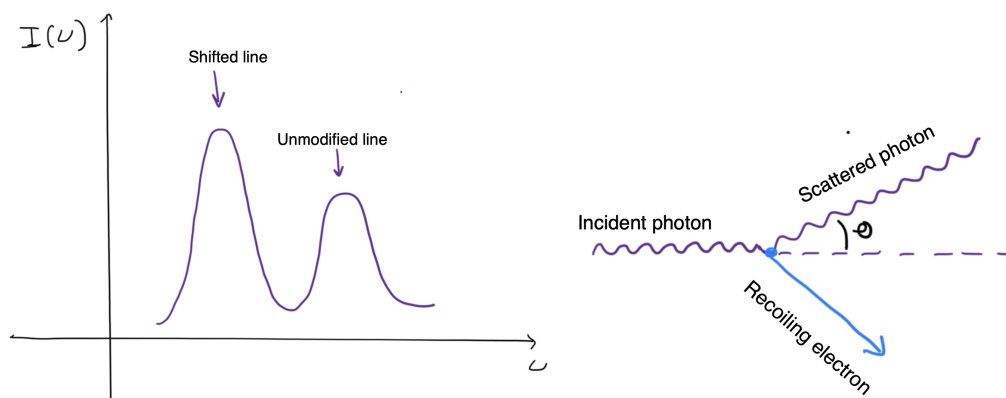
Figure 1.3:   Left: the spectrum of radiation scattered by carbon at a given angle θ, showing the unmodified line on the right and the shifted line on the left. Right: kinematics for the Compton effect.

larger than the electron binding energies, the electrons can effectively be modelled as free electrons. According to classical theory, the mechanism that explains it is the re-radiation of light by the electrons, which are set into forced oscillations by the incident radiation, and this leads to prediction of intensity observed at an angle θ that varies as $(1 + \cos^2 \theta)$ and has the frequency as the incident radiation. Compton found that the radiation scattered through a given angle actually consists of two components: the expected one, whose frequency is the same as that of the incident radiation, and the unexpected one, whose frequency was shifted relative to the incident frequency by an amount that depended on the angle θ (see left panel in Fig. 1.3 for a schematic illustration). Compton was able to explain the angle-dependent component by treating the incoming radiation as a beam of photons of energy $\hbar\omega$. Thus, the scattering results from collisions between a single photon in one beam and a single electron in the other, in which energy and momentum

are transferred between the photon and the electron. (A relativistic treatment of this scattering process was given in the IA Dynamics and Relativity course, see right panel of Fig. 1.3 to recollect it.) This explanation is consistent with the observed scattering data and with conservation of (relativistic) energy and momentum. Provided we assume that a photon of angular frequency ω has a definite momentum $\mathbf{p} = \hbar\mathbf{k}$, where $\mathbf{k}$ is the wavevector of the corresponding electromagnetic wave, so that $|\mathbf{p}| = \hbar|\mathbf{k}| = \hbar\omega/c$, then

$$\frac{1}{\omega'} = \frac{1}{\omega} + \frac{\hbar}{mc^2}(1 - \cos\theta). \tag{1.15}$$

Again, in the limit of $\hbar \to 0$, the result tends to the one predicted by classical mechanics. This is yet another example of correspondence principle. Compton's experiment convinced physicists that light can be treated as a stream of particle-like objects, the photons, whose energy is proportional to the light wave's frequency.

1.3   Atomic spectra

After discovering the electron in 1897, J.J. Thomson proposed a model of the atom as a sort of "plum pudding" with Z point-like electrons of charge $-e$ embedded in a sphere of positive charge $+Ze$. In 1908, Geiger and Marsden's famous experiment, carried out at
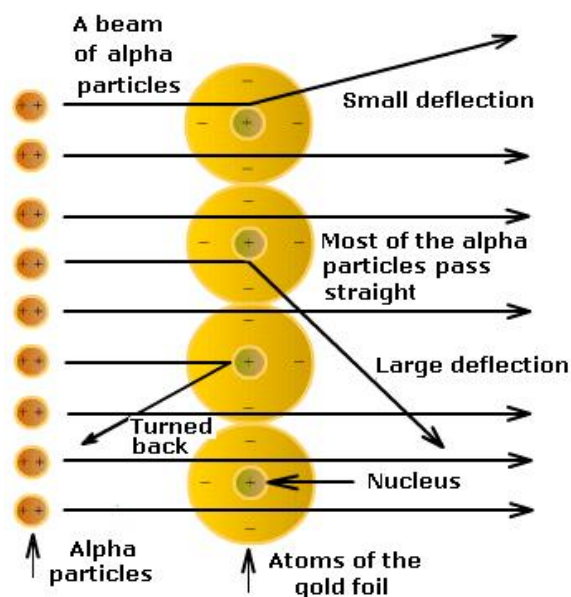

Figure 1.4: Schematic illustration of the Rutherford experiment. From http://physicsopenlab.org/2017/04/11/the-rutherford-geiger-marsden-experiment.

Rutherford's suggestion, tested for large angle scattering of a beam of α-particles directed at gold foil. One would not expect significant scattering from a loosely distributed low charge density "plum pudding" atom. Instead some α-particles were observed to be scattered through very large angles of up to 180°. In Rutherford's famous phrase, "It was as if you fired a 15-inch shell at tissue paper and it came back and hit you." See Fig. 1.4 for a schematic illustration. The scattering suggests a high density positive charge within the atom. Rutherford thus postulated a new model of the atom, with a heavy nearly point-like nucleus, of charge $+Ze$, surrounded by Z electrons in orbit.

Although the Rutherford atom was more compatible with the observed scattering data than was the "plum pudding" model, it had a number of theoretical defects. First, according to Maxwell's electrodynamics, electrons in orbit around a nucleus would radiate, since they are continually undergoing acceleration[3]. This would cause them to lose energy and fall in towards the nucleus. Stationary electrons would also fall into the nucleus because of electrostatic attraction. This would suggest that atoms must be unstable, which they generally are not. Second, the model fails to explain why atoms have characteristic line spectra corresponding to discrete frequencies at which they absorb or emit light. For example, hydrogen has frequencies given by the Rydberg formula (Rydberg, 1890):

$$\omega_{mn} = 2\pi c R_0 \left(\frac{1}{n^2} - \frac{1}{m^2} \right) \qquad \text{for} \quad m > n, \tag{1.16}$$

with R_0 being Rydberg constant $R_0 = 1.097 \times 10^7 \text{m}^{-1}$ and the frequency ω_{mn} corresponding to the line spectra for the decay from the atomic state m to the atomic state n. Third, it fails to explain why atoms belong to a finite number of chemical species, with all members of the same species behaving identically. For instance, if a hydrogen atom can have an electron in any type of orbit around its nucleus, one would expect infinitely many different types of hydrogen atom, corresponding to the infinitely many different possible orbits, and one would expect the atoms to have different physical and chemical properties, depending on the details of the orbit.

In 1913 Niels Bohr observed that these problems could be resolved in a way consistent with Planck's and Einstein's earlier postulates, if we suppose that the electron orbits of hydrogen atoms are quantised so that the orbital angular momentum takes one of a discrete set of values

$$L = n\hbar \qquad \text{with} \quad n \in \mathbb{N},\, n > 0. \tag{1.17}$$

As a consequence of the angular momentum quantisation, if we take an electron e moving with velocity v in a circular orbit of radius r about a proton, $F = m_e a$ gives us that the Coulomb force:

$$F_{\text{Coulomb}} = \frac{e^2}{4\pi\epsilon_0} \frac{1}{r^2} = m_e \frac{v^2}{r}. \tag{1.18}$$

---

[3]Physics teaches us that any charged particle travelling in curved paths emits the so-called synchrotron radiation, thus it loses energies.

In a circular motion, the angular momentum is given by

$$L = m_e v r, \tag{1.19}$$

but, because of Eq. (1.17), we can derive

$$v \equiv v_n = \frac{n\hbar}{m_e r}. \tag{1.20}$$

Plugging it into Eq. (1.18) and solving for r, we find

$$r \equiv r_n = n^2 \left(\frac{4\pi\epsilon_0}{m_e e^2} \hbar^2 \right) \equiv n^2 a_0, \tag{1.21}$$

where

$$a_0 = \frac{4\pi\epsilon_0 \hbar^2}{m_e e^2} \sim 0.53^{-10} m \tag{1.22}$$

is the Bohr radius. So, as a result of the quantisation of the angular momentum, also the radius and the velocity of the electrons, thus their energy in each orbit are quantised. The energy is given by

$$E_n = \frac{1}{2} m_e v_n^2 - \frac{e^2}{4\pi\epsilon_0} \frac{1}{r_n} = -\frac{e^2}{8\pi\epsilon_0 a_0} \frac{1}{n^2} = -\frac{e^4 m_e}{32\pi^2 \epsilon_0^2 \hbar^2} \frac{1}{n^2} \equiv \frac{E_1}{n^2}, \tag{1.23}$$

where $E_1 = -(e^4 m_e)/(32\pi^2 \epsilon_0^2 \hbar^2) \sim -13.6$ eV. Thus we have $n = 1$ with energy $E = E_1$ defining the lowest possible energy state, or ground state, of the Bohr atom. The higher energy excited states, so called because they can be created by "exciting" the ground state atom with radiation, correspond to $n > 1$. These can decay to the ground state: the ground state has no lower energy state to decay to, and so is stable. (The Bohr model does not allow a state with zero orbital angular momentum, which would correspond to $n = 0$, $r = 0$ and $E = -\infty$.)

The energy emitted for a transition from the m-th to the n-th Bohr orbital is $E_{mn} = E_m - E_n$, see Fig. 1.5 for a schematic illustration. Using $E_{mn} = \hbar\omega_{mn}$, where ω_{mn} is the angular frequency of the emitted photon, we have

$$\omega_{mn} = 2\pi c R_0 \left(\frac{1}{n^2} - \frac{1}{m^2} \right), \tag{1.24}$$

where

$$R_0 = \frac{m_e c}{2\hbar} \left(\frac{e^2}{4\pi\epsilon_0 \hbar c} \right)^2 \tag{1.25}$$

agrees well with the experimentally determined values of the Rydberg constant.
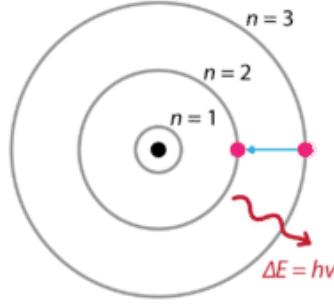
15

Figure 1.5:   Transition between energy levels in the Bohr atomic model. *Licensed under Creative Commons*

Bohr made great use of the notion that the quantum theory that he had developed should merge into classical theory in the limit in which classical theory was known to apply. This idea was formulate as *correspondence principle.* Technically it stated that the classical limit should be reached when the quantum numbers are large, for example, in the atomic model, for $n \to \infty$. To illustrate it better, consider the frequency emitted when an electron makes a "jump" from the orbit with a quantum number $n+1$ to the orbit with quantum number n, where n is very large. This is a good domain to ask for classical limit as the angular momentum $\hbar n$ is indeed much larger than $\hbar$. Classically, an electron moving in a circular orbit with velocity v would be expected to radiate with the frequency of its motion; that is

$$\nu_{\text{cl}} = \frac{v}{2\pi r} = \frac{n\hbar}{m_e} \left(\frac{m_e^2 e^2}{8\pi^2 \epsilon_0 \hbar^2 n^2} \right)^2 = \frac{m_e^3 e^4}{64\pi^4 \epsilon_0^2 \hbar^3} \frac{1}{n^3}. \tag{1.26}$$

On the other hand, the frequency of the radiation associated with the transition is, according to Eq. (1.24),

$$\nu = \frac{\omega}{2\pi} = \frac{m_e^3 e^4}{128\pi^4 \epsilon_0^2 \hbar^3} \left(\frac{1}{n^2} - \frac{1}{(n+1)^2} \right), \tag{1.27}$$

which approaches ν_{cl} for $n \gg 1$. This is a significant result, since it is only the frequency associated with a $n+1 \to n$ transition that corresponds to the fundamental classical theory. The radiation associated with the $n+2 \to n$ transition does not have classical counterparts even in the large n limit. Indeed, there are no $n+2 \to n$ transitions for circular orbits in quantum mechanics, as we shall see in Chapt. 5.

Bohr's model of the atom was rather more successful than its predecessors. It predicts the energy levels of the hydrogen atom and the spectrum of photons emitted and absorbed. It also accounted for spectroscopic data for ionised helium ($He+$) and for some emission and absorption spectra for other atoms. However, as Bohr himself stressed, the model offered no explanation of atomic physics. For example, as Rutherford commented, it's quite mysterious that an electron which jumps from the m-th to the n-th orbit seems to know in

16

advance what frequency to radiate at during the transition. Moreover, the Bohr model is quite wrong about the details of electron orbits, even for the hydrogen atom. Nonetheless, it was an important stepping stone on the path to quantum mechanics, suggesting some link between Planck's constant, atomic spectra and atomic structure, and the quantisation of angular momentum and other dynamical quantities.

1.4   The wave-like behaviour of particles

After the interpretation of the photoelectric effect and of the Compton scattering, light, that even since Maxwell had been understood to be a wave of electric and magnetic fields, was also manifested in a particle, the photon. As a result, a question arose quite naturally: is it possible that something like the electron, that had always been regarded as a particle, could also be manifested as some sort of wave? This was suggested in 1923 by Louis de Broglie, a doctoral student in Paris. Any kind of wave of angular frequency ω and wave number $\mathbf{k}$ has a spacetime dependence $\exp(i\mathbf{k} \cdot \mathbf{x} - i\omega t)$. Lorentz invariance requires that $(\omega, \mathbf{k})$ transforms as a four-vector, just like the momentum four-vector $(E, \mathbf{p})$. For light, according to Einstein, the energy of a photon is $E = \hbar\omega$ and its momentum has a magnitude $|\mathbf{p}| = E/c = h\nu/c = |\mathbf{k}|$, so de Broglie was led to suggest that in general a particle of any mass is associated with a wave having the four-vector $(\omega, \mathbf{k})$ equal to $1/\hbar$ times the four-vector $(E, \mathbf{p})$:

$$\omega = \frac{E}{\hbar} \qquad \mathbf{k} = \frac{\mathbf{p}}{\hbar}. \tag{1.28}$$

Just as vibrational waves on a violin string are quantised by the condition that, since the string is clamped at both ends, it must contain an integer number of half-wavelengths, so an important consequence of de Broglie's hypothesis is that the wave associated with an electron in a circular orbit of radius r must have a wavelength that just fits into the orbit a whole number n of times, so $2\pi r = n\lambda$, and therefore $p = \hbar k = \hbar \times 2\pi/\lambda = n\hbar/r$, thus the quantisation of the angular momentum In a sense, De Broglie's hypothesis, although *a posteriori*, made the orbital angular momentum $L = rp = n\hbar$. somewhat less mysterious.

If electrons really are waves, then they should exhibit the usual behaviour of waves, such as diffraction and interference. In the late 19$^{\text{th}}$ century double-slit experiments had been carried out with many different sources of light. At different points constructive interference or destructive interference were observed. The same experiment was repeated on electrons. Not only this experiment allowed scientists to verify whether electrons were actually behaving like waves. They could also figure out the electron's wavelength λ by means of an experiment. Practically, the actual experiment for electrons is slightly more complicated. Since the wavelength of an electron is rather small, to obtain the diffraction pattern, it was not enough to poke holes in sheets. Instead, crystals had to be used as

diffraction grating. Basically, when waves are scattered by a period structure, there will be a phase difference between waves coming from adjacent scattering planes. The phase difference is given by $(2\pi/\lambda)2a\sin\theta$, and there will be constructive interference whenever the phase difference is equal to $2\pi n$, with n an integer. The results $n\lambda = 2a\sin\theta$ is called the Bragg condition, because it was discovered by W. L. Bragg in connection with his work on the scattering of X rays by crystals. These experiments carried out on electron beams showed that electrons do diffract, and their wavelength is exactly the de Broglie wavelength. This was first confirmed in 1923-7 by Davisson and Germer and independently
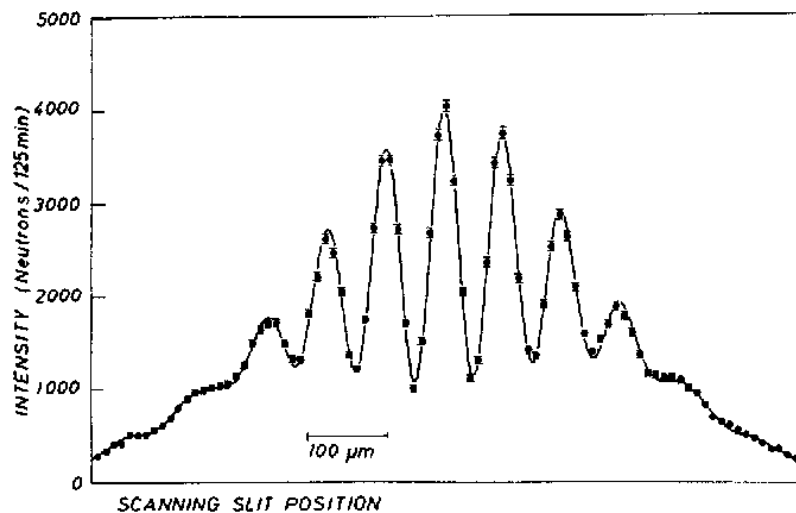
Figure 1.6:   Double slit diffraction pattern for neutron with wavelength $\lambda \sim 18.5$ A. *From A. Zeiliger, Rev. Mod. Phys. **71**:S288 (1999)*

in 1927 by G.P. Thomson, who observed diffractive scattering of electrons from metallic crystals, with diffraction patterns consistent with the de Broglie wavelength $\lambda = 2\pi\hbar/p$ [4]. The interference pattern observed in electron scattering could be correlated with the Bragg condition, provided the De Broglie relation. This constituted a major development in quantum mechanics. Many diffraction experiments with electrons, neutrons and other particles have since been carried out, all confirming de Broglie's prediction. A particularly interesting experiment was carried out by A. Zeiliger and collaborators with slow neutrons and with molecular beams of hydrogen and helium, see Fig. 1.6 for the clear diffractive pattern that was observed.

These experiments also have a conceptual importance. For regular waves, diffraction is

---

[4]G.P. Thomson was the son of J.J. Thomson, who in 1897 discovered the electron, in experiments in which it behaves as (and so was then understood to be) a particle. It is a pleasing historical quirk that G.P. Thomson was the co-discoverer of the wave-like behaviour of electrons in diffraction experiments.

something we can make sense of. However, here we are talking about electrons. We know that if we fire many many electrons, the distribution will follow the pattern described above. But what if we just fire a single electron? On average, it should still follow the distribution. However, for this individual electron, we cannot know where it will actually land. We can only provide a probability distribution of where it will end up. In quantum mechanics, everything is inherently probabilistic. In this respect, quantum mechanics is vastly different from classical mechanics.

1.5   Conclusions

As we have seen in this short historical introduction, quantum mechanics introduces a completely novel point of view in the way we see Nature. In particular it recasts the concepts of *predictivity* and *reality*. Indeed, one of the principles of quantum mechanics is that physics cannot, in general, predict future events, rather it can only estimate the probability for them to happen. On the other hand, as we have mentioned talking about electron diffraction and as we will understand more in depth in the course of these lectures, quantum mechanics is incompatible with the principle of *local realism*, *i.e.* with the idea that the results of experimental measurements are the manifestation of intrinsic features of the measured objects themselves *before* being measured.

The underlying idea of a fundamental physics theory is also radically transformed. Unavoidably, the new concepts of quantum mechanics overcome a *mechanical reductionism* that led Maxwell to hope that – analogously to thermodynamics that can be reduced to statistical mechanics – electromagnetism could also be reduced to classical mechanics. The hypothesis that, thanks to some universal principles, the whole realm of physics can be reduced to the principles of classical mechanics is incompatible with quantum mechanics. Rather, quantum mechanics does not provide a new universal theory, it rather provides a new universal framework, *i.e.* a set of principles that any theory must satisfy.

One of the reasons why quantum mechanics is considered incomprehensible (so much so that Feynman observed "I think I can safely say that nobody really understands quantum mechanics") is that for some time people tried in vain to understand it in terms of classical concepts, rather than realise that the new quantum framework cannot be re-conduced to any of the classical concepts. For example, some textbooks summarise what we described in this chapter by saying that electrons (photons, etc.) exhibit something called "wave-particle duality". This term can mislead, if it is interpreted as a sort of explanation of what is going on. The fact is that our classical wave and particle models are fundamentally inadequate descriptions. It isn't correct to say that an electron (or photon, etc.) is both a wave and a particle in the classical sense of either of those words. The electron is something different again, though it has some features in common with both. To grasp the depth of nature, we need a new grammar to scribe it, the grammar of quantum mechanics, which

you will become more familiar with by the end of the course.

Chapter 2

Foundations of quantum mechanics

In this section, starting from the motion of a non-relativistic particle in a continuous coordinate space, we will introduce the basic postulates and the basic formalism of quantum mechanics. Although ultimately only a non-relativistic part of a bigger theory, quantum mechanics already teaches us that our universe follows laws that involve beautiful and intricate mathematics.

2.1 Wavefunctions and probabilistic interpretation

To lay the mathematical foundation of quantum mechanics, let us look at a single particle and ask "how do we describe a quantum mechanical particle?". Classically, a point particle has a definitive position $\mathbf{x}$ (and momentum $\mathbf{p} = m\dot{\mathbf{x}}$) at each time. To completely specify a particle, it suffices to write down these two vectors. In quantum mechanics, this is more complicated.

In quantum mechanics, a particle is described by a *state* ψ. The state of a particle is determined by its wavefunction. Note that the notion of the *state* ψ as an individual entity, is not to be conflated with the notion of *wavefunction* $\psi(\mathbf{x}, t)$. In quantum mechanics, $\psi(\mathbf{x}, t)$ is actually the complex coefficient of the state in the continuous and complete basis of $\mathbf{x}$. In other words, we say that $\psi(\mathbf{x}, t)$ is the state ψ expressed in the $\mathbf{x}$ representation. We could also represent the state ψ in the $\mathbf{p}$ momentum representation or in other representations. As far as IB Quantum Mechanics is concerned we will always work in the $\mathbf{x}$ representation, therefore the concept of state and wavefunction can be safely identified. [1].

---

[1] In Part II Principles of Quantum Mechanics you will reformulate QM in a more powerful, abstract and flexible formalism, the Dirac formalism, which clarifies the concept of states and of their representation in

<u>Definition:</u> The wavefunction $\psi(\mathbf{x}, t)$ is a complex-valued function

$$\psi : \qquad \mathbb{R}^3 \qquad \rightarrow \qquad \mathbb{C}, \tag{2.1}$$

that satisfies a number of mathematical property, as we shall see very shortly. As we will see, if you know the wavefunction at some time, say t_0, then you have all the information that you need to determine the state at all other times. The description in terms of the wavefunction is not a small amendment to classical mechanics. The three position coordinates $\mathbf{x} \in \mathbb{R}^3$ are replaced with an infinite amount of information. Moreover, we haven't specified anything about the particle's velocity; that information must also be contained, in some manner, in the wavefunction $\psi(\mathbf{x}, t)$.

The wavefunction has a very simple interpretation. Or, more precisely, the modulus square of the wavefunction has a very simple interpretation. It tells us the probability that we will find a particle at a given position. The probability density ρ for a particle to sit at point $\mathbf{x}$ at a given time t is given by

$$\rho(\mathbf{x}, t) = |\psi(\mathbf{x}, t)|^2, \tag{2.2}$$

and from the probability density, you can compute actual probabilities by multiplying by a volume: the probability that the particle sits in some small volume dV centred around point $\mathbf{x}$ is $\rho(\mathbf{x}, t)dV$. This is knows as *Born's rule*.

2.1.1   Normalisation

The probabilistic interpretation dictates the mathematical properties of wavefunction. The particle has to be somewhere, and this translates to the requirement that the wavefunction has to be normalisable (or square integrable) over all space $\mathbb{R}^3$

$$\int_{\mathbb{R}^3} \psi^*(\mathbf{x}, t)\, \psi(\mathbf{x}, t)dV \;=\; \int_{\mathbb{R}^3} |\psi(\mathbf{x}, t)|^2\, dV = \mathcal{N} < \infty, \tag{2.3}$$

with $\mathcal{N} \neq 0$, because the null state is not a state of a physical system. Given that the total probability has be one, we can always define normalised wavefunction as

$$\bar{\psi}(\mathbf{x}, t) = \frac{1}{\sqrt{\mathcal{N}}}\psi(\mathbf{x}, t), \tag{2.4}$$

so that

$$\int_{\mathbb{R}^3} |\bar{\psi}(\mathbf{x}, t)|^2\, dV = 1. \tag{2.5}$$

This means that

$$\rho(\mathbf{x}, t) = |\bar{\psi}(\mathbf{x}, t)|^2 \in \mathbb{R}, \tag{2.6}$$

---

different bases.

is a probability density function in the usual sense.

From now on, we will drop bars to denote normalised wavefunctions, as it will be clear from the context whether wavefunctions are normalised or not. Quite often in these lectures, it will turn out the be useful to work with un-normalised wavefunctions and then remember to include the normalisation factor only at the end when computing probabilities.

There is one other relation between wavefunctions that is important: two wavefunctions that differ by a constant, complex phase should actually be viewed as describing equivalent states

$$\psi(\mathbf{x}, t) \equiv e^{i\alpha}\psi(\mathbf{x}, t) \tag{2.7}$$

for any constant $\alpha \in \mathbb{R}$. Note, in particular, that this doesn't change the probability distribution $|\psi|^2$. Nor, it will turn out, does it change anything else either[2]. Combining the need for normalisation, together with the phase ambiguity, is sometimes useful think of states as the collection of normalisable, complex functions with the equivalence relation

$$\psi(\mathbf{x}, t) = \lambda\psi(\mathbf{x}, t) \tag{2.8}$$

for any non-zero complex number λ. A more precise statement is that states correspond to rays in the vector space of the wavefunctions. A ray, $[\psi]$, is an equivalence class of a vector ψ under the equivalence relation $\psi_1 \sim \psi_2 \Leftrightarrow \psi_1 = \lambda\psi_2$ for some $\lambda \in \mathbb{C} - \{0\}$.

2.1.2   The superposition principle

The set of wavefunctions form a vector space. This means that if ψ_1 and ψ_2 correspond to allowed states of the system then so does the state ψ, which is defined as

$$\psi = a_1\psi_1 + a_2\psi_2 \neq 0 \tag{2.9}$$

for arbitrary complex numbers a_1 and a_2. This is known as the superposition principle.

The superposition principle is an essential feature of quantum mechanics, which does not generally apply in classical physics. It makes no sense in Newtonian mechanics to add a linear combination of two orbits of a planet around the Sun: this doesn't define another possible orbit. But in quantum theory, taking sums of physical wavefunctions, for example those of an electron orbiting the nucleus of a hydrogen atom, gives us another wavefunction that has a direct physical meaning.

As we have just mentioned, to obtain a sensible probability distribution and make physical predictions, we need to normalise a wavefunction. So to make physical predictions from a superposition, we generally need to normalise $\psi(\mathbf{x}, t)$.

---

[2]The "anything else" argument is important here. As we'll see later, if you multiplied the wavefunction by a spatially varying phase $e^{i\alpha(x)}$ then it doesn't change the probability distribution but it does change other observable quantities and so multiplying by such a factor does not give back the same state.

<u>Theorem 2.1:</u> If $\psi_1(\mathbf{x}, t)$ and $\psi_2(\mathbf{x}, t)$ are both normalisable,

$$\int_{\mathbb{R}^3} |\psi_1(\mathbf{x}, t)|^2 \, dV = N_1 < \infty, \qquad \int_{\mathbb{R}^3} |\psi_2(\mathbf{x}, t)|^2 \, dV = N_2 < \infty. \qquad (2.10)$$

then their linear combination is also normalisable.

Proof: For any two complex number z_1 and z_2, the triangle inequality states that

$$|z_1 + z_2| \leq |z_1| + |z_2|, \qquad (2.11)$$

and

$$(|z_1| - |z_2|)^2 \geq 0 \Rightarrow 2|z_1||z_2| \leq |z_1|^2 + |z_2|^2. \qquad (2.12)$$

If we apply these relations for $z_1 = a_1 \psi_1$ and $z_2 = a_2 \psi_2$, we get

$$
\begin{aligned}
\int_{\mathbb{R}^3} |\psi(\mathbf{x}, t)|^2 dV &= \int_{\mathbb{R}^3} |a_1 \psi_1(\mathbf{x}, t) + a_2 \psi_2(\mathbf{x}, t)|^2 \, dV \\
&\leq \int_{\mathbb{R}^3} (|a_1 \psi_1(\mathbf{x}, t)| + |a_2 \psi_2(\mathbf{x}, t)|)^2 \, dV \\
&= \int_{\mathbb{R}^3} (|a_1 \psi_1(\mathbf{x}, t)|^2 + |a_2 \psi_2(\mathbf{x}, t)|^2 + 2|a_1 \psi_1(\mathbf{x}, t)||a_2 \psi_2(\mathbf{x}, t)|) \, dV \\
&\leq \int_{\mathbb{R}^3} (2|a_1 \psi_1(\mathbf{x}, t)|^2 + 2|a_2 \psi_2(\mathbf{x}, t)|^2) \, dV \\
&= 2|a_1|^2 N_1 + 2|a_2|^2 N_2 < \infty. \qquad \square
\end{aligned}
\qquad (2.13)
$$

2.2   Hilbert space and inner product

From the mathematical point of view, wavefunctions satisfy the defining conditions for abstract vectors, and, as we will see, operators act on them as linear transformations. As a result, the natural language of quantum mechanics is linear algebra.

However it is not a form of linear algebra, which you are immediately familiar with. In an n-dimensional space it is simplest to represent a vector v by the n-tuple if its components $\{v_n\}$ with respect to a specified orthonormal basis

$$v \;\rightarrow\; \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}. \qquad (2.14)$$

But the "vectors" we encounter in quantum mechanics are *functions*, and they live in *infinite*-dimensional spaces As we mentioned in Sect 2.1.1, to represent a possible physical state, the wavefunction must be normalised.

<u>Definition:</u> The set of all square-integrable functions in $\mathbb{R}^3$ is called $L^2(\mathbb{R}^3)$ by mathematicians, *Hilbert space* by physicists. In quantum mechanics, wavefunctions live in Hilbert space $\mathcal{H}$.

We can naturally define an inner product on this vector space, in analogy with the finite-dimensional case. In linear algebra, you learn that the inner product (v, w) of two vectors (generalising the dot product in three dimensions) is the complex number

$$(v, w) = v_1^* w_1 + v_2^* w_2 + \cdots + v_n^* w_n. \tag{2.15}$$

<u>Definition:</u> In quantum mechanics, the inner product of two wavefunctions $\psi(\mathbf{x}, t)$ and $\phi(\mathbf{x}, t)$ at a time t is given by

$$(\psi, \phi) \equiv \int_{\mathbb{R}^3} \psi^*(\mathbf{x}, t)\, \phi(\mathbf{x}, t)\, dV. \tag{2.16}$$

<u>Theorem 2.2:</u> If ψ and ϕ are both square-integrable (that is they are both in Hilbert space), their inner product is guaranteed to exist, i.e. the integral above converges to a finite number.

  Proof: Let us take $\psi(\mathbf{x}, t)$ and $\phi(\mathbf{x}, t)$. From the integral Schwartz inequality, we have

$$\left| \int_{\mathbb{R}^3} \psi^*(\mathbf{x}, t)\phi(\mathbf{x}, t)\, dV \right| \leq \sqrt{ \int_{\mathbb{R}^3} |\psi(\mathbf{x}, t)|^2\, dV \int_{\mathbb{R}^3} |\phi(\mathbf{x}, t)|^2\, dV }. \tag{2.17}$$

Therefore, if both terms on the right-hand side of the above equation are finite, then also the inner product is finite. $\square$

Without giving any proof for this, we can extend the definition of the inner product (,) to any ψ and ϕ for which the integral in Eq. (2.16) is well-defined, whether or not ψ or ϕ are normalisable.

You can now check that Eq. (2.16) satisfies all conditions for an inner product. In particular

1. $(\psi, \phi) = (\phi, \psi)^*$.

2. The inner product is anti-linear in the first entry and linear in the second:

$$(a_1\psi_1 + a_2\psi_2, \phi) = a_1^*(\psi_1, \phi) + a_2^*(\psi_2, \phi)$$
$$(\psi, a_1\phi_1 + a_2\phi_2) = a_1(\psi, \phi_1) + a_2(\psi, \phi_2)$$

3. The inner product of ψ with itself

$$(\psi, \psi) = \int_{\mathbb{R}^3} |\psi(\mathbf{x}, t)|^2\, dV,$$

is real and non-negative and, for continuous functions, it is zero only when $\psi = 0$, which is a not an admissible physical state.

Now, using the definition of *inner product*, we can define (or redefine) a series of important properties of wavefunctions:

Definition: The *norm* of a wavefunction is the squared root of the inner product of the wavefunction with itself $||\psi|| = \sqrt{(\psi, \psi)}$.

Definition: A wavefunction ψ is called *normalised* if $||\psi|| = 1$.

Definition: Two wavefunctions ψ, ϕ are *orthogonal* is their inner product (ψ, ϕ) is zero.

Definition: A set of wavefunctions $\{\psi_n\}$ is *orthonormal* is they are normalised and mutually orthogonal:

$$(\psi_m, \psi_n) = \delta_{mn}.$$

Definition A set of wavefunctions $\{\psi_n\}$ is *complete* if any other wavefunction in Hilbert space can be expressed as a linear combination of them:

$$\phi = \sum_{n=0}^{\infty} c_n \psi_n. \tag{2.18}$$

Lemma 2.3: If the wavefunctions $\{\psi_n\}$ that form a complete set are also an orthonormal set, the coefficients are given by the Fourier's trick:

$$c_n = (\psi_n, \phi). \tag{2.19}$$

  Proof:

$$
\begin{aligned}
(\psi_n, \phi) &= (\psi_n, \sum_{m=0}^{\infty} c_m \psi_m) \\
&= \sum_{m=0}^{\infty} c_m (\psi_n, \psi_m) = \sum_{m=0}^{\infty} c_m \delta_{mn} = c_n,
\end{aligned}
\tag{2.20}
$$

where in the second line we used the linearity of the inner product in the second entry. $\square$

As far as the physics interpretation is concerned, if $\phi(\mathbf{x}, t)$ is the desired outcome, then the probability of measuring such an outcome given the wavefunction $\psi(\mathbf{x}, t)$ at a time t is given by Born's Rule

$$|(\psi, \phi)|^2 = \left| \int_{\mathbb{R}^3} \psi^*(\mathbf{x}, t)\, \phi(\mathbf{x}, t)\, dV \right|^2. \tag{2.21}$$

We say that (2.21) is the probability amplitude of ϕ to be found in ψ at time t, so in a sense it measures the overlap of the two wavefunctions.

2.3   Time-dependent Schrödinger equation (TDSE)

The wavefunction gives us a description of the state of the system. The next step is to understand how these states evolve in time. The time evolution of wavefunction $\psi(\mathbf{x}, t)$ is governed by the time-dependent Schrödinger equation. For a non-relativistic particle of mass m moving in a potential $U(\mathbf{x})$, this reads

$$i\hbar \frac{\partial \psi}{\partial t}(\mathbf{x}, t) = -\frac{\hbar^2}{2m}\nabla^2 \psi(\mathbf{x}, t) + U(\mathbf{x})\psi(\mathbf{x}, t) \tag{2.22}$$

Note that, being a first order equation in t, $\psi(\mathbf{x}, t)$ is uniquely determined by Eq. (2.22) and the initial value $\psi(\mathbf{x}, 0)$. On the other hand Eq. (2.22) is a second order differential equation in $\mathbf{x}$. The asymmetry between $\mathbf{x}$ and t is intrinsically related to the fact that the TDSE is a non-relativistic equation. Eq. (2.22) is a postulate of quantum mechanics. If you want a heuristic justification in 1D, recall that de Broglie postulated that particles are described by waves, and that the energy and momentum are related to the angular frequency and wave vector by $E = \hbar\omega$ and $p = \hbar k$. We can express this by associating to a particle of energy E and momentum p a wave of the form (called the de Broglie wave)

$$\psi(x, t) = \exp\left[i(kx - \omega t)\right] = \exp\left[\frac{i}{\hbar}\left(px - \frac{p^2}{2m}t\right)\right] \tag{2.23}$$

The simplest wave equation to which the above wave is the general solution is the time-dependent 1D Schrödinger equation (2.22) for a free particle, *i.e.* for $U(x) = 0$. We will see in Sect. 3.2 that the De Broglie wave is non-normalisable, and thus cannot be an acceptable solution of the Schrödinger equation. We will thus introduce the concept of wave package to achieve a correct solution of the free particle problem.

Let's return to the statistical interpretation of the wavefunction, of Eq. (2.5). This might seem in apparent contradiction with the Schrödinger equation. After all, the wavefunction is supposed to be determined by the Schrödinger equation, so we cannot impose an external condition on $\psi(\mathbf{x}, t)$ without checking that the two are consistent. Suppose that we have normalised the wavefunction at time $t = 0$. How do we know that the wavefunction will stay normalisable as time goes? Fortunately, the Schrödinger equation has the remarkable property that it automatically preserves the normalisation of the wavefunction. Without this crucial feature, Eq. (2.5) and Eq. (2.22) would be mutually incompatible, as probability would not be conserved in time. This point is important and it deserves a careful proof. To begin with

$$\frac{d}{dt}\int_{\mathbb{R}^3} |\psi(\mathbf{x}, t)|^2 \, dV \;=\; \int_{\mathbb{R}^3} \frac{\partial}{\partial t}|\psi(\mathbf{x}, t)|^2 \, dV. \tag{2.24}$$

By the product rule,

$$\frac{\partial}{\partial t}|\psi|^2 = \frac{\partial}{\partial t}(\psi^*\psi) \;=\; \psi^*\frac{\partial \psi}{\partial t} + \psi\frac{\partial \psi^*}{\partial t}. \tag{2.25}$$

27

Because of the TDSE, Eq. (2.22), we have

$$\frac{\partial}{\partial t}\psi = \frac{i\hbar}{2m}\nabla^2\psi - \frac{i}{\hbar}U\psi \tag{2.26}$$

and

$$\frac{\partial}{\partial t}\psi^* = -\frac{i\hbar}{2m}\nabla^2\psi^* + \frac{i}{\hbar}U\psi^*, \tag{2.27}$$

so

$$\frac{\partial}{\partial t}|\psi|^2 = \frac{i\hbar}{2m}\left(\psi^*\nabla^2\psi - \psi\nabla^2\psi^*\right) = \nabla\cdot\left[\frac{i\hbar}{2m}\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right)\right]. \tag{2.28}$$

The integral in Eq. (2.24) can now be evaluated explicitly and gives

$$\left.\frac{i\hbar}{2m}\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right)\right|_{\delta V} = 0, \tag{2.29}$$

given that $\psi(\mathbf{x}, t)$ must be zero at the boundaries δV, *i.e.* at $|\mathbf{x}| \to \infty$.
Hence a wavefunction's normalisation is constant and the *probability conservation* holds

$$\frac{\partial}{\partial t}\rho(\mathbf{x}, t) + \nabla\cdot\mathbf{j}(\mathbf{x}, t) = 0, \tag{2.30}$$

where

$$\mathbf{j}(\mathbf{x}, t) = -\frac{i\hbar}{2m}\left(\psi^*(\mathbf{x}, t)\nabla\psi(\mathbf{x}, t) - \psi(\mathbf{x}, t)\nabla\psi^*(\mathbf{x}, t)\right) \tag{2.31}$$

is the conserved probability current of the physics states that obey the Schrödingerequation.

2.4   Measurements in Quantum Mechanics

In Sect. 2.1 we have seen that the modulus square of the wavefunction gives the probability distribution of the position of the particle. How about other information such as a particle's momentum or energy? It turns out that all the information about the particle is contained in the wavefunction (which is why we call it the "state" of the particle). We call *observable* each property of the particle which we can measure. Each observable is represented by an *operator* acting on the state ψ. Each measurement is represented by an *expectation value*.

2.4.1   An heuristic explanation

In this section we start by giving an heuristic explanation that introduces the concept of measurements continuing along the lines of the probabilistic interpretation of the wavefunction. Afterwards we provide a more mathematical definition by introducing the concepts of operators and expectation values.

Given the probabilistic interpretation of the wavefunction provided by the Born's rule, for a particle in state ψ, the expectation value of x (for simplicity let's keep working in one dimension) is given by

$$\langle x \rangle = \int_{-\infty}^{+\infty} x |\psi(x,t)|^2 \, dx, \tag{2.32}$$

where $\psi(x,t)$ is the normalised wavefunction. The above equation is nothing but the value of the position x weighted over the probability density in the $x-$space $|\psi|^2$ of the particle associate to the state ψ. What exactly does this mean? It does not mean that if you measure the position of one particle over and over again, $\langle x \rangle$ is the average result you will get. On the contrary: the first measurement (whose outcome is not determinate) will collapse the wavefunction to a spike at the value actually obtained and the subsequent measurements will simply give the same result. Rather, $\langle x \rangle$ is the average of measurements performed on particles all in the same state ψ. Basically you prepare a whole ensemble of particle, each in the same state ψ, and measure the position of all of them: $\langle x \rangle$ is the average of these results. To summarise, the expectation value is the average of repeated measurements on an ensemble of identically prepared systems, not the average of repeated measurements on the same system.

Now, as time goes on, $\langle x \rangle$ will change (because of the time-dependence of $\psi(x,t)$) and we might be interested in knowing what is the momentum of the particle. In our heuristic explanation we start from what we are familiar with and define the momentum as the mass times velocity, i.e. as the mass times the derivative of the position. This way we get

$$
\begin{aligned}
\langle p \rangle &= m \frac{d\langle x \rangle}{dt} = m \frac{d}{dt} \int_{-\infty}^{+\infty} x |\psi(x,t)|^2 \, dx = m \int_{-\infty}^{+\infty} x \frac{\partial}{\partial t} (\psi^*(x,t)\psi(x,t)) \, dx \\
&= \frac{i\hbar}{2} \int_{-\infty}^{+\infty} x \frac{\partial}{\partial x} \left(\psi^*(x,t) \frac{\partial \psi(x,t)}{\partial x} - \psi(x,t) \frac{\partial \psi^*(x,t)}{\partial x} \right) \, dx,
\end{aligned}
\tag{2.33}
$$

where in the last passage we have used the T.D.S.E. for a free particle and its complex conjugate version. The above expression can be further simplified using integration-by-parts and throwing away the boundary terms (on the grounds that $\bar{\psi}$ must go to zero at $\pm\infty$ faster than $1/x$) and we obtain the expectation value for the momentum p

$$
\begin{aligned}
\langle p \rangle &= \frac{-i\hbar}{2} \int_{-\infty}^{+\infty} \left(\psi^*(x,t) \frac{\partial \psi(x,t)}{\partial x} - \psi(x,t) \frac{\partial \psi^*(x,t)}{\partial x} \right) \, dx \\
&= -i\hbar \int_{-\infty}^{+\infty} \psi^*(x,t) \frac{\partial \psi(x,t)}{\partial x} \, dx,
\end{aligned}
\tag{2.34}
$$

where in the last passage we performed a further integration by part. The expressions for

the expectation values $\langle x \rangle$ and $\langle p \rangle$ can be written as

$$\langle x \rangle \;\; = \;\; \int_{-\infty}^{+\infty} \psi^*(x,t) x \psi(x,t)\, dx, \tag{2.35}$$

$$\langle p \rangle \;\; = \;\; \int_{-\infty}^{+\infty} \psi^*(x,t) \left(-i\hbar \frac{\partial}{\partial x} \right) \psi(x,t)\, dx. \tag{2.36}$$

We now say that the operator $\hat{x} = x$ represents the position of a particle and that the operator $\hat{p} = -i\hbar \partial/\partial x$ represents the momentum of a particle (both in the $x-$space representation).

In general, to calculate the expectation value of any quantity $\mathcal{Q}(x,p)$, we simply replace p by $(-i\hbar \partial/\partial x)$ and insert the resulting operator between ψ^* and ψ and integrate over the whole space

$$\langle \mathcal{Q}(\mathbf{x}, \mathbf{p}) \rangle = \int_{\mathbb{R}^3} \psi^* \mathcal{Q}\left(\mathbf{x}, -i\hbar \nabla\right) \psi\, dV \tag{2.37}$$

The above equation is a recipe for computing the expectation values of any dynamical quantities, for a particle in state ψ. It might be useful that the quantity $Q(x,p)$ should have an implicit ordering associated with it, otherwise you can get different operators for the same classical quantities. For example $Q(x,p) = xp$ and $Q(x,p) = px$ give rise to different operators, even though they are the same as classical quantities. Fixing an order (e.g. saying all x's occur before all p's) sorts out the problem.

In the next two sections we will give a more mathematical definition of operators and expectation values.

2.4.2 Hermitian operators

After having given a more heuristic explanation (that physicists usually like) we now give a more rigorous definition of operators and in the next section of measurements. To do that, we continue the analogy with linear algebra.

It is well-known that in a n-dimensional complex space linear transformations T are represented by matrices (with respect to the specific basis), which act on vectors (to produce new vectors) by the ordinary rules of matrix multiplication.

$$w = Tv \;\rightarrow\; \mathbf{w} = \mathbf{T} \mathbf{v} = \begin{pmatrix} t_{11} & t_{12} & \ldots & t_{1n} \\ t_{21} & t_{22} & \ldots & t_{2n} \\ \vdots & \vdots & & \vdots \\ t_{n1} & t_{n2} & \ldots & t_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \tag{2.38}$$

In quantum mechanics, linear transformations are represented by linear operators $\hat{O}$, which act on wavefunctions to produce new wavefunctions of the Hilbert space $\mathcal{H}$.

Definition: We define an operator $\hat{O}$ to be any linear map from the Hilbert space to itself *i.e.* any map such that

$$\hat{O}(a_1\psi_1 + a_2\psi_2) = a_1\hat{O}\psi_1 + a_2\hat{O}\psi_2, \tag{2.39}$$

for all complex numbers a_1, a_2 and all ψ_1, $\psi_2 \in \mathcal{H}$[3].
To simplify the notation, without any loss of generality, we give examples of linear operators in 1D, such as

- Finite differential operators

$$\sum_{n=0}^{N} p_n(x)\frac{\partial^n}{\partial x^n}$$

  where the $p_n(x)$ are polynomials. in x. Note that this class includes the $\hat{p}$ and $\hat{x}$.

- the translation operators
$$S_a: \quad \psi(x) \to \psi(x-a)$$

- the parity operator
$$P: \quad \psi(x) \to \psi(-x)$$

There is a special class of operators in quantum mechanics , the Hermitian operators. They are special because physical observables are associated to Hermitian operators.

Definition: We define the Hermitian conjugate $\hat{A}^\dagger$ of an operator $\hat{A}$ to be the operator such that

$$(\hat{A}^\dagger\psi_1, \psi_2) = (\psi_1, \hat{A}\psi_2) \tag{2.40}$$

for all normalisable wavefunctions ψ_1, $\psi_2 \in \mathcal{H}$.

Exercise: Verify the following properties of Hermitian conjugation

1. $(a_1\hat{A}_1 + a_2\hat{A}_2)^\dagger = a_1^*\hat{A}_1^\dagger + a_2^*\hat{A}_2^\dagger$
2. $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$

Definition: An operator is Hermitian if

$$\hat{A} = \hat{A}^\dagger. \tag{2.41}$$

All quantum dynamical variables or observables (i.e. all quantities that we can measure) are represented by Hermitian operators, and viceversa. Examples of operators defining

---

[3]Sometimes it is also useful to expand the definition to include anti-linear operators (specifically in the case of time-reversal symmetry).

quantum observables such as the position, momentum and energy operators, are defined (in 1D) as:

- position

$$\hat{x}: \quad \psi(x,t) \to x\psi(x,t), \tag{2.42}$$

- momentum

$$\hat{p}: \quad \psi(x,t) \to -i\hbar\frac{\partial\psi(x,t)}{\partial x}. \tag{2.43}$$

- kinetic energy

$$\hat{T}: \quad \psi(x,t) \to \frac{\hat{p}^2}{2m}\,\psi(x,t) = -\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2}, \tag{2.44}$$

- potential energy

$$\hat{U}: \quad \psi(x,t) \to U(\hat{x})\,\psi(x,t) = U(x)\,\psi(x,t) \tag{2.45}$$

- total energy, sum of the kinetic and the potential energies

$$\hat{H}: \quad \psi \to -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x,t) + U(x)\psi(x,t). \tag{2.46}$$

Exercise: Verify explicitly that all the above operators are Hermitian, *i. e.* that $(\hat{x}\psi, \phi) = (\psi, \hat{x}\phi)$, $(\hat{p}\psi, \phi) = (\psi, \hat{p}\phi)$ and $(\hat{H}\psi, \phi) = (\psi, \hat{H}\phi)$ for all $\psi, \phi \in \mathcal{H}$.

We will now describe several crucial properties of Hermitian operators.

<u>Theorem 2.4:</u> The eigenvalues of a Hermitian operator are real.

    Proof: Let $\hat{A}$ be Hermitian and ψ be a normalised eigenfunction with eigenvalue a: $\hat{A}\psi = a\psi$. We have

$$\begin{aligned}
(\psi, \hat{A}\psi) &= (\psi, a\psi) = a(\psi, \psi) = a, \\
(\psi, \hat{A}\psi) &= (\hat{A}^\dagger\psi, \psi) = (\hat{A}\psi, \psi) \\
&= (a\psi, \psi) = a^*(\psi, \psi) = a^*.
\end{aligned} \tag{2.47}$$

Hence $a^* = a$, so the eigenvalues of a Hermitian operator are real.   □

<u>Theorem 2.5:</u> Let $\hat{A}$ be a Hermitian operator and ψ_1, ψ_2 be normalised eigenfunctions with different eigenvalues a_1, a_2. Then ψ_1 and ψ_2 are orthogonal.

    Proof: We have $\hat{A}\psi_1 = a_1\psi_1$ and $\hat{A}\psi_2 = a_2\psi_2$, and by the theorem above we have that a_1, a_2 are real.

$$\begin{aligned}
a_1(\psi_1, \psi_2) &= a_1^*(\psi_1, \psi_2) = (a_1\psi_1, \psi_2) \\
&= (\hat{A}\psi_1, \psi_2) = (\hat{A}^\dagger\psi_1, \psi_2) \\
&= (\psi_1, \hat{A}\psi_2) = (\psi_1, a_2\psi_2) = a_2(\psi_1, \psi_2)
\end{aligned} \tag{2.48}$$

Hence $(a_1 - a_2)(\psi_1, \psi_2)$. But since $a_1 \neq a_2$, then $(\psi_1, \psi_2) = 0$, *i.e.* ψ_1 and ψ_2 are orthogonal.
$\square$

Our discussion is complicated by the fact that the eigenfunctions of Hermitian operators (*i.e.* of quantum observables) are not necessarily all normalisable. For example the position operator $\hat{x}$ has normalisable eigenfunctions $\delta(x - q)$, with eigenvalue q

$$\hat{x}\delta(x - q) = x\delta(x - q) = q\delta(x - q). \tag{2.49}$$

However they do not exist in the space of wavefunctions (since they are generalised functions, i.e. distributions, rather than true functions). On the other hand the momentum operator $\hat{p} = -i\hbar\partial/\partial x$ has eigenfunctions $\exp(ipx/\hbar)$ with eigenvalue p

$$\hat{p}\, e^{\frac{ipx}{\hbar}} = p\, e^{\frac{ipx}{\hbar}}, \tag{2.50}$$

which however are not normalisable.

In Chapt. 3 we will see that the Hamiltonian, i.e. the energy operator, has normalisable eigenfunctions for all systems corresponding to bound states and not normalisable eigenfunctions for all systems corresponding to unbound states or scattering solutions. In general, a Hermitian operator may have both a set of normalisable eigenfunctions with discrete eigenvalues and a set of non-normalisable eigenfunctions with continuous eigenvalues, and the theorems stated in this section apply to both sets[4].

Theorem 2.6: The discrete and continuous sets of eigenfunctions of any Hermitian operator together form a complete orthogonal basis of the physical wavefunctions, i.e. of the normalisable complex-valued wavefunctions $\psi(\mathbf{x})$[5].

Note that we say the Hermitian operator $\hat{A}$ has a degenerate eigenvalue λ if it has more than one linearly independent eigenfunction with eigenvalue λ. If $\hat{A}$ has degenerate eigenvalues, we define an orthogonal basis of its eigenfunctions by choosing orthogonal bases for the eigenfunctions associated with each degenerate eigenvalue λ.

In the next Chapter, we will see that, if we apply Theorem 2.6 to a very special Hermitian operator, the Hamiltonian $\hat{H}$, we obtain the important corollary, that links the solutions of the TDSE to the eigenfunctions of $\hat{H}$.

---

[4]*Non-examinable technical note: it is well beyond our scope here but, in fact, it turns out one can find a more general notion of normalisability which covers both sets of eigenfunctions, and more general versions of the theorems can be precisely framed in terms of this condition. This definition includes bound states and scattering solutions to the time-independent SE, but not solutions which blow up exponentially. A discussion can be found in, for example, Messiah, "Quantum Mechanics", vol. 1, chap V.9.*

[5]This theorem is quite hard to prove in complete generality. We will assume it without proof in this course.

2.4.3   Expectation values

In this section we list a series of important postulates that have to do with quantum measurements.

i) Every quantum observable O is represented by a Hermitian operator $\hat{O}$.

ii) The possible outcomes of a measurement of O are the eigenvalues of $\hat{O}$.

iii) If $\hat{O}$ has a discrete (either finite or infinite) set of normalised eigenfunctions $\{\psi_i\}$ with corresponding distinct eigenvalues $\{\lambda_i\}$, and a measurement of O is carried out on a particle with a normalised wavefunction

$$\psi = \sum_i a_i\,\psi_i, \tag{2.51}$$

then the probability of the outcome λ_i is equal to $|a_i|^2$.
Of course, if the wavefunction ψ is an eigenfunction ψ_i of $\hat{O}$, then the measurement outcome will be λ_i with probability one. For example, a stationary state obeying $\hat{H}\psi = E\psi$ will always give outcome E if one measures the energy of the state. But unless the wavefunction ψ is an eigenfunction of the measured observable, the measurement outcome is not definitely predictable. In contrast to classical mechanics, a quantum observable does not generally have a definite value on a quantum state.

iv) We can extend iii) to the case when $\hat{O}$ has degenerate eigenvalues. If O is measured on a state $\psi = \sum a_i\psi_i$, where $\{\psi_i\}$ are orthonormal eigenfunctions of the corresponding operator $\hat{O}$ and $\{\psi_i\}_{i\in I}$ are a complete set of orthonormal eigenfunctions with the same eigenvalue λ, then the probability of outcome λ is

$$P(O = \lambda) = \sum_{i\in I} |a_i|^2. \tag{2.52}$$

v) The total probability of all possible outcomes is one

$$\sum_i |a_i|^2 \;=\; \sum_i (a_i\psi_i, a_i\psi_i) \;=\; \sum_{i,j}(a_i\psi_i, a_j\psi_j) \;=\; (\psi, \psi) = 1, \tag{2.53}$$

as it should be. As a result, the postulates are consistent: the sum of the probabilities of all possible outcomes is 1, and so if you carry out a measurement you will certainly get some outcome and you will only get one outcome. (We already verified this in the case of the Born rule for the measurement of the position.)

vi) The projection postulate: If a measurement of the observable O is carried out on a particle with normalised wavefunction $\psi(\mathbf{x}, t)$ at time t and the outcome λ_i is obtained, the wavefunction instantaneously after the measurement becomes $\psi_i(\mathbf{x})$.

This is sometimes referred to as the "collapse of the wavefunction". The wavefunction then evolves according to the Schrödinger, with initial state $\psi_i(\mathbf{x})$ at time t, until the next measurement.

vii) We can extend this statement to the case when $\hat{O}$ has degenerate eigenvalues. If $\psi = \sum_i a_i \psi_i$ is measured where $\{\psi_i\}$ are orthonormal eigenfunctions of $\hat{O}$ and $\{\psi_i\}_{i \in I}$ are a complete set of orthonormal eigenfunctions with the same eigenvalue λ, then the state resulting after a measurement with outcome λ is given by $\sum_{i \in I} a_i \psi_i$ (up to normalisation).

The projection postulate is so called because it implies that the post-measurement wavefunction $\psi_i(\mathbf{x}, t)$ is obtained from the pre-measurement wavefunction $\psi(\mathbf{x}, t)$ by the action of the projection operator $\hat{P}_i$ defined by

$$\hat{P}_i : \psi \quad \rightarrow \quad (\psi_i, \psi)\psi_i, \tag{2.54}$$

up to normalisation. We call $\hat{P}_i$ a projection since it maps any state onto its component in a particular linear subspace, namely the subspace spanned by ψ_i – an action analogous to, for instance, the projection of a 3D vector (x, y, z) onto its x-component $(x, 0, 0)$.

A physics implication of the projection postulate is that, if an observable O is measured twice, with infinitesimal time separating the two measurements, then if the first outcome is λ_i the second will, with probability one, also be λ_i. This is the reason why, in Chapt. 2, we stressed that the expectation value of the position operator, $\langle x \rangle$, is to be interpreted as the average of measurements performed on an ensemble of particles all in the same state ψ. Basically the expectation value is the average of repeated measurements on an ensemble of identically prepared systems, not the average of repeated measurements on the same system, otherwise the first measurement (whose outcome is not determinate) will collapse the wavefunction to a spike at the value actually obtained and the subsequent measurements will simply give the same result. This has the important consequence that quantum measurements resemble classical measurements in at least one sense: they establish a property of the system that can be repeatedly verified. If we measure something, and then quickly measure it again, we get the same answer. If this was not true, it would be hard to find any good reason for the use of the term "measurement" in quantum mechanics.

We can now formulate more in depth the concept of expectation value. Consider a measurement of the observable O on the state ψ. If the corresponding Hermitian operator $\hat{O}$ has only a discrete set of normalisable eigenfunctions $\{\psi_i\}$, we have just learnt that the possible outcomes are the corresponding eigenvalues λ_i, with outcome probabilities are $P_i = |(\psi, \psi_i)|^2$. The *expectation value* of the measured outcome, in the standard statistical

sense of the term, is thus

$$
\begin{aligned}
\sum_i P_i \lambda_i &= \sum_i |(\psi, \psi_i)|^2 \, \lambda_i \\
&= \left(\sum_i (\psi, \psi_i)\psi_i, \sum_j \lambda_j (\psi, \psi_j)\psi_j \right) = (\psi, \hat{O}\psi)
\end{aligned}
\tag{2.55}
$$

Thus we have

$$
\langle \hat{O} \rangle_\psi = (\psi, \hat{O}\psi),
\tag{2.56}
$$

where by $\langle \hat{O} \rangle_\psi$ we indicate the expectation value of a measurement of the observable O on the state ψ. By the above definition, we can justify this definition of expectation value for the position operator $\hat{x}$, that we introduced in Sect. 2.3 from the probabilistic interpretation. Recall that the probability of obtaining a position measurement outcome in the infinitesimal volume dV is given by $|\psi(\mathbf{x}, t)|^2 dV$ (Born's rule). The expectation value of a position measurement is thus

$$
\int_{\mathbb{R}^3} \mathbf{x}|\psi(\mathbf{x}, t)|^2 \, dV = \int_{\mathbb{R}^3} \psi^*(\mathbf{x}, t)\, \hat{\mathbf{x}}\, \psi(\mathbf{x}, t) \, dV = (\psi, \hat{\mathbf{x}}\psi).
\tag{2.57}
$$

This encourages us to take Eq. (2.56) as a general definition of expectation value for any observable, whether its eigenvalues are discrete, continuous or a combination of those.

Note that the expectation value is linear with respect to real scalars *i.e.*

$$
\langle a\hat{A} + b\hat{B} \rangle_\psi = a\langle \hat{A} \rangle_\psi + b\langle \hat{B} \rangle_\psi,
\tag{2.58}
$$

for any Hermitian operators $\hat{A}$, $\hat{B}$ and any real numbers a, b. We restrict to a, b real here because the interpretation of $\langle \hat{A} \rangle_\psi$ as an expectation value of an observable requires that $\hat{A}$ is Hermitian, since observables are always represented by Hermitian operators. A complex multiple of a Hermitian operator is not generally Hermitian: if $\hat{A}$ is Hermitian then $(c\hat{A})^\dagger = c^*\hat{A}^\dagger = c^*\hat{A} \neq c\hat{A}$ if $c^* \neq c$.

Chapter 3

Solutions of the Schrödinger equation

In this chapter we introduce the time-independent Schrödinger equation and stationary stated. We then familiarise ourselves with it, by looking at various examples of solutions of the one-dimensional (1D) Schrödinger equation for one particle. We start with a particle confined in a region in x and only then move to solving the Schrödinger equation for a free particle, which is much trickier than for a confined particle. We will then look at the solutions we get when a particle is in a finite potential well, and at several cases of scattering problems, such as a potential step and a potential barrier. We will conclude by looking into the solution of the Schrödinger equation for a harmonic oscillator.

3.1   Time-independent Schrödinger equation (TISE)

Les us now go back to the time-dependent Schrödinger equation (2.22). We rewrite it explicitly in one dimension as

$$\hat{H}\chi(x) = -\frac{\hbar^2}{2m}\chi''(x) + U(x)\chi(x) = E\chi(x), \tag{3.1}$$

with $\chi''(x) \equiv d^2\chi/dx^2$. We want to find the allowed energy eigenvalues for a particle in 1D given a potential U. If $U(x)$ is time independent, we can then use the method of separation of variables to find solutions, by trying the ansatz (try solution)

$$\psi(x,t) = \chi(x)T(t). \tag{3.2}$$

Plugging it into (3.1) gives

$$(\hat{H}\chi(x))T(t) = \chi(x)i\hbar\frac{\partial T(t)}{\partial t} \quad \Rightarrow \frac{\hat{H}\chi(x)}{\chi(x)} = \frac{1}{T(t)}\left(i\hbar\frac{\partial T(t)}{\partial t}\right). \tag{3.3}$$

Since the left hand side depends only on x and the right hand side only on t, both must equal a constant, which we call E. Thus we have

$$T(t) \; = \; \exp\left(-\frac{i}{\hbar}Et\right), \tag{3.4}$$

and the time-independent Schrödinger Equation (TISE)

$$\hat{H}\chi = E\chi \Leftrightarrow -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\chi(x) + U(x)\chi(x) \; = \; E\chi(x). \tag{3.5}$$

Note that the constant E has to be real, otherwise $T(t)$ would not be defined at all times t. Indeed, if $E = E_0 + i\Gamma$, with $E_0, \Gamma \in \mathbb{R}$, then

$$T(t) \; = \; \exp\left(-\frac{i}{\hbar}E_0 t\right)\exp\left(\frac{1}{\hbar}t\Gamma\right), \tag{3.6}$$

which would go to ∞ for $t \to \infty$. Hence the energy E is a real number. From Eq. (3.5), we argue that $\chi(x)$ is an eigenfunction of the Hamiltonian operator $\hat{H}$ which corresponds to a physical state whose energy is given by the eigenvalue E. The terminology is a natural extension to infinite-dimensional spaces (of functions) of the definitions of eigenvector and eigenvalue for finite-dimensional matrices.

3.2   Stationary States

Definition: The particular class of solutions that have the form

$$\Psi(x,t) = \chi(x)\exp\left(-\frac{i}{\hbar}Et\right), \tag{3.7}$$

where $\chi(x)$ is an eigenfunction of the Hamiltonian with eigenvalue E, are called *stationary* states. We will refer to $\chi(x)$ as the stationary state wavefunction. Stationary states are states of definite energy E. For stationary states the probability density

$$\rho(x,t) = |\Psi(x,t)|^2 \; = \; |\chi(x)|^2, \tag{3.8}$$

is time independent, hence the name stationary state.

By applying theorem 2.6 to the Hamiltonian operatior we get an important corollary.
Theorem 3.1: Every solution of the TDSE can be written as a superposition of stationary states $\chi(x)T(t)$.
For systems that have a discrete set of energy eigenfunctions

$$\psi(x,t) = \sum_{n=1}^{\infty} a_n \chi_n(x) e^{-iE_n t/\hbar}, \tag{3.9}$$

while for systems that have a continuous set of energy eigenfunctions labelled by a continuous index α

$$\psi(x,t) \;=\; \int_\Delta A(\alpha)\,\chi_\alpha(x)e^{-iE_\alpha t/\hbar}\,d\alpha. \tag{3.10}$$

Stationary states are very important in quantum mechanics. Indeed the TISE yields a collection of solutions $\chi_1(x), \chi_2(x), \chi_3(x)...$, each with its associated value of the separation constant $E_1, E_2, E_3...$ Thus there is a different wavefunction for each allowed energy:

$$\Psi_1(x,t) = \chi_1(x)e^{-iE_1 t/\hbar}, \qquad \Psi_2(x,t) = \chi_2(x)e^{-iE_2 t/\hbar}, ... \tag{3.11}$$

Now, given that the TDSE has the property that any linear combination of solutions is itself a solution, once we have found the separable solutions, then we can immediately construct a much more general solution of the form

$$\psi(x,t) = \sum_{n=1}^{\infty} a_n \chi_n(x)e^{-iE_n t/\hbar}. \tag{3.12}$$

Note that the state defined in (3.12) is *not* a stationary state. We can see this explicitly. Take a state ψ, which is the superposition of two stationary states ψ_1 and ψ_2 with two real coefficient of proportionality a_1 and a_2 and real stationary state wavefunctions $\chi_1(x)$ and $\chi_2(x)$ (for simplicity). If we compute the probability density associated to this state, we see that the it is not a constant, rather it oscillates with time. Indeed

$$
\begin{aligned}
|\psi(x,t)|^2 &= |a_1\Psi_1(x,t) + a_2\Psi_2(x,t)|^2 \\
&= a_1^2|\chi_1(x)|^2 + a_2^2|\chi_2(x)|^2 + a_1 a_2 \chi_1(x)\chi_2(x)[\exp(i\Delta E t/\hbar) + \exp(-i\Delta E t/\hbar)] \\
&= a_1^2|\chi_1(x)|^2 + a_2^2|\chi_2(x)|^2 + 2a_1 a_2 \chi_1(x)\chi_2(x)\,\cos(\Delta E t/\hbar),
\end{aligned}
\tag{3.13}
$$

with $\Delta E = E_1 - E_2$.

As for Theoremo 3.1, *every* solution of the TDSE can be written like Eq. (3.12). It is only a problem of finding the right complex constants a_n so as to fit the initial condition of the problem at hand. So, the stationary states form a basis of the state space. Hence, by understanding the stationary states, we can understand a lot about a quantum system. In general $\psi(x,t)$ is not a stationary state and thus does not have a definite energy. Instead, the probability of measuring the particle energy as E_n is equal to $|a_n|^2$. This is intrinsically related to the fact that the stationary states are an orthonormal basis of the Hilbert space.

* In general, a discrete set of eigenvalues is associated with normalisable wavefunctions, while a continuous spectrum is associated with non-normalisable wavefunctions. The former give rise to *bound states*, *i.e.* of states in which the particle cannot be too far from
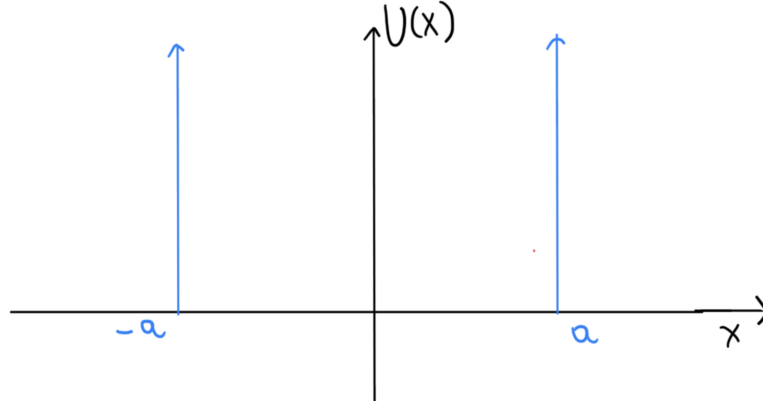
Figure 3.1: The infinite square well potential.

the origin. How do we see that? It is quite simple, as (take a one-dimensional state for simplicity):

$$\int_{-\infty}^{+\infty} dx\,|\chi|^2 = \mathcal{N} < \infty \;\Rightarrow\; \lim_{R\to\infty}\int_{|x|>R} dx|\chi(x)|^2 = 0. \tag{3.14}$$

Using the same kind of reasoning, we can immediately see that the above equation does not hold for non normalisable functions, which instead are associated with *scattering states*, and the particle can be at $\pm\infty$ *.

3.3 Bound states

3.3.1 Infinite potential well

The simplest case to consider is the infinite potential square well. Here the potential is infinite outside the region $[-a, a]$ and the particle is free inside this region.

$$U(x) = \begin{cases} 0 & |x| \le a \\ \infty & |x| > a, \end{cases} \tag{3.15}$$

see Fig. 3.1 for a sketch. For $|x| > a$, we must have $\chi(x) = 0$, or else $U(x)\chi(x)$ would be infinite. We require then $\chi(x) = 0$ for $|x| \ge a$ and $\chi(x)$ to be continuous at $x = \pm a$. Within $x \le a$ Eq. (3.1) becomes

$$-\frac{\hbar^2}{2m}\chi''(x) = E\chi(x). \tag{3.16}$$

We simplify this to become

$$\chi''(x) + k^2\chi(x) = 0, \qquad \text{with } k^2 = \frac{2mE}{\hbar^2}. \tag{3.17}$$

Given that $E > 0$, the general solution is the general oscillatory solution

$$\chi(x) = A \sin(kx) + B \cos(kx), \tag{3.18}$$

with two arbitrary complex constants A and B, which depends on the boundary conditions. To find A and B we match solutions at the boundaries $x = -a$, $x = a$ by imposing continuity

$$A \sin(ka) \pm B \cos(ka) = 0 \quad \Rightarrow A \sin(ka) = 0 \wedge B \cos(ka) = 0 \tag{3.19}$$

Since $\sin(ka)$ and $\cos(ka)$ cannot be simultaneously 0, either $A = 0$ or $B = 0$. So the two possibilities are

1. $A = 0 \Rightarrow \chi_n = B \cos(kx)$ and $k = n\pi/(2a)$, with $n = 1, 3, ...$ (<u>even</u> eigenfunctions).

2. $B = 0 \Rightarrow \chi_n = A \sin(kx)$ with $k = n\pi/(2a)$, with $n = 2, 4, ...$ (<u>odd</u> eigenfunctions).

Hence, the allowed energy levels (i.e. the eigenvalues of the Hamiltonian) are

$$E_n = \frac{\hbar^2 \pi^2 n^2}{8ma^2}, \qquad \text{with } n = 1, 2, 3, .. \tag{3.20}$$

and the wavefunctions (i.e. the eigenfunctions of the Hamiltonian) are

$$\chi_n(x) = \sqrt{\frac{1}{a}} \begin{cases} \cos \frac{n\pi x}{2a} & n = 1, 3, ... \\ \sin \frac{n\pi x}{2a} & n = 2, 4, ... \end{cases} \tag{3.21}$$

where it is straightforward to verify that $A = B = \sqrt{1/a}$ derives from the normalisation condition

$$\int_{-a}^{a} |\chi_n(x)|^2 \, dx = 1. \tag{3.22}$$

In Fig. 3.2 we present an illustration of the lowest energy wavefunctions. Notice that, unlike classical mechanics, the lowest energy state E_1 is non-zero. The lowest energy state of any system is called ground state or vacuum state. Wavefunctions alternate between even $(n = 1, 3, ...)$ and odd $(n = 2, 4, ...)$ under reflection on the y axis. Also, the wavefunction χ_n has $(n + 1)$ zeros (or nodes) where $\rho_n(x) = |\chi_n(x)|^2$ vanishes. Finally, in the limit $n \to \infty$ the probability approaches a constant, which is what we would expect from the classical result. This is yet another example of the *correspondence principle*.

The infinite square well was a rather simple and nice example. We have an infinite well, and the particle is well-contained inside the box. The solutions just look like standing waves on a string with two fixed end points — something we are familiar with. Note that

$$\chi_n(-x) = (-1)^{n+1} \chi_n(x). \tag{3.23}$$

This is actually a general feature of energy eigenfunctions of a symmetric potential. This is known as *parity*.
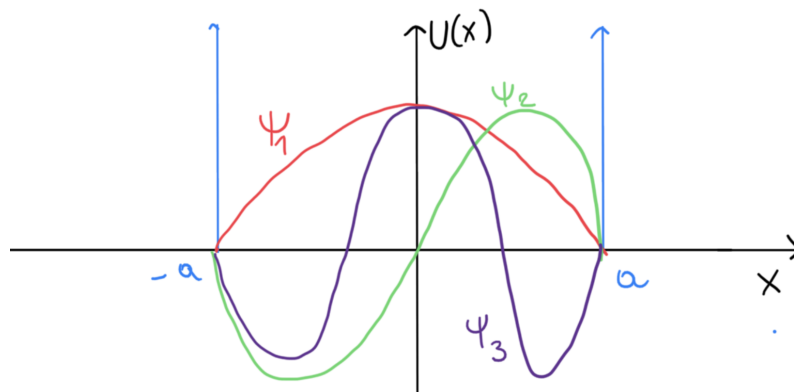
Figure 3.2: Lowest energy wavefunctions of the infinite square well.

Proposition: Suppose that the spectrum of bound states is non-degenerate

$$E_i \neq E_j \qquad \text{for} \qquad i \neq j. \tag{3.24}$$

If $U(x) = U(-x)$ then the eigenfunctions of the Hamiltonian must be either even or odd.

Proof: If $U(x) = U(-x)$, then the time-independent Schrödinger equation is reflection invariant. Therefore if $\chi(x)$ is a solution with eigenvalue E, $\chi(-x)$ is also a solution with the same eigenvalue. The non-degeneracy of the spectrum implies that the two solutions must be the same up to a multiplicative factor, *i.e.* $\chi(-x) = \alpha\chi(x)$ for some non-zero complex α. For consistency

$$\chi(x) = \chi(-(-x)) = \alpha\chi(-x) = \alpha^2\chi(x). \tag{3.25}$$

Thus $\alpha^2 = 1 \Rightarrow \alpha = \pm 1$ and $\chi(-x) = \pm\chi(-x)$, i.e. all stationary state wavefunctions must be either even or odd. $\qquad \square$

3.3.2   Finite potential well

In the case of a finite potential well, shown in Fig. 3.3,

$$U(x) = \begin{cases} 0 & |x| \leq a \\ U_0 & |x| > a, \end{cases} \tag{3.26}$$

the stationary states obey

$$-\frac{\hbar^2}{2m}\chi''(x) \; + \; U(x)\chi(x) = E\chi(x). \tag{3.27}$$

42
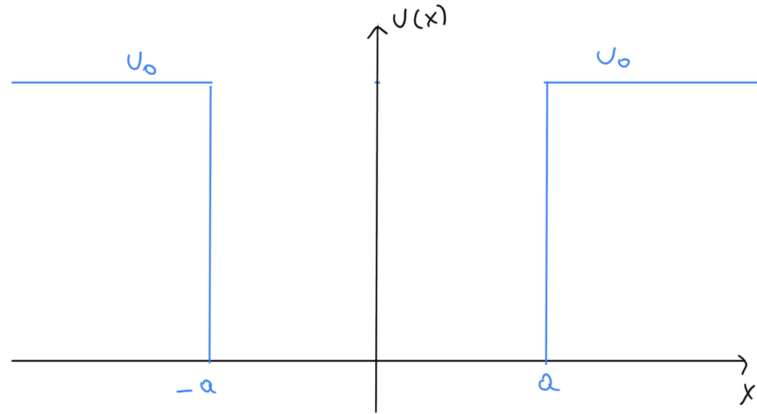
Figure 3.3: The finite square well.

Consider *even-parity* bound states[1]

$$\chi(-x) \;=\; \chi(x)$$

with energy E such that $0 \le E \le U_0$, and define the real constants

$$k = \sqrt{\frac{2mE}{\hbar^2}} \;\ge\; 0 \qquad\qquad \kappa = \sqrt{\frac{2m(U_0 - E)}{\hbar^2}} \;\ge 0. \tag{3.28}$$

For $|x| < a$, the Schrödinger equation becomes,

$$\chi''(x) + k^2\chi(x) = 0, \tag{3.29}$$

and the general solution takes the oscillatory form,

$$\chi(x) \;=\; A\cos(kx) + B\sin(kx). \tag{3.30}$$

Imposing the <u>even</u> parity condition $\chi(-x) = \chi(x)$, we get

$$B = 0 \qquad \Rightarrow \qquad \chi(x) = A\cos(kx). \tag{3.31}$$

For $|x| \ge a$

$$\chi''(x) - \kappa^2\chi(x) = 0, \tag{3.32}$$

and the general solution for $x > +a$ has the form,

$$\chi(x) \;=\; C\exp(+\kappa x) + D\exp(-\kappa x). \tag{3.33}$$

---

[1]The odd-parity bound states can be determined in a similar way, you can do it at home as an exercise.

Normalisability yields $C = 0$ thus,

$$\chi(x) \quad = \quad D \exp(-\kappa x) \tag{3.34}$$

for $x > +a$. Similarly for $x < -a$ (by even parity) we have,

$$\chi(x) \quad = \quad D \exp(+\kappa x) \tag{3.35}$$

Imposing continuity of $\chi(x)$ at $x = \pm a$ gives,

$$A \cos(ka) \quad = \quad D \exp(-\kappa a) \tag{3.36}$$

and continuity of $\chi'(x)$ at $x = \pm a$ gives,

$$-kA \sin(ka) \quad = \quad -\kappa D \exp(-\kappa a). \tag{3.37}$$

Dividing Eq. (3.37) by (3.36) yields,

$$k \tan(ka) \quad = \quad \kappa \tag{3.38}$$

From the definitions in Eq. (3.28) we find a second equation relating k and κ,

$$k^2 + \kappa^2 \quad = \quad \frac{2mU_0}{\hbar^2} \tag{3.39}$$

Now define rescaled variables $\xi = ka$ and $\eta = \kappa a$ and the constant $r_0 = a\sqrt{(2mU_0)/\hbar^2}$. Eqs. (3.38) and (3.39) become,

$$\begin{cases} \xi \tan \xi = \eta \\ \xi^2 + \eta^2 = r_0^2 \end{cases} \tag{3.40}$$

It is not possible to solve these transcendental equations in closed form. Instead one may easily establish some qualitative features of the solutions via a graphical solution as shown in Fig. 3.4. Here the two equations are plotted in the (ξ, η)-plane. The solutions correspond to the intersection points $\{\xi_1, \xi_2, \ldots, \xi_p\}$, with $p \geq 1$. We observe that the number of solutions (i.e. the numbers of discrete eigenvalues of the Hamiltonian) increases with the width and the depth of the well, as r_0 grows. Each solution determines an energy level via,

$$E_n \quad = \quad \frac{\hbar^2 \xi_n^2}{2ma^2} \qquad \text{with } n = 1, \ldots, p. \tag{3.41}$$

We always have at least one solution for $U_0 > 0$. In fact it can be proved that attractive potentials in one dimension <u>always</u> have at least one bound state. From the graph we see that,

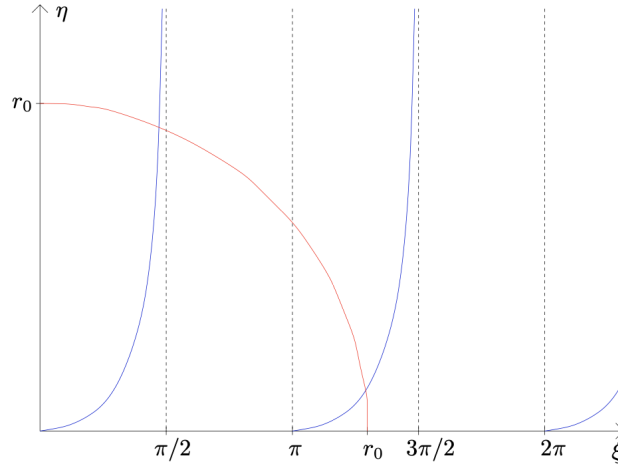$$(n-1)\pi \quad \leq \xi_n \leq \quad \left(n - \frac{1}{2}\right)\pi$$

Figure 3.4: Graphical solution of Eq. (3.40).

The limit of infinite square well $U_0 \to \infty$ implies $r_0 \to \infty$ and therefore $\xi_n \to (n - 1/2)\pi$. The resulting energy levels are given by

$$E_n \quad = \quad \frac{\hbar^2 \xi_n^2}{2ma^2} = \frac{\hbar^2 (2n - 1)^2 \pi^2}{8ma^2}. \tag{3.42}$$

This agrees with earlier result for even levels of infinite well of width $2a$, Eq. (3.20). As an exercise, you may explicitly check that wavefunction goes over to the wavefunctions of an infinite square potential well in the limit of infinite well $U_0 \to \infty$. Also, you may use boundary conditions (3.36) and (3.37) to eliminate constant D in terms of A and determine the constant A by imposing the normalisation condition,

$$\int_{-\infty}^{+\infty} |\chi(x)|^2 dx \quad = \quad 1$$

Resulting ground state probability distribution $|\chi_1(x)|^2$ is plotted in Fig. 3.5. Note that there is a non-zero probability of finding particle in the classically forbidden region $|x| > a$, which is not the case for the infinite potential well.

The solutions for the wavefunctions in case $E > U_0$ are not dealt with explicitly in this section, as they are an example of scattering states, which we will deal with in the next two sections.
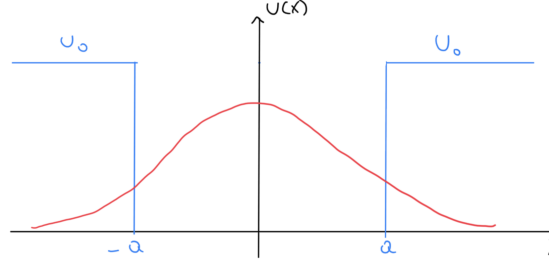
Figure 3.5: Ground state probability density for the finite square potential well.

3.3.3   The harmonic oscillator

The harmonic oscillator potential (see left panel of Fig. 3.6) is crucial in modelling the dynamics of many physics systems.

$$U(x) \;\; = \;\; \frac{1}{2}k\,x^2 = \frac{1}{2}m\omega^2 x^2, \tag{3.43}$$

where k is the often referred to as elastic constant and $\omega = \sqrt{k/m}$ is the harmonic oscillator angular frequency. Let us remind ourselves the solution for a particle subject to an harmonic potential in classical mechanics (which you have seen in Part IA Dynamics and Relativity). Newton's second law implies

$$\frac{d^2 x(t)}{dt^2} \;\; = \;\; -\omega^2 x(t). \tag{3.44}$$

The general solution is given by

$$x(t) \;\; = \;\; A\,\sin(\omega t) + B\,\cos(\omega t) \tag{3.45}$$

The particle oscillates around the minimum of the potential at $x = 0$, with period $T = 2\pi/\omega$.

In quantum mechanics, the stationary states are described by TISE

$$-\frac{\hbar^2}{2m}\chi''(x) \;+\; \frac{1}{2}m\omega^2 x^2\,\chi(x) \;\; = \;\; E\chi(x) \tag{3.46}$$

subject to the normalisability condition,

$$\int_{-\infty}^{+\infty} dx\,|\chi(x)|^2 \;\; \mathcal{N} < \;\; \infty \quad \text{and} \quad \mathcal{N} \neq 0$$

Given that a particle subject to an harmonic potential is a bound state, we expect to find a discrete set of normalisable eigenfunctions and, because of the potential is symmetric,

46
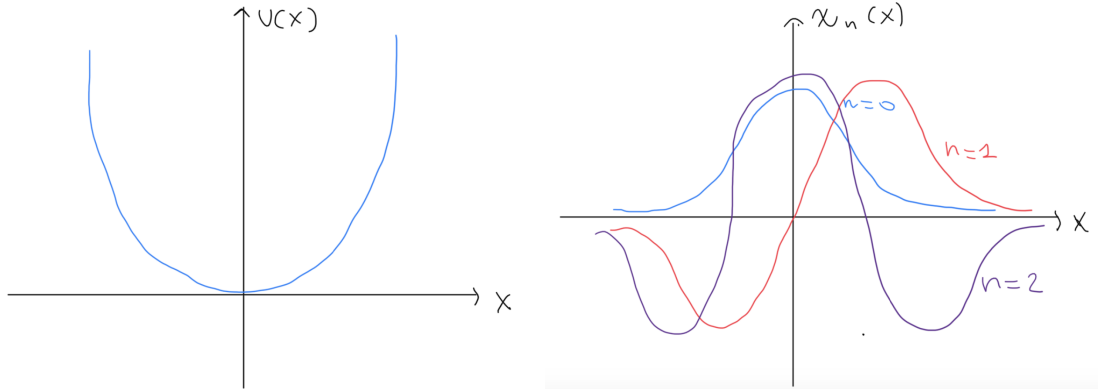
Figure 3.6: The harmonic oscillator potential (left) and lowest energy eigenfunctions (right).

we expect to find eigenfunctions that are either even or odd. To solve the above equation, we define the rescaled variables

$$\xi^2 = \frac{m\omega}{\hbar}\, x^2 \qquad \text{and} \quad \epsilon = \frac{2E}{\hbar\omega}. \tag{3.47}$$

In terms of these variables Eq. (3.46) becomes,

$$-\frac{d^2\chi(\xi)}{d\xi^2} + \xi^2\chi(\xi) = \epsilon\chi(\xi) \tag{3.48}$$

Let us start with a the special case, $\epsilon = 1$, which corresponds to the energy eigenvalue that we indicate by $E_0 = \hbar\omega/2$. By inspection, we find normalisable solution

$$\chi_0(x) = \exp\left(-\frac{1}{2}\xi^2\right) \tag{3.49}$$

If we plug Eq. (3.49) into Eq. (3.48) for $\epsilon = 1$, we see that $\chi_0(\xi)$ is indeed a solution of the rescaled TISE, as

$$\frac{d\chi(\xi)}{d\xi} = -\xi\, \exp\left(-\frac{1}{2}\xi^2\right)$$

$$\frac{d^2\chi(\xi)}{d\xi^2} = (\xi^2 - 1)\, \exp\left(-\frac{1}{2}\xi^2\right)$$

$$\Rightarrow \qquad -\frac{d^2\chi(\xi)}{d\xi^2} + \xi^2\chi(\xi) = \chi(\xi) \qquad \square$$

The corresponding eigenfunction $\chi_0(x)$ with energy $E_0 = \hbar\omega/2$ can be written as

$$\chi_0(x) = A\, \exp\left(-\frac{m\omega}{2\hbar}x^2\right), \tag{3.50}$$

47

in which we have added an overall contact A that is completely determined by requiring that $\chi_0(x)$ is normalised to 1.

To find the other eigenvalues and corresponding eigenfunction of the Hamiltonian, we now look for the general solution of the form,

$$\chi(\xi) = f(\xi) \, \exp\left(-\frac{1}{2}\xi^2\right)$$

$$\Rightarrow \qquad \frac{d\chi}{d\xi} = \left(\frac{df}{d\xi} - \xi \, f\right) \exp\left(-\frac{1}{2}\xi^2\right)$$

$$\Rightarrow \qquad \frac{d^2\chi}{d\xi^2} = \left(\frac{d^2 f}{d\xi^2} - 2\xi\frac{df}{d\xi} + (\xi^2 - 1) \, f\right) \exp\left(-\frac{1}{2}\xi^2\right),$$

where I omitted the dependence of the functions χ and f on ξ. Then (3.48) becomes,

$$-\frac{d^2 f}{d\xi^2} + 2\xi\frac{df}{d\xi} + (1 - \epsilon)f = 0 \tag{3.51}$$

You can easily check that $f = 1$ is a solution when $\epsilon = 1$.

If we now apply the standard power series method to find the general form of f ($\xi = 0$ is a regular point), we set,

$$f(\xi) = \sum_{n=0}^{\infty} a_n \, \xi^n \tag{3.52}$$

To plug the series (3.52) into (3.51), we compute

$$f(\xi) = \sum_{n=0}^{\infty} a_n \, \xi^n \qquad (3.52)$$

$$\frac{df}{d\xi} = \sum_{n=0}^{\infty} n \, a_n \, \xi^{n-1}$$

$$\xi\frac{df}{d\xi} = \sum_{n=0}^{\infty} n \, a_n \, \xi^n$$

Then

$$\frac{d^2 f}{d\xi^2} = \sum_{n=0}^{\infty} n \, (n-1) \, a_n \, \xi^{n-2} = \sum_{n=0}^{\infty} (n+1) \, (n+2) \, a_{n+2} \, \xi^n$$

Finally,

$$\frac{d^2 f}{d\xi^2} - 2\xi\frac{df}{d\xi} + (\epsilon - 1)f = \sum_{n=0}^{\infty} [(n+1)(n+2) \, a_{n+2}$$
$$-2n \, a_n + (\epsilon - 1)a_n] \, \xi^n$$

48

Thus

$$\frac{d^2 f}{d\xi^2} \;-\; 2\xi \frac{df}{d\xi} \;+\; (\epsilon - 1)f \;=\; 0 \qquad (3.51)$$

implies,

$$(n+1)(n+2)\, a_{n+2} - 2n\, a_n + (\epsilon - 1)a_n \;=\; 0$$

$$\Rightarrow \qquad a_{n+2} \;=\; \frac{(2n - \epsilon + 1)}{(n+1)(n+2)}\, a_n \qquad\qquad (3.53)$$

It is important to notice that the harmonic potential is reflection invariant and therefore $\chi(-x) = \pm\chi(x)$, which implies $f(-\xi) = \pm f(\xi)$. Hence we have a further constraint on the coefficient of the series. With $m = 1, 2, 3\ldots$, we can set,

- **either** $a_n = 0$ for $n = 2m - 1 \Leftrightarrow f(-\xi) = f(\xi)$.

- **or** $a_n = 0$ for $n = 2m \Leftrightarrow f(-\xi) = -f(\xi)$.

Finally, the series in (3.52) can either be finite or infinite. There are two possibilities

- The series (3.52) terminates. In other words $\exists\, N > 0$ such that $a_n = 0 \;\forall\, n > N$.

- The series (3.52) does not terminate. In other words $\nexists\, N > 0$ such that $a_n = 0 \;\forall\; n > N$.

Proposition If the series (3.52) does not terminate, the solution does not yield normalisable eigenfunctions. As a consequence, the series must terminate.

Proof:
Suppose series (3.52) does not terminate. The consider the large-ξ behaviour of the function,

$$f(\xi) \;=\; \sum_{n=0}^{\infty} a_n\, \xi^n.$$

This is determined by the asymptotic behaviour of the coefficients a_n as $n \to \infty$. If the series does not terminate then (3.53) determines the asymptotic behaviour of the coefficients as,

$$\frac{a_{n+2}}{a_n} \;\to\; \frac{2}{n}$$

This is identical to the asymptotic behaviour of the coefficients of the Taylor series for the function,

$$\exp(+\xi^2) \; = \; \sum_{m=0}^{\infty} \frac{\xi^{2m}}{m!} \tag{3.54}$$

Indeed if we write the series as,

$$\exp(+\xi^2) \; = \; \sum_{n=0}^{\infty} b_n \, \xi^n$$

with coefficients,

$$\begin{aligned} b_n \; &= \; \frac{1}{m!} && \text{for} \;\; n = 2m \\ &= \; 0 && \text{for} \;\; n = 2m+1 \end{aligned}$$

we immediately find (for $n = 2m$)

$$\frac{b_{n+2}}{b_n} \; = \; \frac{(n/2)!}{(n/2+1)!} \; = \; \frac{2}{n+2} \; \rightarrow \; \frac{2}{n}$$

as $n \to \infty$. The fact that the coefficients of the two series (3.52) and (3.54) have the same behaviour as $n \to \infty$ means that the respective sums have the same asymptotics as $\xi \to \infty$. Thus, if the series does not terminate, we must have,

$$f(\xi) \; \sim \; \exp(+\xi^2)$$

or equivalently,

$$\chi(\xi) = f(\xi) \, \exp\left(-\frac{\xi^2}{2}\right) \; \sim \; \exp\left(+\frac{\xi^2}{2}\right)$$

as $\xi \to \infty$, which corresponds to a non-normalisable wavefunction. $\square$.

Therefore the series must terminate and $\exists$ an integer $N \geq 0$ such that $a_{N+2} = 0$ with $a_N \neq 0$. Thus, from (3.53) we find,

$$(2N - \epsilon + 1) \; = \; 0$$

Recalling that $\epsilon = 2E/\hbar\omega$ we immediately obtain the energy spectrum of the quantum harmonic oscillator,

$$E = E_N \; = \; \left(N + \frac{1}{2}\right) \hbar\omega$$

The ground state energy is given by $E_0 = \hbar\omega/2$. The energy levels are equally spaced with $E_{N+1} - E_N = \hbar\omega$. This means that the system can absorb or emit photons whose angular frequency is an integer multiple of $\omega \Rightarrow$ equally-spaced spectral lines.

The corresponding wavefunction is

$$\chi_N(x) \;=\; f_N(\xi) \, \exp\left(-\frac{\xi^2}{2}\right) \tag{3.55}$$

Notice that $f_N(\xi)$ is an even/odd function of $\xi = \sqrt{m\omega x^2/\hbar}$ for N even/odd,

$$\chi_N(-x) \;=\; (-1)^N \, \chi_N(x)$$

$f_N(\xi)$ is an N'th order polynomial (known as the N'th Hermite polynomial) in ξ and therefore the wavefunction has N nodes or zeros. The first few levels (see the right plot on Fig.3.6) are given explicitly in the table below.

N	E_N	$\chi_N(\xi)$
0	$\frac{1}{2}\hbar\omega$	$\exp\left(-\frac{\xi^2}{2}\right)$
1	$\frac{3}{2}\hbar\omega$	$\xi \, \exp\left(-\frac{\xi^2}{2}\right)$
2	$\frac{5}{2}\hbar\omega$	$\left(1 - 2\xi^2\right) \, \exp\left(-\frac{\xi^2}{2}\right)$
3	$\frac{7}{2}\hbar\omega$	$\left(\xi - \frac{2}{3}\xi^3\right) \, \exp\left(-\frac{\xi^2}{2}\right)$

In fact one can derive a general expression (see e.g. Schiff, "Quantum Mechanics", 3rd edition if you are interested in the explicit proof):

$$f_n(\xi) \;=\; (-1)^n \, e^{\xi^2} \, \frac{d^n}{d\xi^n} e^{-\xi^2} \tag{3.56}$$

3.4   Free particle and Gaussian wavepacket

A free particle solution corresponds to a solution of the TISE with $U(x) \equiv 0$. From the discussion in the previous section, we see immediately that the free particle is problematic. Indeed it can be seen as the limit $a \to \infty$ of a infinite square well potential, in which the normalisation condition (3.22) cannot be satisfied. Let us go through the TISE solution to see where the problem explicitly arises and how it can be solved.

For $U(x) = 0$ the stationary state wavefunction $\chi(x)$ satisfies ,

$$-\frac{\hbar^2}{2m}\chi''(x) \;=\; E\chi(x) \tag{3.57}$$

This equation has a plane-wave solution, which in one spatial dimension ($x \in \mathbb{R}$)

$$\chi(x) \;=\; A \, \exp(ikx). \tag{3.58}$$

This satisfies (3.57) provided

$$E \;=\; \frac{\hbar^2 k^2}{2m}$$

The complete wavefunction,

$$
\begin{aligned}
\psi_k(x,t) \;&=\; \chi(x)\exp\left(-\frac{iEt}{\hbar}\right) \\
&=\; A\,\exp(ikx)\;\times\;\exp\left(-i\frac{\hbar k^2 t}{2m}\right)
\end{aligned}
\tag{3.59}
$$

coincides with the "De Broglie wave" that we introduced in (2.23). Such a solution is non-normalisable and thus does not give an acceptable probability density. There are several ways to resolve this problem. One way is to treat the plane wave solution ψ_k as a limiting case of a Gaussian wave packet describing a localised particle, as we will see in Sect. 3.3.1. Alternatively, the plane wave solution can be interpreted as describing a beam of particles rather than a single particle. We will discuss this second possible interpretation in Sect. 3.3.2.

3.4.1   Gaussian wavepacket

Starting from Eq. (3.59), we can construct new solutions of the Schrödinger equation by taking a linear superposition of $\psi_k(x,t)$. Given that k is a continuous variable we can make a linear superposition by integration, as in Eq. (**??**),

$$
\begin{aligned}
\psi(x,t) \;&=\; \int dk\; A(k)\,\psi_k(x,t) \\
&=\; \int dk\; A(k)\;\exp(ikx)\;\exp\left(-i\frac{\hbar k^2}{2m}t\right)
\end{aligned}
\tag{3.60}
$$

where $A(k)$ is a function of k that should go to zero sufficiently fast that the integral exists. The Gaussian wavepacket corresponds to the choice,

$$A_{\mathrm{GP}}(k) \;=\; \exp\left[-\frac{\sigma}{2}(k-k_0)^2\right] \tag{3.61}$$

where $\sigma > 0$, which looks like a Gaussian distribution of wave numbers k centred at $k = k_0$ with width $\sim 1/\sqrt{\sigma}$ (see Fig. 3.7) The resulting wavefunction can be obtained by substituting (3.61) for $A(k)$ in (3.60),

$$\psi_{\mathrm{GP}}(x,t) \;=\; \int_{-\infty}^{+\infty} dk\;\exp\left(F(k)\right)$$
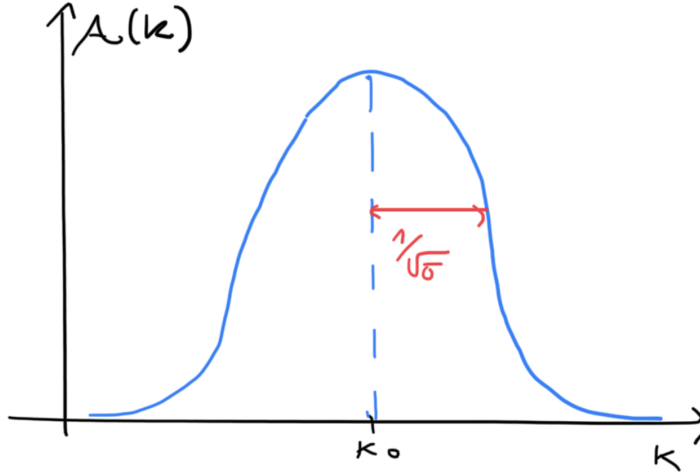
Figure 3.7: Gaussian wave packet.

where GP stands for Gaussian wavePacket and the exponent in the integrand is,

$$
\begin{aligned}
F(k) &= -\frac{\sigma}{2}(k - k_0)^2 + i\left(kx - \frac{\hbar k^2}{2m}t\right) \\
&= -\frac{1}{2}\left(\sigma + \frac{i\hbar t}{m}\right)k^2 + (k_0\sigma + ix)k - \frac{\sigma}{2}k_0^2.
\end{aligned} \tag{3.62}
$$

Completing the square gives,

$$
F(k) = -\frac{\alpha}{2}\left(k - \frac{\beta}{\alpha}\right)^2 + \frac{\beta^2}{2\alpha} + \delta, \tag{3.63}
$$

where

$$
\alpha = \sigma + \frac{i\hbar t}{m} \qquad \beta = k_0\sigma + ix \qquad \delta = -\frac{\sigma}{2}k_0^2 \tag{3.64}
$$

Hence,

$$
\begin{aligned}
\psi_{\mathrm{GP}}(x,t) &= \exp\left(\frac{\beta^2}{2\alpha} + \delta\right)\int_{-\infty}^{+\infty} dk \exp\left(-\frac{1}{2}\alpha\left(k - \frac{\beta}{\alpha}\right)^2\right) \\
&= \exp\left(\frac{\beta^2}{2\alpha} + \delta\right)\int_{-\infty-i\nu}^{+\infty-i\nu} d\tilde{k} \exp\left(-\frac{1}{2}\alpha\tilde{k}^2\right)
\end{aligned}
$$

where $\tilde{k} = k - \beta/\alpha$ and $\nu = \Im[\beta/\alpha]$. The integral can be related to the standard Gaussian integral

$$
\mathcal{I}(a) = \int_{-\infty}^{+\infty} dx \ \exp(-a\,x^2) = \sqrt{\frac{\pi}{a}} \tag{3.65}
$$

53

by a straightforward application of the Cauchy residue theorem. The result is,

$$\psi_{\text{GP}}(x,t) \;=\; \sqrt{\frac{2\pi}{\alpha}} \; \exp\left(\frac{\beta^2}{2\alpha} + \delta\right). \tag{3.66}$$

This wavefunction decays exponentially at $x \to \pm\infty$ and is therefore normalisable. The resulting position probability density is,

$$\rho_{\text{GP}}(x,t) \;=\; |\bar{\psi}_{\text{GP}}(x,t)|^2 = \bar{\psi}^*_{\text{GP}}(x,t)\bar{\psi}_{\text{GP}}(x,t) \tag{3.67}$$

where $\bar{\psi}_{\text{GP}}$ is the normalised wavefunction corresponding to $\psi_{\text{GP}}(x,t)$. After some algebra, we obtain,

$$\rho_{\text{GP}}(x,t) \;=\; \frac{C}{\sqrt{\sigma^2 + \frac{\hbar^2 t^2}{m^2}}} \; \exp\left[-\frac{\sigma\left(x - \frac{\hbar k_0}{m}t\right)^2}{\sigma^2 + \frac{\hbar^2 t^2}{m^2}}\right].$$

Note that it is straightforward to prove (you can do it at home as an exercise) that the constant C is fixed by the normalisation condition,

$$\int_{-\infty}^{+\infty} dx \; \rho_{\text{GP}}(x,t) = 1 \qquad \Rightarrow \qquad C = \sqrt{\frac{\sigma}{\pi}}. \tag{3.68}$$

Note that $\rho_{\text{GP}}(x,t)$ defines a *Gaussian probability distribution* for the position of the particle, as represented in Fig. 3.8. The centre $\langle x \rangle$ of the distribution corresponds to the average value of position:

$$\langle x \rangle \;=\; \frac{\hbar k_0}{m} t, \tag{3.69}$$

which moves constant speed

$$v \;=\; \frac{\hbar k_0}{m} = \frac{\langle p \rangle}{m}$$

Here $\langle p \rangle = \hbar k_0$ denotes the average value of the momentum. The width of the distribution, Δx (also known as the standard deviation) corresponds to the uncertainty in the measurement of position,

$$\Delta x \;=\; \sqrt{\langle x^2 \rangle - \langle x \rangle^2} = \sqrt{\frac{1}{2}\left(\sigma + \frac{\hbar^2 t^2}{m^2 \sigma}\right)} \tag{3.70}$$

and it increases with time, which means that the Gaussian wavepacket is not a stationary state, as it is the case for a generic superposition of stationary states, as we discussed in
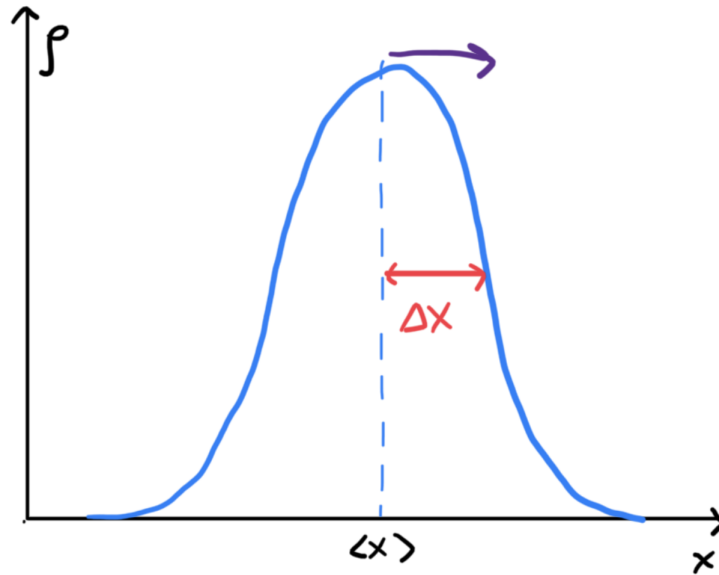
Figure 3.8: The probability distribution for the particle to be at the position x.

Sect. 2.5. Physically, the Gaussian wavepacket corresponds to a state in which the particle is localised near the point $\langle x \rangle$ with an uncertainty Δx in the measurement of its position. Not only the Gaussian wave is moving, it is also spreading, becoming more and more delocalised. The plane wave solution $\psi_k(x,t)$ is a limiting case of the Gaussian wavepacket where the uncertainty in position Δx becomes infinite. This an idealised state in which the momentum takes the definite value $p = \hbar k$. The uncertainty in the momentum of the particle, Δp, therefore vanishes. This is related to the Heisenberg uncertainty principle, which we will discuss in Chapt. 4, when we will show that the Gaussian wavepacket is the state with minimum uncertainty.

3.4.2   Beam interpretation

Another way to use the non-normalisable De Broglie plane waves is to resurrect them by endowing wavefunctions of this kind with a different interpretation. Rather than thinking of them as quantum probabilities for a single particle, we will instead consider them as describing a continuous beam of particles, with the probability density now interpreted as the average density of particles.

Basically we can describe the free particle wavefunctions, Eq. (3.59) as describing a beam of particles of momentum

$$p = \hbar k$$

and energy

$$E = \hbar\omega = \frac{\hbar^2 k^2}{2m} = \frac{p^2}{2m}.$$

The probability density

$$\rho(x,t) = |\psi_k(x,t)|^2 = |A|^2 \tag{3.71}$$

is now interpreted as the constant average density of particles. The probability current is given by

$$\begin{aligned}
j_k(x,t) &= -\frac{i\hbar}{2m}\left(\psi_k^* \frac{d\psi_k}{dx} - \psi_k \frac{d\psi_k^*}{dx}\right)\\
&= -\frac{i\hbar}{2m} \times |A|^2 \times 2ik\\
&= |A|^2 \times \frac{\hbar k}{m} = |A|^2 \times \frac{p}{m}\\
&= \text{average density} \times \text{velocity}\\
&= \text{average flux of particles}, \tag{3.72}
\end{aligned}$$

where this last expression has the interpretation of the average density of particles multiplied by their velocity or, alternatively, as the average flux of particles.

3.5   Scattering states

Consider a particle scattering on a potential barrier like the one shown (in blue) in Fig. 3.9, where U_0 is the potential barrier height and a is its width. An incoming particle of mass m and total energy E hits the barrier from left. In classical mechanics, if $E > U_0$ the particle gets over the barrier and proceeds to $x = +\infty$. On the other hand, if $E < U_0$ the particle is reflected back towards $x = -\infty$. Instead, in quantum mechanics, if we consider a localised Gaussian wavepacket with normalised wavefunction, $\psi_{\mathrm{GP}}(x,t)$

$$\int_{-\infty}^{+\infty} |\psi_{\mathrm{GP}}(x,t)|^2 \, dx = 1. \tag{3.73}$$

The wavepacket is centred at $x = x_0(t) << 0$ at initial time $t = 0$ with average momentum $\langle p \rangle > 0$, as shown in Fig. 3.9 (in red). If we evolve the wave functions according to the TDSE to get final state wavefunction for $t >> 0$ we end up with the resulting probability distribution shown in green in the same Fig. 3.9, *i.e.* the function can be both transmitted and reflected, independently of whether $E > U_0$ or $E < U_0$. We define the reflection and transmission coefficients,

$$\begin{aligned}
R &= \lim_{t\to\infty} \int_{-\infty}^{0} |\psi_{\mathrm{GP}}(x,t)|^2 \, dx\\
T &= \lim_{t\to\infty} \int_{0}^{+\infty} |\psi_{\mathrm{GP}}(x,t)|^2 \, dx, \tag{3.74}
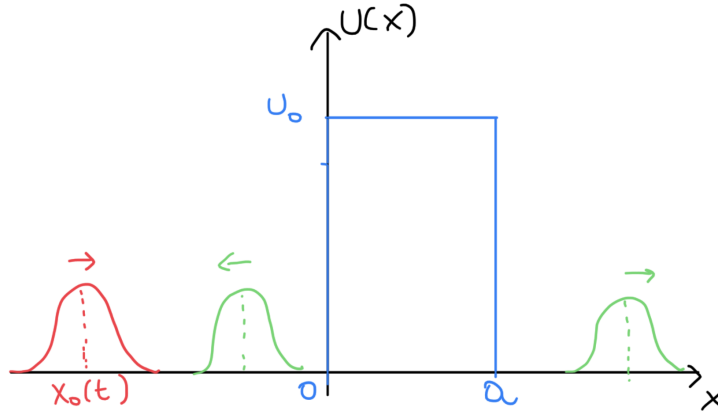\end{aligned}$$

Figure 3.9: The potential barrier (blue), the initial state of a wave packet at the time t (red), the final state of a wave packet (green).

which measure the probabilities of the particle being reflected or transmitted respectively. As total probability is conserved as we have,

$$R + T \;=\; \int_{-\infty}^{+\infty} |\psi_{\mathrm{GP}}(x,t)|^2 \; dx = 1 \tag{3.75}$$

In practice this is too mathematically challenging so will work with *non-normalisable* stationary states (plane waves) instead using the beam interpretation, which we described in Sect. 3.3.2. Fortunately both approaches yield the same answers.

Let's take the plane-wave solution,

$$\psi_k(x,t) \;=\; \chi_k(x) \; \exp\left(-i\frac{\hbar k^2 t}{2m}\right) \tag{3.76}$$

where $\chi_k(x) = A \, \exp(ikx)$ is interpreted as a beam of particles with momentum $p = \hbar k$. The average density of particles is $|A|^2$. The incident particle flux (probability current) is defined as in Eq. (3.72). In what follows, we will explicitly solve the Schrödinger equation for the scattering of particles off a potential step and a potential barrier using the beam interpretation, *i.e* working with non-normalisable eigenfunctions, which are associated with a beam of particles rather than with single particles.

3.5.1   Scattering on a potential step

Consider a beam of particles of mass m scattering on the potential step shown in Fig. 3.10

$$U(x) = \begin{cases} 0 & x \leq 0 \\ U_0 & x > 0, \end{cases} \tag{3.77}$$
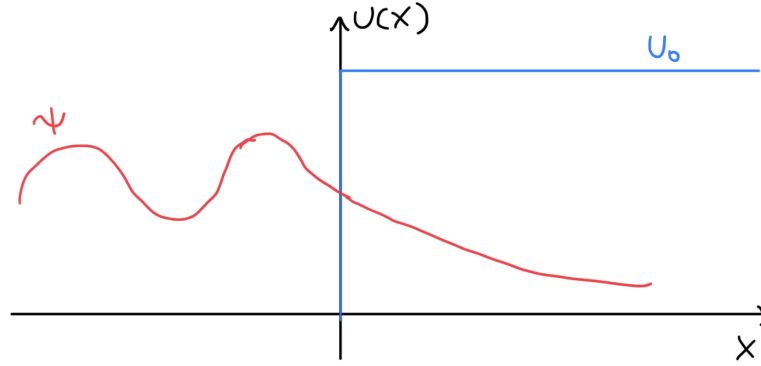
Figure 3.10: The potential step (blue) and wavefunction (red) of an incoming particle with energy $E < U_0$.

The eigenfunctions $\chi_k(x)$ obey the TISE

$$-\frac{\hbar^2}{2m}\chi_k''(x) \;+\; U(x)\chi_k(x) = E\chi_k(x) \tag{3.78}$$

We will start by considering the case where $E > U_0$ and comment on the $E \leq U_0$ at the end.

In the region $x \leq 0$, the Schrödinger equation becomes,

$$\chi_k''(x) + k^2\chi_k(x) \;\;=\;\; 0, \tag{3.79}$$

where $k = \sqrt{2mE/\hbar^2} \geq 0$. For $E > 0$, the general solution takes the form,

$$\chi_k(x) \;\;=\;\; A\exp(ikx) + B\exp(-ikx) \tag{3.80}$$

We can define $\chi_k^{(+)}(x) = A\exp(ikx)$, which corresponds to a beam of particles incident on the step from $x = -\infty$ with momentum $p = \hbar k$. The corresponding particle flux is given by the probability current,

$$j_k^{(+)}(x) \;\;=\;\; |A|^2 \;\times\; \frac{\hbar k}{m}, \tag{3.81}$$

and $\chi_k^{(-)}(x) = B\exp(-ikx)$, which corresponds to a beam of reflected particles moving to the left (*i.e.* towards from $x = -\infty$) with momentum $p = -\hbar k$. The corresponding particle flux is,

$$j_k^{(-)}(x) \;\;=\;\; -|B|^2 \;\times\; \frac{\hbar k}{m} \tag{3.82}$$

In our scattering problem we have incident particles from the left and also expect some particles to be reflected off the barrier. Thus we retain the general solution,

$$\chi_k(x) \;\; = \;\; A\exp(ikx) + B\exp(-ikx) \qquad \text{for} \;\; x \leq 0. \tag{3.83}$$

The resulting expression corresponds to a superposition of the two beams $\chi_k^{(+)}$ and $\chi_k^{(-)}$. The total flux is given by

$$
\begin{aligned}
j_k(x) \;\; &= \;\; -\frac{i\hbar}{2m}\left[\chi_k^*\frac{d\chi_k(x)}{dx} - \chi_k\frac{d\chi_k^*(x)}{dx}\right] \\
&= \;\; \frac{\hbar k}{m}\left(|A|^2 - |B|^2\right) \\
&= \;\; j_k^{(+)}(x) \; + \; j_k^{(-)}(x),
\end{aligned}
\tag{3.84}
$$

where we notice that cross-terms involving $\chi_k^{(+)}$ and $\chi_k^{(-)}$ vanish.

In the region $x > 0$, the Schrödinger equation becomes,

$$\chi_{\tilde{k}}''(x) \; + \; \tilde{k}^2\chi_{\tilde{k}}(x) \;\; = \;\; 0, \tag{3.85}$$

where,

$$\tilde{k} \;\; = \;\; \sqrt{\frac{2m(E - U_0)}{\hbar^2}} \; \geq 0. \tag{3.86}$$

The general solution is

$$\chi_{\tilde{k}}(x) \;\; = \;\; C\exp(i\tilde{k}x) + D\exp(-i\tilde{k}x) \tag{3.87}$$

As we did before, we define $\tilde{\chi}_{\tilde{k}}^{(+)}(x) = C\exp(i\tilde{k}x)$, which corresponds to a beam of particles in region $x > 0$ moving towards $x = +\infty$. This corresponds to a *transmitted wave* in the scattering problem. $\tilde{\chi}_{\tilde{k}}^{(-)}(x) = D\exp(-i\tilde{k}x)$ corresponds to a beam of particles incident on the barrier from the right (*i.e.* from $x = +\infty$). This solution is not relevant for our scattering problem and thus we set $D = 0$ and choose the solution,

$$\chi_{\tilde{k}}(x) \;\; = \;\; C\exp(i\tilde{k}x) \qquad \text{for} \;\; x > 0 \tag{3.88}$$

It remains to enforce the continuity of the stationary-state wavefunction and its derivative at $x = 0$. Comparing the solutions (3.83) and (3.88) we find,

- Continuity of $\chi(x)$ at $x = 0 \Rightarrow A + B = C$

- Continuity of $\chi'(x)$ at $x = 0 \Rightarrow ikA - ikB = i\tilde{k}C$

Solving the equations above we get,

$$B = \frac{k - \tilde{k}}{k + \tilde{k}} \, A \qquad\qquad C = \frac{2k}{k + \tilde{k}} A. \qquad\qquad (3.89)$$

If we now identify the particle flux corresponding to each component of the wave function, we have an incoming flux given by

$$j_{\text{inc}}(x) \;\; = \;\; j^{(+)}(x) = \frac{\hbar k}{m} \, |A|^2, \qquad\qquad (3.90)$$

a reflected flux

$$j_{\text{ref}}(x) \;\; = \;\; -j^{(-)}(x) = +\frac{\hbar k}{m} \, |B|^2 = \frac{\hbar k}{m} \left(\frac{k - \tilde{k}}{k + \tilde{k}} \right)^2 |A|^2 \qquad\qquad (3.91)$$

and a transmitted flux

$$j_{\text{tr}}(x) \;\; = \;\; \frac{\hbar \tilde{k}}{m} \, |C|^2 = \frac{\hbar \tilde{k}}{m} \frac{4k^2}{(k + \tilde{k})^2} \, |A|^2. \qquad\qquad (3.92)$$

They are all constant functions.

To determine the portion of the incident beam, which is reflected/transmitted we compute the corresponding probabilities,

$$
\begin{aligned}
R \;\; &= \;\; \frac{j_{\text{ref}}}{j_{\text{inc}}} = \left(\frac{k - \tilde{k}}{k + \tilde{k}} \right)^2 \\
T \;\; &= \;\; \frac{j_{\text{tr}}}{j_{\text{inc}}} = \frac{4k\tilde{k}}{(k + \tilde{k})^2}.
\end{aligned}
\qquad\qquad (3.93)
$$

Note that the undetermined constant A cancels out. We can check explicitly that $R+T = 1$. Unlike the classical case, there is still a finite probability of reflection for $E > U_0$. However as $E \to \infty$ we have $k - \tilde{k} \to 0$ which implies $R \to 0$, $T \to 1$, corresponding to the classical limit.

Finally we consider the case $E < U_0$. According to classical mechanics the particle should always be reflected. In this case the solution for $x \leq 0$, Eq. (3.83) remains unchanged. Instead, for $x > 0$ the time-independent Schrödinger equation becomes,

$$\chi''(x) - \kappa^2 \chi(x) \;\; = \;\; 0. \qquad\qquad (3.94)$$

where

$$\kappa \;\; = \;\; \sqrt{\frac{2m(U_0 - E)}{\hbar^2}} \;\; > \;\; 0 \qquad\qquad (3.95)$$

. The general solution of this equation can be written as

$$\chi_\kappa(x) \;=\; G\exp(+\kappa x) + F\exp(-\kappa x). \tag{3.96}$$

. The growing exponential diverges at $x \to +\infty$, thus is unphysical. Hence we must set $G = 0$. The solution for $x > 0$ is therefore

$$\chi_\kappa(x) \;=\; F\exp(-\kappa x) \qquad \text{for } x > 0 \tag{3.97}$$

If we now impose the boundary conditions on the solutions (3.83) and (3.97) at $x = 0$, we get

- Continuity of $\chi(x)$ at $x = 0 \Rightarrow A + B = F$.

- Continuity of $\chi'(x)$ at $x = 0 \Rightarrow ikA - ikB = -\kappa F$

Solving the above equations, we get

$$B = \left(\frac{ik+\kappa}{ik-\kappa}\right) A \qquad\qquad F = \frac{2ik}{ik-\kappa}\,A \tag{3.98}$$

If we now identify the particle flux corresponding to each component of the wave function, we get an incoming flux

$$j_{\text{inc}}(x) \;=\; j_k^{(+)}(x) = \frac{\hbar k}{m}\,|A|^2, \tag{3.99}$$

a reflected flux

$$j_{\text{ref}}(x) \;=\; -j_k^{(-)}(x) = +\frac{\hbar k}{m}\,|B|^2 = \frac{\hbar k}{m}\,|A|^2 = j_{\text{inc}}(x). \tag{3.100}$$

The transmitted particle flux vanishes, as

$$j_{\text{tr}}(x) \;=\; -\frac{i\hbar}{2m}\left[\chi_\kappa^*(x)\frac{d\chi_\kappa(x)}{dx} \;-\; \chi_\kappa(x)\frac{d\chi_\kappa^*(x)}{dx}\right] = 0. \tag{3.101}$$

Thus the whole beam is reflected, as

$$R \;=\; \frac{j_{\text{ref}}}{j_{\text{inc}}} = 1$$

$$T \;=\; \frac{j_{\text{tr}}}{j_{\text{inc}}} = 0 \tag{3.102}$$

As in the classical case, the particle is certain to be reflected. However the wavefunction is not equal to zero in the classically forbidden region, rather it decays as shown in Fig. 3.10.

61

3.5.2   Scattering off a potential barrier

Consider an incident particle with energy $E < U_0$ hitting the potential barrier of Fig. 3.9.
We look for stationary state wave function obeying,

$$-\frac{\hbar^2}{2m}\chi''(x) \; + \; U(x)\chi(x) = E\chi(x). \tag{3.103}$$

We have dropped the suffix k, as by now it is clear that we are talking of a plane wave
solutions. Let us define the real constants

$$k = \sqrt{\frac{2mE}{\hbar^2}} \; \geq \; 0 \qquad\qquad \kappa = \sqrt{\frac{2m(U_0 - E)}{\hbar^2}} \; \geq 0. \tag{3.104}$$

The solution has the general form

$$\chi(x) = \begin{cases} \exp(ikx) + A \ \exp(-ikx) & x < 0 \\ B \ \exp(-\kappa x) \; + \; C \ \exp(+\kappa x) & 0 < x < a \\ D \ \exp(+ikx) & x > a, \end{cases} \tag{3.105}$$

where A and D are coefficients of reflected and transmitted waves respectively. The coefficient of incident wave $\exp(+ikx)$ has been normalised to unity.

Let us know impose the boundary conditions:

- Continuity of $\chi(x)$ at $x = 0 \Rightarrow 1 + A = B + C$.

- Continuity of $\chi'(x)$ at $x = 0 \Rightarrow ik - ikA = -\kappa B + \kappa C$

- Continuity of $\chi(x)$ at $x = a \Rightarrow B \ \exp(-\kappa a) + C \ \exp(+\kappa a) = D \exp(ika)$.

- Continuity of $\chi'(x)$ at $x = a \Rightarrow -\kappa B \ \exp(-\kappa a) + \kappa C \ \exp(+\kappa a) = ikD \exp(ika)$

Thus we have four equations for the four unknown constants A, B, C and D, which has
solution,

$$D \;\; = \;\; \frac{-4i\kappa k}{(\kappa - ik)^2 \exp[(\kappa + ik)a] - (\kappa + ik)^2 \exp[-(\kappa - ik)a]} \tag{3.106}$$

The transmitted flux is given by

$$j_{\mathrm{tr}}(x) \;\; = \;\; \frac{\hbar k}{m} \; |D|^2. \tag{3.107}$$

On the other hand, the incident flux is given by

$$j_{\mathrm{in}}(x) \;\; = \;\; \frac{\hbar k}{m} \tag{3.108}$$

Thus the transmission probability is given as,

$$
\begin{aligned}
T = \frac{j_{\text{tr}}}{j_{\text{inc}}} &= |D|^2 \\
&= \frac{4k^2\kappa^2}{(k^2 + \kappa^2)^2 \sinh^2(\kappa a) + 4k^2\kappa^2}
\end{aligned}
\tag{3.109}
$$

In the limit in which we have low energy particle scattering on very tall barrier $U_0 - E >> \hbar^2/2ma^2$, we have $\kappa a >> 1$. In this case (3.109) simplifies to give

$$
T \simeq \left(\frac{16k^2\kappa^2}{(k^2 + \kappa^2)^2}\right) \exp(-2\kappa a) \propto \exp\left[-\frac{2a}{\hbar}\sqrt{2m(U_0 - E)}\right].
\tag{3.110}
$$

There is transmission, even though the energy lies below the top of the barrier. This is a wave phenomenon, and in quantum mechanics it is also one exhibited by particles. It is widely known as *quantum tunneling*.

Not examinable *We pause to analyse an apparent difficulty. The wavefunction does not vanish inside the barrier, and thus there appears to be some probability of finding the particle with negative kinetic energy. How can we make sense of it? We look at the uncertainty relation to remove an apparent paradox that arises from a too classical description of the process. An experiment to study the particle inside a potential barrier must be able to localise it within an accuracy $\Delta x \ll a$. Applying the Heisenberg uncertainty principle

$$
\Delta x \Delta p \geq \frac{\hbar}{2},
\tag{3.111}
$$

the measurement of the position will transfer to a particle momentum, with

$$
\Delta p \gg \frac{\hbar}{2a},
\tag{3.112}
$$

which corresponds to a transfer of energy

$$
\Delta E \gg \frac{\hbar^2}{2ma^2}.
\tag{3.113}
$$

In order to observe the negative kinetic energy, this uncertainty must be much less than $|E - U_0|$, so that

$$
\frac{\hbar^2\kappa^2}{2m} \gg \frac{\hbar^2}{2ma^2},
\tag{3.114}
$$

which implies $\kappa a \gg 1$. In this limit, the quantity to be measured $|T|^2$, Eq. (3.110), is vanishingly small. For a deeper explanation, have a look at the book by S. Gasiorowicz, "Quantum Physics" (Wiley 2003s), Section 4.3. *

3.5.3   Physics examples of quantum tunneling*

In this section we give several physics examples among an impressive number of physics and technology applications that quantum tunneling has in physics.
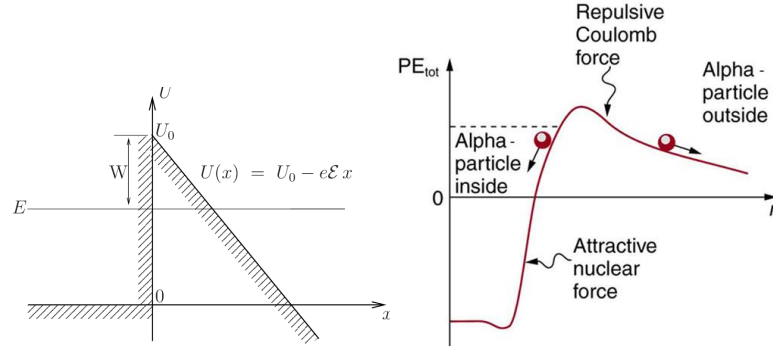

Figure 3.11: Left: the potential seen by the electrons in a metal when an electric filed is turned on. Right: the nuclear potential.

A first example of tunneling is given by *cold emission*. In Sect. 1 we discussed the photo-electric effect and we saw that electrons need a minimum energy W (the work function) to escape from a metal. We may view the most easily liberated electrons as trapped in a large box, with a potential barrier of height W. The photoelectric effect is not the only way in which electrons can be removed from a metal. They can also be removed by heating the metal, or can be removed at room temperature by applying an external electric field $\mathcal{E}$. In the latter case, the phenomenon is called *cold emission*. It occurs because the external field changes the potential seen by the electrons from a macroscopic barrier of height W to one describe by $W - e\mathcal{E}x$, where x is the distance from the walls of the box, see the left panel of Fig. 3.11 for a schematic illustration. The change creates a barrier of finite width, and electrons can tunnel through it. We define the most easily removed electrons as having an energy $E = 0$. Then the transmission coefficient is given by

$$|T|^2 \propto \exp\left[-\int_0^a \sqrt{2m(W - e\mathcal{E}x)/\hbar^2}\, dx\right] = \exp\left[-2\sqrt{2mWa^2/9\hbar^2}\right], \qquad (3.115)$$

which is known as the *Fowler-Nordheim* formula. A version of the cold electron emission described above has found an important application in the *scanning tunneling microscopes*, which are able to get a resolution of $\mathcal{O}(10^{-10})$ m.

Another example is given by the radioactive decay of nuclei. Consider the radioactive decay of an isotope N_Z^A. Here A and Z are the atomic weight and atomic number respectively (see Appendix A.2). The decay proceeds through emission of an α-particle (*i.e.* a Helium nucleus),

$$N_Z^A \quad \rightarrow \quad N_{Z-2}^{A-4} \;+\; He_2^4,$$
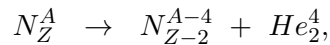
see Fig. 3.12 for an illustration. In a simple model of this process due to Gammow the α-particle feels a potential due to the other particles in the nucleus which has the form shown on the right panel of Fig. 3.11. The potential has a short-range attractive component due to the strong nuclear force and a long-range component due to the electrostatic repulsion between the protons in the α-particle and those in the nucleus. According to Gammow's model, the α-decay occurs when α-particle "tunnels" through potential barrier and the half-life of the nucleus is inversely proportional to the transmission coefficient T, so that the half-life exponentially depends on the height and width of the barrier. This model is extremely successful and it accounts for the huge range of half-lives of radioactive isotopes found in nature (and created in the lab). These range from 3×10^{-7} s to 2×10^{17} years!

Finally, a further example has to do with semiconductor devices such as electronic circuit components or integrated circuits that are designed at nanoscales; hence, the term 'nanotechnology.' For example, a diode (an electric-circuit element that causes an electron current in one direction to be different from the current in the opposite direction, when the polarity of the bias voltage is reversed) can be realized by a tunneling junction between two different types of semiconducting materials. In such a tunnel diode, electrons tunnel through a single potential barrier at a contact between two different semiconductors. At the junction, tunneling-electron current changes nonlinearly with the applied potential difference across the junction and may rapidly decrease as the bias voltage is increased. This is unlike the Ohm's law behavior that we are familiar with in household circuits. This kind of rapid behaviour (caused by quantum tunneling) is desirable in high-speed electronic devices and it is at the basis of the development of nanotechnologies.
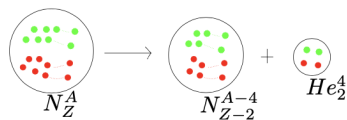

Figure 3.12: Radioactive decay.

Chapter 4

Simultaneous measurements in quantum mechanics

In the last chapter we have encountered a number of interesting properties of simple quantum systems. Some of these are accidental features of specific potentials (the even spacing of energy levels for the harmonic oscillator for example) but other seem to be more general. In this chapter, we discuss the notion of simultaneous measurements of observables and prove two important theorems about these — Ehrenfest's theorem and the Heisenberg uncertainty principle – which are among the pillars of quantum mechanics.

4.1   Commutators

In order to discuss about simultaneous measurements, it is essential to define the operation of commutation between two operators.

Definition:The commutator operator of two operators is defined as

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}. \tag{4.1}$$

Note that the commutator operator depends linearly on both entries. It is easy to verify the following identities:

$$
\begin{aligned}
{[\hat{A}, \hat{A}]} &= 0 \\
{[\hat{A}, \hat{B}]} &= -[\hat{B}, \hat{A}] \\
{[\hat{A}, \hat{B}\hat{C}]} &= [\hat{A}, \hat{B}]\hat{C} + \hat{B}\,[\hat{A}, \hat{C}] \tag{4.2} \\
{[\hat{A}\hat{B}, \hat{C}]} &= \hat{A}\,[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\,\hat{B}. \tag{4.3}
\end{aligned}
$$

The commutator plays a crucial role in describing symmetries of quantum mechanical systems, as we will see when we consider angular momentum. It also gives a way of

calibrating how close two operators are to having simultaneously determinable eigenvalues, as we will see in the following part of this chapter, in particular when discussing the uncertainty principle.

Exercise: Compute the commutator of the position and momentum operator in 1D.

The position and momentum operators are represented in the x-space in 1D as $\hat{p} = -i\hbar\partial/\partial x$ and $\hat{x} = x$. We can work out their commutator by considering their action on a general wavefunction ψ:

$$
\begin{aligned}
\hat{x}\hat{p}\psi &= -ix\hbar\frac{\partial\psi}{\partial x} \\
\hat{p}\hat{x}\psi &= -i\hbar\frac{\partial(x\psi)}{\partial x} = -i\hbar\psi - ix\hbar\frac{\partial\psi}{\partial x} \\
\Rightarrow \quad & [\hat{x},\hat{p}]\psi = i\hbar\psi,
\end{aligned}
\tag{4.4}
$$

and as this is true for all ψ we have

$$
[\hat{x},\hat{p}] = i\hbar\hat{I}, \qquad \square \tag{4.5}
$$

where $\hat{I}$ is the identity operator. Eq. (4.5) is the *canonical commutation relation*. It can be shown (though not in this course: see Part II Principles of Quantum Mechanics) that these commutation relations essentially characterise the position and momentum operators. That is, any pair of operators satisfying these relations is equivalent (in a sense that can be made precise) to position and momentum.

<u>Definition:</u> We say two Hermitian operators $\hat{A}$ and $\hat{B}$ are simultaneously diagonalisable if the space of normalisable wavefunctions has a complete basis of joint eigenfunctions $\{\psi_i\}$ *i.e.* of eigenfunctions such that $\hat{A}\psi_i = a_i\psi_i$ and $\hat{B}\psi_i = b_i\psi_i$ for some real numbers $a_i, b_i \in \mathbb{R}$.

<u>Theorem 4.1:</u> Two Hermitian operators $\hat{A}$ and $\hat{B}$ are simultaneously diagonalisable if and only if $[\hat{A}, \hat{B}] = 0$.

Proof: If $\hat{A}$ and $\hat{B}$ are simultaneously diagonalisable, the space of normalisable wavefunctions has a complete basis of joint eigenfunctions $\{\psi_i\}$. Now for any such eigenfunction

$$
[\hat{A},\hat{B}]\psi_i = \hat{A}\hat{B}\psi_i - \hat{B}\hat{A}\psi_i = (a_ib_i - b_ia_i)\psi_i = 0. \tag{4.6}
$$

Now, if the basis $\{\psi_i\}$ is complete, any wavefunction ψ can be written as $\psi = \sum_i c_i\psi_i$, so

$$
[\hat{A},\hat{B}]\psi = \sum_i c_i[\hat{A},\hat{B}]\psi_i = 0. \tag{4.7}
$$

As a result, we have the operator equation $[\hat{A}, \hat{B}] = 0$, where 0 is the zero operator, which maps any wavefunction to the zero function.

Conversely, if $[\hat{A}, \hat{B}] = 0$ and $\hat{A}\psi_i = a_i\psi_i$, then

$$0 = [\hat{A}, \hat{B}]\psi_i = \hat{A}(\hat{B}\psi_i) - a_i(\hat{B}\psi_i) \quad \Rightarrow \quad \hat{A}(\hat{B}\psi_i) = a_i(\hat{B}\psi_i) \tag{4.8}$$

so $\hat{B}\psi_i$ is also an eigenfunction of $\hat{A}$ with eigenvalue a_i. Thus $\hat{B}$ maps the eigenspace E of $\hat{A}$ with eigenvalue a_i to itself. If we indicate by $\hat{B}|_E$ for the operator $\hat{B}$ restricted to E, then clearly $\hat{B}|_E$ is a Hermitian operator on E. Thus, by Theorem 4.3, we can find a basis of E in which $\hat{B}$ acts diagonally. Since this is true for all eigenspaces of $\hat{A}$, we can find a complete basis of simultaneous eigenfunctions of $\hat{A}$ and $\hat{B}$. $\square$

4.2   Heisenberg's uncertainty principle

We define the *uncertainty* $(\Delta_\psi A)$ in a measurement of A on the state ψ by

$$(\Delta_\psi A)^2 = \langle (\hat{A} - \langle \hat{A} \rangle_\psi)^2 \rangle_\psi = \langle \hat{A}^2 \rangle_\psi - (\langle \hat{A} \rangle_\psi)^2, \tag{4.9}$$

according to the definition of the statistical variance of the probability distribution for the possible outcomes of the measurement of A on ψ.

Lemma 4.2: The uncertainty $(\Delta_\psi A) \geq 0$ and $(\Delta_\psi A) = 0$ if and only if ψ is an eigenfunction of $\hat{A}$.

Proof: We can write

$$\begin{aligned}
(\Delta_\psi A)^2 &= \langle (\hat{A} - \langle \hat{A} \rangle_\psi)^2 \rangle_\psi \\
&= ((\hat{A} - \langle \hat{A} \rangle_\psi)\psi, (\hat{A} - \langle \hat{A} \rangle_\psi)\psi)
\end{aligned} \tag{4.10}$$

If we now call $\phi = (\hat{A} - \langle \hat{A} \rangle_\psi)\psi$, we have $(\phi, \phi) \geq 0$, with equality only if $\phi = 0$, i.e.

$$\hat{A}\psi = \langle \hat{A} \rangle_\psi \psi, \tag{4.11}$$

which implies that ψ is an eigenfunction of $\hat{A}$.
Conversely, if ψ is an eigenfunction of $\hat{A}$ then $\langle \hat{A} \rangle_\psi = (\psi, \hat{A}\psi) = a(\psi, \psi) = a$ and $\langle \hat{A}^2 \rangle_\psi = (\psi, \hat{A}^2\psi) = a^2(\psi, \psi) = a^2$, so $\Delta_\psi A = \langle \hat{A}^2 \rangle_\psi - (\langle \hat{A} \rangle_\psi)^2 = a^2 - a^2 = 0$. $\square$

Theorem 4.3 (Schwarz's inequality): If ϕ and ψ are any two normalisable wavefunctions, then $|(\phi, \psi)|^2 \leq (\phi, \phi)(\psi, \psi)$. We have equality if and only if $\phi = a\psi$ some complex number a.

Proof: For any a we have

$$0 \leq (\phi - a\psi, \phi - a\psi). \tag{4.12}$$

If in particular we take $a = (\psi, \phi)/(\psi, \psi)$, we have

$$
\begin{aligned}
0 \;\leq\; & (\phi, \phi) - 2\frac{|(\psi, \phi)|^2}{(\psi, \psi)} + \frac{|(\psi, \phi)|^2}{(\psi, \psi)} \\
\Rightarrow\; & |(\phi, \psi)|^2 \leq (\phi, \phi)(\psi, \psi),
\end{aligned}
\tag{4.13}
$$

with

$$
|(\phi, \psi)|^2 = (\phi, \phi)(\psi, \psi) \Leftrightarrow \phi = a\psi,
\tag{4.14}
$$

because the norm of a wavefunction is zero only if the wavefunction is zero.   □

Theorem 4.4 (Generalised uncertainty theorem): If A and B are any two observables, and ψ is any state, then

$$
(\Delta_\psi A)\,(\Delta_\psi B) \;\geq\; \frac{1}{2}|(\psi, [\hat{A}, \hat{B}]\psi)|
\tag{4.15}
$$

Proof: We have

$$
\begin{aligned}
(\Delta_\psi A)^2 &= \langle(\hat{A} - \langle\hat{A}\rangle_\psi)^2\rangle_\psi \\
&= ((\hat{A} - \langle\hat{A}\rangle_\psi)\psi, (\hat{A} - \langle\hat{A}\rangle_\psi)\psi) \\
(\Delta_\psi B)^2 &= ((\hat{B} - \langle\hat{B}\rangle_\psi)\psi, (\hat{B} - \langle\hat{B}\rangle_\psi)\psi).
\end{aligned}
\tag{4.16}
$$

So, writing $\hat{A}' = \hat{A} - \langle\hat{A}\rangle_\psi$ and $\hat{B}' = \hat{B} - \langle\hat{B}\rangle_\psi$, we have

$$
\begin{aligned}
(\Delta_\psi A)^2(\Delta_\psi B)^2 &= (\hat{A}'\psi, \hat{A}'\psi)(\hat{B}'\psi, \hat{B}'\psi) \\
&\geq |(\hat{A}'\psi, \hat{B}'\psi)|^2 \\
&= |(\psi, \hat{A}'\hat{B}'\psi)|^2,
\end{aligned}
\tag{4.17}
$$

where in the second-to-last line we used Theorem 4.6 and in the last line we used the fact that $\hat{A}'$ is Hermitian. We now want to write $\hat{A}'\hat{B}'$ in terms of the commutator and the anti-commutator of the two operators, where the anti-commutator is defined as

$$
\{\hat{A}', \hat{B}'\} = \hat{A}'\hat{B}' + \hat{B}'\hat{A}'.
\tag{4.18}
$$

We can easily see that

$$
\hat{A}'\hat{B}' = \frac{1}{2}\left([\hat{A}', \hat{B}'] + \{\hat{A}', \hat{B}'\}\right).
\tag{4.19}
$$

Using the definition of Hermitian conjugation, and the fact that $(\hat{A}'\hat{B}')^\dagger = \hat{B}'^\dagger\hat{A}'^\dagger$, we can easily see that

$$
\begin{aligned}
(\psi, \{\hat{A}', \hat{B}'\}\psi) &= (\{\hat{A}', \hat{B}'\}^\dagger\psi, \psi) = (\{\hat{A}', \hat{B}'\}\psi, \psi) = (\psi, \{\hat{A}', \hat{B}'\}\psi)^* \quad &(4.20) \\
(\psi, [\hat{A}', \hat{B}']\psi) &= ([\hat{A}', \hat{B}']^\dagger\psi, \psi) = -([\hat{A}', \hat{B}']\psi, \psi) = -(\psi, \{\hat{A}', \hat{B}'\}\psi)^* \quad &(4.21)
\end{aligned}
$$

and that therefore $(\psi, [\hat{A}', \hat{B}']\psi)$ is purely imaginary and $(\psi, \{\hat{A}', \hat{B}'\}\psi)$ is purely real. As a result

$$|(\psi, \hat{A}'\hat{B}'\psi)|^2 = \frac{1}{4}(|(\psi, [\hat{A}', \hat{B}']\psi)|^2 + |(\psi, \{\hat{A}', \hat{B}'\}\psi)|^2) \tag{4.22}$$

Combining Eq. (4.17) and (4.22), we have

$$(\Delta_\psi A)^2 (\Delta_\psi B)^2 \geq \frac{1}{4}(|(\psi, [\hat{A}, \hat{B}]\psi)|^2, \tag{4.23}$$

and taking the square root, we have proven the theorem. $\square$

The first thing to outline is that if two observables A and B are associated to operators $\hat{A}$ and $\hat{B}$ that commute ($[\hat{A}, \hat{B}] = 0$), i. e. of two operators that can be simultaneously diagonalisable, then they can be simultaneously measured with arbitrary precision. This is why the commutator of two operators has to do with simultaneous measurements.
However, the most important corollary of the theorem that we have just proven is its application to the position and momentum operator, which gives us the famous *Heisenberg uncertainty principle*. If we take $\hat{A} = \hat{x}$ and $\hat{B} = \hat{p}$, from Eq. (4.5) we have $[\hat{A}, \hat{B}] = i\hbar\hat{I}$, and therefore:

Corollary 4.4.1 (The Heisenberg uncertainty principle)

$$\Delta_\psi \hat{x} \, \Delta_\psi \hat{p} \geq \hbar/2. \tag{4.24}$$

Thus, the smaller the uncertainty in position, $\Delta_\psi \hat{x}$, the greater the minimum possible uncertainty in momentum, $\Delta_\psi \hat{p}$, and vice versa. An intuitive way to interpret the Heisenberg's uncertainty principle is the following. To resolve particle position to accuracy Δx, we need to use a light of wavelength $\lambda \sim \Delta x$. According to the De Broglie relation this corresponds to photons with momentum of magnitude $p = h/\lambda \sim h/\Delta x$. The recoil of measured particle introduces uncertainty in its momentum of order $\Delta p \sim p \sim h/\Delta x$. Thus the estimated uncertainties obey $\Delta x \Delta p \sim h$.

The plane wave solution $\psi_k(x, t)$ corresponds to the unphysical state in which the uncertainty in position Δx becomes infinite and $\Delta p = 0$, as the momentum takes the definite value $p = \hbar k$. On the other hand, the Gaussian wavepacket defined in Eq. (3.68), with C defined in (3.68), can be shown to be the state with minimum uncertainty. This can be either done brute force by explicitly computing $(\Delta_{\psi_{\text{GP}}} x)$ and $(\Delta_{\psi_{\text{GP}}} p)$, or in a more elegant way, as we will do by proving the next two lemmas.

Lemma 4.5:   If

$$\hat{x}\psi = ia\hat{p}\psi \tag{4.25}$$

for some real parameter a, then

$$\Delta_\psi \hat{x} \, \Delta_\psi \hat{p} = \hbar/2. \tag{4.26}$$

Proof: If $\hat{x}\psi = ia\hat{p}\psi$, we have

$$
\begin{aligned}
(\psi, \{\hat{x}, \hat{p}\}\psi) &= (\psi, \hat{x}\hat{p}\psi) + (\psi, \hat{p}\hat{x}\psi) \\
&= (\hat{x}\psi, \hat{p}\psi) + (\hat{p}\psi, \hat{x}\psi) \\
&= (ia\hat{p}\psi, \hat{p}\psi) + (\hat{p}\psi, ia\hat{p}\psi) \\
&= (ia - ia)(\hat{p}\psi, \hat{p}\psi) = 0
\end{aligned}
\tag{4.27}
$$

which is the condition for the first term on the RHS of Eq. (4.22) to vanish. We also have that $\langle \hat{x} \rangle_\psi = ia\langle \hat{p} \rangle_\psi$ and, since both expectations are real, this implies that $\langle \hat{x} \rangle_\psi = 0$ and $\langle \hat{p} \rangle_\psi = 0$. Hence

$$
(\hat{x} - \langle \hat{x} \rangle_\psi)\psi = ia(\hat{p} - \langle \hat{p} \rangle_\psi)\psi,
\tag{4.28}
$$

which means we have equality in the Schwarz's inequality (Theorem 4.6) used to derive Eq. (4.24). $\square$

<u>Lemma 4.6:</u> . The condition (4.25) holds if and only if $\psi(x) = C \exp(-bx^2)$ for some constants $b, C \in \mathbb{R}$.

Proof: If $\hat{x}\psi = ia\hat{p}\psi$ for some real a, we have $x\psi = a\hbar \partial \psi / \partial x$ and so $\psi(x) = C \exp(-bx^2)$ for some real $b = -1/(2a\hbar)$, and because we have equality in Eq. (4.24) we know the uncertainty is minimised.
Conversely, any wavefunction of the form $\psi(x) = C \exp(-bx^2)$ satisfies $\hat{x}\psi = ia\hat{p}\psi$ for some real a. $\square$

Note that for the wavefunction to be normalisable, we require $b > 0$ and $C \neq 0$. We can take $C = |C| > 0$ by multiplying ψ by a phase factor (which does not alter any physical quantity: the probabilities of outcomes for any measurement are unaffected). It can be actually shown that the condition (4.25) for minimum uncertainty in Lemma 4.8 is necessary as well as sufficient, so that the normalisable minimum uncertainty states are precisely the wavefunctions defined by Gaussian wavepackets.

4.3 Ehrenfest Theorem

<u>Theorem 4.7:</u> (Ehrenfest's theorem): The expectation value $\langle \hat{A} \rangle_\psi$ of an operator $\hat{A}$ on a state ψ evolves by

$$
\frac{d}{dt}\langle \hat{A} \rangle_\psi = \frac{i}{\hbar}\langle [\hat{H}, \hat{A}] \rangle_\psi + \langle \frac{\partial \hat{A}}{\partial t} \rangle_\psi
\tag{4.29}
$$

Proof: In 1D, omitting the explicit dependence of the wavefunction $\psi(x, t)$ on x and

t, we have

$$
\begin{aligned}
\frac{d}{dt}\langle \hat{A}\rangle_\psi &= \frac{d}{dt}\int_{-\infty}^{+\infty}\psi^*\hat{A}\psi dx \\
&= \int_{-\infty}^{+\infty}\left(\frac{\partial\psi^*}{\partial t}\hat{A}\psi + \psi^*\frac{\partial\hat{A}}{\partial t}\psi + \psi^*\hat{A}\frac{\partial\psi}{\partial t}\right)dx \\
&= \frac{i}{\hbar}\int_{-\infty}^{+\infty}\left(\psi^*\hat{H}\hat{A}\psi - \psi^*\hat{A}\hat{H}\psi\right)dx + \langle\frac{\partial\hat{A}}{\partial t}\rangle_\psi \\
&= \frac{i}{\hbar}\langle[\hat{H},\hat{A}]\rangle_\psi + \langle\frac{\partial\hat{A}}{\partial t}\rangle_\psi \qquad \square
\end{aligned}
\tag{4.30}
$$

The Ehrenfest's theorem applied to the Hamiltonian operator itself yields a very interesting result. Indeed, if we compute explicitly the commutator of the Hamiltonian operator with $\hat{x}$, $\hat{p}$ and itself, we get (you can compute them as an exercise)

$$
\begin{aligned}
{[\hat{H},\hat{p}]} &= i\hbar\frac{dU(\hat{x})}{dx}, \\
{[\hat{H},\hat{x}]} &= \frac{-i\hbar}{m}\hat{p}, \\
{[\hat{H},\hat{H}]} &= 0.
\end{aligned}
\tag{4.31}
$$

Given that $\hat{H}$, $\hat{x}$ and $\hat{p}$ do not depend explicitly on time, the RHS of Eq. (4.29) vanishes in each case, giving

$$
\begin{aligned}
\frac{d}{dt}\langle \hat{p}\rangle_\psi &= -\langle\frac{dU}{dx}\rangle_\psi \\
\frac{d}{dt}\langle \hat{x}\rangle_\psi &= \frac{1}{m}\langle\hat{p}\rangle_\psi \\
\frac{d}{dt}\langle \hat{H}\rangle_\psi &= 0,
\end{aligned}
\tag{4.32}
$$

which are quantum versions of the classical laws $dp/dt = -dU/dx$, $p = mv$ and of the conservation of total energy respectively.

4.4 *The harmonic oscillator revisited*

[1] By considering commutation relations, we can give a much nicer and more illuminating derivation of the energy spectrum of the harmonic oscillator. This derivation forms part of the material for the Part II Principles of Quantum Mechanics course. It is non-examinable material, in the sense that if you are asked to derive the energy spectrum (without any method being stipulated) then the derivation given earlier is a perfectly adequate answer. However, the derivation below is simpler and slicker, and of course it also may be used in this context. Recall that the harmonic oscillator hamiltonian is

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2, \tag{4.33}$$

which can be written as

$$\hat{H} = \frac{1}{2m}(\hat{p} + im\omega\hat{x})(\hat{p} - im\omega\hat{x}) + \frac{i\omega}{2}[\hat{p}, \hat{x}]. \tag{4.34}$$

Substituting the canonical commutation relation for $\hat{x}$ and $\hat{p}$, we get

$$\hat{H} = \frac{1}{2m}(\hat{p} + im\omega\hat{x})(\hat{p} - im\omega\hat{x}) + \frac{\hbar\omega}{2}. \tag{4.35}$$

Define the operator (sometimes referred to as *ladder operator* or *raising and lowering operators*)

$$\hat{a} = \frac{1}{\sqrt{2m}}(\hat{p} - im\omega\hat{x}) \tag{4.36}$$

$$\hat{a}^\dagger = \frac{1}{\sqrt{2m}}(\hat{p} + im\omega\hat{x}), \tag{4.37}$$

where in computing the Hermitian conjugate we used the fact that $\hat{x}$ and $\hat{p}$ are both Hermitian. Note that instead $\hat{a}$ is not Hermitian, and thus it does not correspond to a physical observable, while $\hat{a}^\dagger\hat{a}$ is Hermitian. In terms of $\hat{a}$ and $\hat{a}^\dagger$, the Hamiltonian can be written as

$$\hat{H} = \hat{a}^\dagger\hat{a} + \frac{\hbar\omega}{2}. \tag{4.38}$$

You can easily verify that

$$\begin{aligned}
[\hat{a}, \hat{a}^\dagger] &= \hbar\omega\hat{I} \\
[\hat{H}, \hat{a}] &= -\hbar\omega\hat{a} \\
[\hat{H}, \hat{a}^\dagger] &= \hbar\omega\hat{a}^\dagger.
\end{aligned} \tag{4.39}$$

---

[1]Non-examinable section

Suppose now that χ is a harmonic oscillator eigenfunction of energy E, $\hat{H}\chi = E\chi$. We then have

$$
\begin{aligned}
\hat{H}\hat{a}\chi &= [\hat{H}, \hat{a}]\chi + \hat{a}\hat{H}\chi = (E - \hbar\omega)\hat{a}\chi \\
\hat{H}\hat{a}^\dagger\chi &= [\hat{H}, \hat{a}^\dagger]\chi + \hat{a}^\dagger\hat{H}\chi = (E + \hbar\omega)\hat{a}\chi
\end{aligned}
\tag{4.40}
$$

so that $\hat{a}\chi$ and $\hat{a}^\dagger\chi$ are eigenfunctions of energies $(E - \hbar\omega)$ and $(E + \hbar\omega)$ respectively. We can use this to prove by induction that $\hat{a}^n\chi$ and $(\hat{a}^\dagger)^n\chi$ are eigenfunctions of energy $(E - n\hbar\omega)$ and $(E + n\hbar\omega)$. For example, we know that $E_0 = E$. The proof by induction is complete by showing that, if $(\hat{a}^{n-1})\chi$ is the eigenfunction with energy $(E - (n-1)\hbar\omega)$ then $\hat{H}\hat{a}^n\chi$ is the eigenfunction with energy $(E - n\hbar\omega)$, which is straightforward given that

$$
\hat{H}\hat{a}^n\chi = \hat{a}\hat{H}(\hat{a}^{n-1})\chi = (E - n\hbar\omega)\hat{a}^n\chi.
\tag{4.41}
$$

The same proof can be achieved when considering $(\hat{a}^\dagger)^n\chi$. As a consequence, if it were true that $\hat{a}^n\chi \neq 0$ for all n, there would be eigenfunctions of arbitrarily low energy, and so there would be no ground state. However, given any physical wavefunction ψ, we have that

$$
\langle\hat{H}\rangle_\psi = \int_{-\infty}^{+\infty} \psi^*(x) \left(\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2 \right) \psi(x)\, dx \geq 0,
\tag{4.42}
$$

since both terms are non-negative. This argument can obviously be generalised to show that, if we have any potential $U(x)$ such that $U(x) \geq 0$ for all x, then $\langle\hat{H}\rangle_\psi \geq 0$ for all states ψ. As a result, there cannot be negative energy eigenfunctions for the harmonic potential. Thus there must be a lowest energy (i.e. a ground state) eigenfunction χ_0 such that

$$
0 = \hat{a}\chi_0 \Rightarrow -i\hbar\frac{d\chi_0}{dx} = im\omega x\chi_0,
\tag{4.43}
$$

and hence

$$
\chi_0(x) = C\exp\left(-\frac{m\omega x^2}{2\hbar}\right)
\tag{4.44}
$$

which is indeed the ground state wavefunction we obtained in Sect. 3.5, but this time we derived it much more simply. Now, since $\hat{H} = \hat{a}^\dagger\hat{a} + \hbar\omega/2$ and $\hat{A}\chi_0 = 0$ we have $\hat{H}\chi_0 = \hbar\omega\chi_0/2$, giving us the previously obtained value of $\hbar\omega/s$ for the ground state energy. The excited states can be obtained by applying the operator $\hat{a}^\dagger$ n times:

$$
(\hat{a}^\dagger)^n\chi_0 = \frac{C}{(2m)^{n/2}}(\hat{p} + im\omega\hat{x})^n \exp\left(-\frac{m\omega x^2}{2\hbar}\right)
\tag{4.45}
$$

$$
\Rightarrow \chi_n = \frac{C}{(2m)^{n/2}}\left(-i\hbar\frac{\partial}{\partial x} + im\omega\hat{x}\right)^n \exp\left(-\frac{m\omega x^2}{2\hbar}\right)
\tag{4.46}
$$

and we see immediately that their energies are $(n + 1/2)\hbar\omega$ as previously obtained by a less direct argument. Indeed

$$
\begin{aligned}
\hat{H}\chi_n &= \left(\hat{a}\hat{a}^\dagger + \frac{\hbar\omega}{2}\right)(\hat{a}^\dagger)^n\chi_0 \\
&= \hat{a}^\dagger\,[\hat{a},(\hat{a}^\dagger)^n]\chi_0 + \frac{\hbar\omega}{2}(\hat{a}^\dagger)^n\chi_0 \\
&= \left(n + \frac{1}{2}\right)\hbar\omega\chi_n,
\end{aligned}
\tag{4.47}
$$

where we used the (easily verifiable) identity

$$
[\hat{a},(\hat{a}^\dagger)^n] = n\hbar\omega\,(\hat{a}^\dagger)^{n-1}
\tag{4.48}
$$

We can similarly show that there cannot be eigenfunctions with energies taking values other than $(n+1/2)\hbar\omega$. If there were such eigenfunctions $\tilde{\chi}$, then $\hat{a}^m\tilde{\chi}$ cannot vanish for any m, since χ_0 is the unique wavefunction annihilated by $\hat{a}$. So there would be negative energy eigenfunctions, which contradicts the result shown above. With a little more thought, we can similarly also show that the eigenspaces must all be non-degenerate: *i.e.* there is (up to scalar multiplication) only one eigenfunction of each energy.

The derivation of the harmonic oscillator spectrum in this subsection illustrates an important general feature: symmetries or regularities in a quantum mechanical spectrum (such as the regular spacing of the harmonic oscillator energy levels) suggest the existence of a set of operators whose commutation relations define the symmetry or explain the regularity (in this case, the operators $\hat{H}$, $\hat{a}$ and $\hat{a}^\dagger$).

Chapter 5

The 3D Schrödinger equation

In this chapter, we will finally put everything that we have learned together to investigate realistic QM problems in our three-dimensional world, by focussing on spherically symmetric potential, in particular on the hydrogen atom, the capstone of our lectures.

5.1 3D Schrödinger equation for spherically symmetric potentials

We have already introduced the time-independent Schrödinger equation in three spatial dimensions,

$$-\frac{\hbar^2}{2m}\nabla^2\chi(\mathbf{x}) \ + U(\mathbf{x})\chi(\mathbf{x}) \ = \ E\chi(\mathbf{x}) \tag{5.1}$$

As we mentioned in Chap. 2, in Cartesian coordinates the Laplace operator is given by:

$$\nabla^2 \ = \ \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{5.2}$$

It is useful to write it down in spherical polar coordinates, see Fig. 5.1,

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

where,

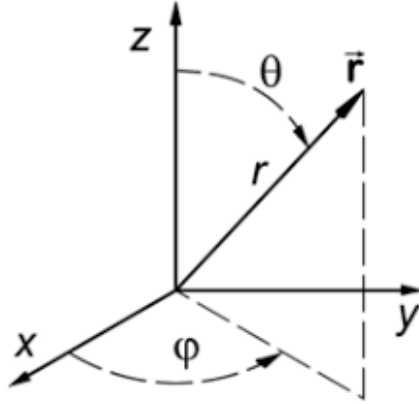$$0 \le r < \infty, \ \ 0 \le \theta \le \pi, \ \ 0 \le \phi < 2\pi.$$

Figure 5.1: Spherical polar coordinates.

In spherical polar coordinate, the Laplace operator becomes

$$\nabla^2 \;=\; \frac{1}{r}\frac{\partial^2}{\partial r^2}\,r \;+\; \frac{1}{r^2\sin^2(\theta)}\left[\sin(\theta)\frac{\partial}{\partial\theta}\,\sin(\theta)\frac{\partial}{\partial\theta} \;+\; \frac{\partial^2}{\partial\phi^2}\right], \tag{5.3}$$

as you have seen in part IA Vector Calculus. Also, recall that,

$$\int_{\mathbb{R}^3} dV \;=\; \int_0^{2\pi} d\phi \int_{-1}^{+1} d\left(\cos(\theta)\right) \int_0^{\infty} r^2\, dr$$

In this Section we only focus on spherically symmetric potential, for which the potential is only function of r

$$U(r,\theta,\phi) \;\equiv\; U(r).$$

These are both the simplest potentials to study and, happily, the kinds of potentials that are most useful for physics.

As a result, we may want to look for spherically symmetric stationary state

$$\chi(r,\theta,\phi) \;\equiv\; \chi(r)$$

for which,

$$\nabla^2\chi \;=\; \frac{1}{r}\frac{\partial^2}{\partial r^2}\,(r\chi) = \frac{1}{r}\frac{d^2}{dr^2}\,(r\chi) \tag{5.4}$$

and thus the time-independent Schrödinger equation becomes,

$$-\frac{\hbar^2}{2mr}\frac{d^2}{dr^2}\,(r\chi(r)) \;+\; U(r)\,\chi(r) \;=\; E\chi(r) \tag{5.5}$$
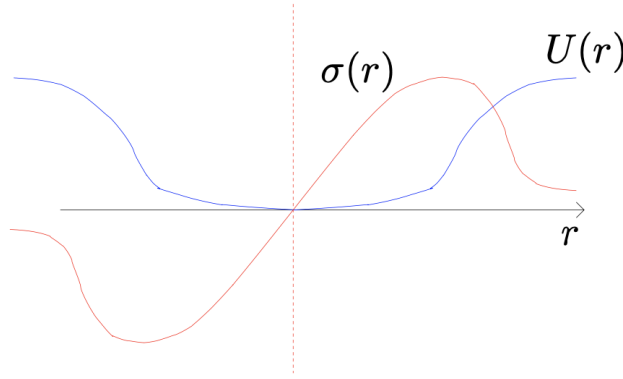
78

Figure 5.2: One-dimensional problem.

or

$$-\frac{\hbar^2}{2m}\left(\frac{d^2\chi(r)}{dr^2} \ + \ \frac{2}{r}\frac{d\chi(r)}{dr}\right) \ + \ U(r)\,\chi(r) \ = \ E\chi(r). \tag{5.6}$$

We will show later that the ground state (the lowest energy bound state, if there is one) of any 3D quantum system with spherically symmetric potential is itself spherically symmetric[1]. Hence we can always use the method that we outline in this section to obtain the ground state.

Because of the normalisability condition, the spacial part of the wavefunction $\chi(r)$ must be finite at $r = 0$ [2] and

$$\int_{\mathbb{R}^3} |\psi|^2\ dV \ < \ \infty \ \Rightarrow \ \int_0^\infty |\chi(r)|^2\ r^2\ dr \ < \ \infty$$

which requires that $\chi(r) \to 0$ sufficiently fast as $r \to \infty$.

It might be useful is to define $\sigma(r) = r\chi(r)$. Eq. (5.6) becomes,

$$-\frac{\hbar^2}{2m}\frac{d^2\sigma(r)}{dr^2} \ + \ U(r)\sigma(r) \ = \ E\sigma(r) \tag{5.7}$$

This is one-dimensional Schrödinger equation on the half-line $r \geq 0$. A trick is to solve Schrödinger equation on whole line $-\infty < r < +\infty$ by defining $U(r)$ also on the negative half life by the symmetry condition $U(-r) = U(r)$ (as in Fig. 5.2). The bound state

---

[1]Cf. the 1D result that the ground state of a symmetric potential always has even parity.
[2]This is an additional condition as compared to the solutions in 1D

wavefunctions on the whole real line must be either even or odd, but in this case the only acceptable solutions are the odd parity ones $\sigma^{(-)}(-r) = -\sigma^{(-)}(r)$ given that

$$\sigma(0) = 0 \qquad \text{and} \qquad \int_0^\infty |\sigma(r)|^2 \, dr \; < \; \infty$$

which yields a solution to the original problem because the wavefunction $\chi(r) = \sigma(r)/r$ finite at $r = 0$. This follows from $\sigma(0) = 0$ provided $\sigma'(0)$ is finite (use L' Hôpital's rule) On the other hand the normalisability condition,

$$\int_0^\infty |\chi(r)|^2 \, r^2 \, dr \; < \; \infty \tag{5.8}$$

follows for $\chi(r) = \sigma(r)/r$.

Now that we have proven that the only acceptable solutions for Eq. (5.7) are the odd-parity one, when we extend it on the whole real line, we can use this trick to solve a number of problems, for example the spherically-symmetric harmonic oscillator, or a spherically-symmetric square well. In the latter case,

$$U(r) \;=\; \begin{cases} 0 & \text{for} \quad r < a \\ U_0 & \text{for} \quad r > a \end{cases} \tag{5.9}$$

To find odd-parity bound states states of 3D square well, analogously to what we did in 1D, we define the constants,

$$k = \sqrt{\frac{2mE}{\hbar^2}} \;\geq\; 0, \qquad \bar{k} = \sqrt{\frac{2m(U_0 - E)}{\hbar^2}} \;\geq\; 0$$

and select solutions of TISE of form,

$$\sigma(r) \;=\; \begin{cases} A\sin(kr) & |r| \leq a \\ B\exp(-\bar{k}r) & r > a \end{cases} \tag{5.10}$$

Applying the boundary conditions, i.e. asking for the continuity of σ and σ' at $r = a$, we get

$$\begin{aligned} A\sin(ka) &= B\exp(-\bar{k}a) \\ \text{and} \qquad kA\cos(ka) &= -\bar{k}B\exp(-\bar{k}a) \end{aligned}$$

$$\Rightarrow \qquad -k\cot(ka) \;=\; \bar{k}$$

We can then define the rescaled variables,

$$\xi = ka, \qquad \eta = \bar{k}a, \qquad r_0 = \sqrt{\frac{2mU_0}{\hbar^2}}\, a \tag{5.11}$$

80
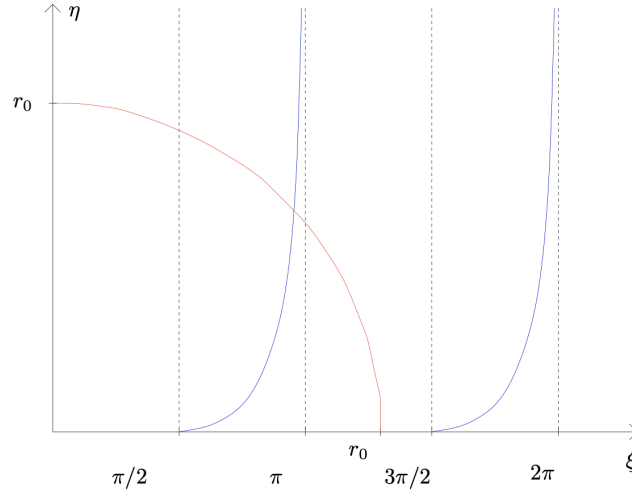
Figure 5.3: Graphical solution: odd-parity levels of spherically-symmetric potential well.

as we did in the 1D case. We end up with two equations relating unknowns ξ and η.

$$\begin{aligned} \xi^2 + \eta^2 &= r_0^2 \\ -\xi\cot(\xi) &= \eta \end{aligned}$$

As in 1D, the finite number of bound states is determined by number of intersections between the two curves in the graphical solution. Comparing Fig. 5.3 to Fig. 3.4 in Chapt. 3, we see that, contrarily to what happens in 1D, in 3D there are no bound states if $r_0 < \pi/2$ or equivalently if

$$U_0 \ \leq \ \frac{\pi^2 \hbar^2}{8ma^2}$$

unlike one-dimensional case where we always find at least one bound state.

5.2   The angular momentum

So far, we have focussed on spherically symmetric eigenfunctions $\chi(r)$ of the Hamiltonian. In this section, we introduce the angular momentum operator, which we need in order to obtain the full eigenstate spectrum of the hydrogen atom. In Classical Mechanics, the angular momentum is a vector defined as

$$\mathbf{L} \ = \ \mathbf{x} \times \mathbf{p}.$$

As you might recollect from Part IA Dynamics and Relativity, the vector $\mathbf{L}$ is a conserved quantity for systems with a spherically symmetric potentials $U(r, \theta, \phi) \equiv U(r)$.

Definition: In Quantum Mechanics, the orbital angular momentum is an observable which corresponds to the operator,

$$
\begin{aligned}
\hat{\mathbf{L}} &= \hat{\mathbf{x}} \times \hat{\mathbf{p}} \\
&= -i\hbar \mathbf{x} \times \nabla
\end{aligned}
$$

In index notation for Cartesian coordinates $\mathbf{x} = (x_1, x_2, x_3)$

$$
\hat{L}_i = \varepsilon_{ijk} \hat{x}_j \hat{p}_k = -i\hbar \, \varepsilon_{ijk} x_j \frac{\partial}{\partial x_k}
$$

where ε_{ijk} is the Levi-Civita alternating tensor. Explicitly,

$$
\hat{\mathbf{L}} = -i\hbar \left(x_2 \frac{\partial}{\partial x_3} - x_3 \frac{\partial}{\partial x_2}, \;\; x_3 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_3}, \;\; x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1} \right)
$$

Properties of the angular momentum operator:

- The operators $\hat{L}_i$ are Hermitian, as all operators associated to physical observables.

- The different components of the angular momentum operator do not commute with each other: $[\hat{L}_i, \hat{L}_j] \neq 0$ for $i \neq j$. Thus, different components of angular momentum cannot be measured simultaneously.
  We can **check** this explicitly by computing the commutator,

  $$
  \begin{aligned}
  &[\hat{L}_1, \hat{L}_2] \, f(x_1, x_2, x_3) = \\
  &-\hbar^2 \left[\left(x_2 \frac{\partial}{\partial x_3} - x_3 \frac{\partial}{\partial x_2} \right) \left(x_3 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_3} \right) - \left(x_3 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_3} \right) \left(x_2 \frac{\partial}{\partial x_3} - x_3 \frac{\partial}{\partial x_2} \right) \right] f(x_1, x_2, x_3)
  \end{aligned}
  $$

  In the above equation many term cancels leaving,

  $$
  \begin{aligned}
  [\hat{L}_1, \hat{L}_2] \, f &= -\hbar^2 \left(x_2 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_2} \right) f(x_1, x_2, x_3) \\
  &= +i\hbar \, \hat{L}_3 f(x_1, x_2, x_3) \qquad\qquad\qquad (5.12)
  \end{aligned}
  $$

  A similar calculation for the other components confirms the commutation relations,

  $$
  [\hat{L}_2, \hat{L}_3] = i\hbar \hat{L}_1 \qquad \text{and} \qquad [\hat{L}_1, \hat{L}_3] = -i\hbar \hat{L}_2 \qquad\qquad (5.13)
  $$

- $[\hat{L}_i, \hat{L}_j] = i\hbar \varepsilon_{ijk} \hat{L}_k$, which combines the three independent commutation relations that we found above, using index notation.

In Classical Mechanics the magnitude of the angular momentum is $L = |\mathbf{L}|$. Thus,

$$L^2 \;=\; L_1^2 \;+\; L_2^2 \;+\; L_3^2$$

Definition: In Quantum Mechanics we define the **total angular momentum operator**,

$$\hat{L}^2 \;=\; \hat{L}_1^2 \;+\; \hat{L}_2^2 \;+\; \hat{L}_3^2$$

Properties of the total angular momentum operator

- The total angular momentum $\hat{L}^2$ commutes with each of the components of angular momentum $\hat{L}_i$, $i = 1, 2, 3$.

  We can **check** this explicitly by taking the commutator with the component $i = 1$. Now evaluate the commutators:

$$[\hat{L}_1, \hat{L}^2] \;=\; [\hat{L}_1, \hat{L}_1^2 + \hat{L}_2^2 + \hat{L}_3^2]. \tag{5.14}$$

Clearly

$$[\hat{L}_1, \hat{L}_1^2] \;=\; 0 \tag{5.15}$$

while

$$[\hat{L}_1, \hat{L}_2^2] \;=\; [\hat{L}_1, \hat{L}_2]\hat{L}_2 + \hat{L}_2[\hat{L}_1, \hat{L}_2]$$

Then using (5.12) we obtain,

$$[\hat{L}_1, \hat{L}_2^2] \;=\; i\hbar \left[\hat{L}_3\hat{L}_2 + \hat{L}_2\hat{L}_3 \right] \tag{5.16}$$

and,

$$\begin{aligned}
[\hat{L}_1, \hat{L}_3^2] \;&=\; [\hat{L}_1, \hat{L}_3]\hat{L}_3 + \hat{L}_3[\hat{L}_1, \hat{L}_3] \\
&=\; -i\hbar \left[\hat{L}_3\hat{L}_2 + \hat{L}_2\hat{L}_3 \right]
\end{aligned} \tag{5.17}$$

Finally adding equations (5.14), (5.15), (5.16) and (5.17) we obtain,

$$[\hat{L}_1, \hat{L}^2] = [\hat{L}_1, \hat{L}_1^2] + [\hat{L}_1, \hat{L}_2^2] + [\hat{L}_1, \hat{L}_3^2] \;=\; 0$$

An identical calculation of $[\hat{L}_2, \hat{L}^2]$ and $[\hat{L}_3, \hat{L}^2]$ confirms that,

$$[\hat{L}_i, \hat{L}^2] \;=\; 0 \tag{5.18}$$

for $i = 1, 2, 3$ $\square$.

- The total angular momentum operator commutes with an Hamiltonian featuring a spherically symmetric potential. This means that the angular momentum operators care nothing for how the wavefunctions depend on the radial coordinate r. The angular momentum of a state is, perhaps unsurprisingly, encoded in how the wavefunction varies in the angular directions θ and Φ. To **check** the above property, we have to prove first the following commutation relations,

$$
\begin{aligned}
[\hat{L}_i, \hat{x}_j] &= \varepsilon_{ikr}[\hat{x}_k \hat{p}_r, \hat{x}_j] \\
&= \varepsilon_{ikr}\hat{x}_k[\hat{p}_r, \hat{x}_j] = \varepsilon_{ikr}\hat{x}_a(-i\hbar \delta_{rj}\hat{I}) \\
&= i\hbar\, \varepsilon_{ijk}\hat{x}_k,
\end{aligned}
\tag{5.19}
$$

where in the first line we substituted the definition of the angular momentum operator using the index notation and in the second we used the canonical commutation relation $[\hat{x}_i, \hat{p}_j] = i\hbar \delta_{ij}\hat{I}$. Analogously you can verify that

$$
[\hat{L}_i, \hat{p}_j] = i\hbar\, \varepsilon_{ijk}\hat{p}_k
\tag{5.20}
$$

From these obtain,

$$
[\hat{L}_i, \hat{x}_1^2 + \hat{x}_2^2 + \hat{x}_3^2] = 0
$$

$$
[\hat{L}_i, \hat{p}_1^2 + \hat{p}_2^2 + \hat{p}_3^2] = 0
$$

The Hamiltonian for a particle on mass m moving in a spherically symmetric potential has the form,

$$
\begin{aligned}
\hat{H} &= -\frac{\hbar^2}{2m}\nabla^2 + U(r) \\
&= \frac{|\hat{\mathbf{p}}|^2}{2m} + U(\hat{r})
\end{aligned}
$$

Here $\hat{r}$ is the operator which acts on functions $f(\mathbf{x})$ as $r\hat{\mathbb{I}}$ where r is the radial coordinate and $\hat{\mathbb{I}}$ is the unit operator. Using the above commutation relations show that $\hat{H}$ commutes with $\hat{L}_i$ for $i = 1, 2, 3$ and therefore also with $\hat{L}^2$:

$$
[\hat{H}, \hat{L}_i] = [\hat{H}, \hat{L}^2] = 0
\tag{5.21}
$$

The properties that we have listed so far give us an important physics insight. In particular, the commutation relations (5.18) and (5.21) imply that $\hat{H}$, $\hat{L}^2$ and any one of the three operators $\hat{L}_i$, $i = 1, 2, 3$ form a set of three mutually commuting operators. We must choose only one of the $\hat{L}_i$ because they do not commute with each other (5.13,5.13). By convention we choose $\hat{L}_3$. Labelling the Cartesian coordinates in the usual way as $x_1 = x$,$x_2 = y$,

$x_3 = z$, we also denote this operator as $\hat{L}_z$ or the "z-component of angular momentum". Thus we choose a set of mutually commuting operators,

$$\left\{ \ \hat{H} \ , \ \hat{L}^2, \ \hat{L}_3 \ \right\} \tag{5.22}$$

It is important to outline the following facts:

- As the operators commute we can find simultaneous eigenstates of all three (See Theorem 4.4).

- The corresponding eigenvalues are the observables energy, total angular momentum and the z-component of angular momentum.

- The set (5.22) is maximal, in other words, we cannot construct another independent operator (other than the unit operator) which commutes with each of $\hat{H}$, $\hat{L}^2$ and $\hat{L}_3$.

We can now conclude this section by finding the simultaneous eigenfunctions of the total angular momentum and the z component of the angular momentum operators. In spherical polar coordinates we have (see AppendixA.6),

$$\hat{L}^2 \ = \ -\frac{\hbar^2}{\sin^2(\theta)} \left[\sin(\theta)\frac{\partial}{\partial\theta} \ \left(\sin(\theta)\frac{\partial}{\partial\theta} \right) \ + \ \frac{\partial^2}{\partial\phi^2} \right] \tag{5.23}$$

$$\hat{L}_3 \ = \ -i\hbar\frac{\partial}{\partial\phi} \tag{5.24}$$

Considering the eigenvalue equation for $\hat{L}_3$, we look for simultaneous eigenfunctions of $\hat{L}^2$ and $\hat{L}_3$ of the form $y(\theta)\exp(im\phi)$,

$$\hat{L}_3 \exp(im\phi) \ = \ \hbar m \exp(im\phi)$$

But wavefunctions must be single-valued functions on $\mathbb{R}^3$ and should therefore be invariant under $\phi \to \phi + 2\pi$. The function $\exp(im\phi)$ is invariant provided,

$$\exp(2\pi im) \ = \ 1 \qquad \Rightarrow \qquad m \in \mathbb{Z}$$

Thus the eigenvalues of $\hat{L}_3$ have the form $\hbar m$ for integer m. Equivalently, the z-component of angular momentum is quantised in integer multiples of $\hbar$. This agrees with Bohr's quantisation condition.

Similarly we must have,

$$\hat{L}^2 \ y(\theta) \ \exp(im\phi) \ = \ \lambda \ y(\theta) \ \exp(im\phi)$$

for some eigenvalue λ. Using the explicit form (5.23) for $\hat{L}^2$ we find that $y(\theta)$ must obey the equation,

$$\frac{1}{\sin(\theta)}\frac{\partial}{\partial\theta} \ \left(\sin(\theta)\frac{\partial y(\theta)}{\partial\theta} \right) \ - \ \frac{m^2}{\sin^2(\theta)} \ y(\theta) \ = \ -\frac{\lambda}{\hbar^2} \ Y(\theta) \tag{5.25}$$

This is the *associated Legendre equation*. The non-singular solutions are known as *associated Legendre functions*,

$$
\begin{aligned}
y(\theta) &= P_{l,m}\left(\cos(\theta)\right) \\
&= (\sin(\theta))^{|m|}\ \frac{d^{|m|}}{d\left(\cos(\theta)\right)^{|m|}}\ P_l\left(\cos(\theta)\right) \quad\quad (5.26)
\end{aligned}
$$

where $P_l(\cos(\theta))$ are (ordinary) Legendre polynomials (see IB Methods). Without proof, the expression (5.26) solves equation (5.25) with eigenvalue,

$$
\lambda = l(l+1)\ \hbar^2 \quad\quad \text{with} \quad l = 0,1,2\ldots
$$

There is also a further constraint on the integers l and m which reads,

$$
-l \le m \le +l
$$

The simultaneous eigenfunctions of $\hat{L}^2$ and $\hat{L}_3$ are therefore labelled by two integers $l > 0$ with $-m \le l \le +m$ and take the form,

$$
Y_{l,m}(\theta,\phi) = P_{l,m}\left(\cos(\theta)\right)\ \exp(im\phi)
$$

They obey,

$$
\begin{aligned}
\hat{L}^2 Y_{l,m}(\theta,\phi) &= l(l+1)\hbar^2 Y_{l,m}(\theta,\phi) \\
\hat{L}_3 Y_{l,m}(\theta,\phi) &= m\hbar Y_{l,m}(\theta,\phi)
\end{aligned}
$$

The functions $Y_{l,m}$ are known as *Spherical Harmonics* The indices l and m are called quantum numbers, in particular the integer l is called the total angular momentum quantum number while the integer m is called the azimuthal quantum number. Finally the constraint $-l \le m \le +l$ is the quantum version of the classical inequality,

$$
-|\mathbf{L}| \ \le \ L_z \ \le \ +|\mathbf{L}|
$$

which follows because $L_z = |\mathbf{L}|\cos(\theta)$.

To conclude, we give the explicit functional form of some spherical harmonics:

$$
\begin{aligned}
Y_{0,0}(\theta, \phi) &= \sqrt{\frac{1}{4\pi}} \\
Y_{1,0}(\theta, \phi) &= \sqrt{\frac{3}{4\pi}} \cos(\theta) \\
Y_{1,\pm 1}(\theta, \phi) &= \mp\sqrt{\frac{3}{8\pi}} \sin(\theta) \exp(-i\phi) \\
Y_{2,0}(\theta, \phi) &= \sqrt{\frac{5}{16\pi}} \left(3\cos^2(\theta) - 1\right) \\
Y_{2,\pm 1}(\theta, \phi) &= \mp\sqrt{\frac{15}{8\pi}} \sin(\theta) \cos(\theta) \exp(\mp i\phi) \\
Y_{2,\pm 2}(\theta, \phi) &= \sqrt{\frac{15}{32\pi}} \sin^2(\theta) \exp(\mp 2i\phi),
\end{aligned}
$$

where the normalisation constants are determined by the normalisation condition

$$
\int_0^{2\pi} d\Phi \int_{-1}^{+1} d\cos\theta \, Y_{l,m}^*(\theta, \phi) \, Y_{l,m}(\theta, \phi) \; = \; 1. \tag{5.27}
$$

One can also verify explicitly the orthogonality of spherical harmonics, i. e.

$$
\int_0^{2\pi} d\Phi \int_{-1}^{+1} d\cos\theta \, Y_{l',m'}^*(\theta, \phi) \, Y_{l,m}(\theta, \phi) \; = \; \delta_{ll'}\delta_{mm'}. \tag{5.28}
$$

5.3   The Hydrogen Atom

The hydrogen (H) atom consists of a single proton p^+ and an electron e^-. As before treat the proton as stationary at the origin of spherical polar coordinates, i.e. we model the hydrogen atom by treating the proton as infinitely massive and at rest at the origin. The Coulomb attraction,

$$
F(r) = -\frac{\partial U(r)}{\partial r} \;\; = \;\; -\frac{e^2}{4\pi\varepsilon_0 r^2}
$$

corresponds to a potential

$$
U(r) \;\; = \;\; -\frac{e^2}{4\pi\varepsilon_0 r}, \tag{5.29}
$$

which is sketched in Fig. 5.4.

We notice that the potential is infinitely deep. The energy is defined so that the electron is at rest at $r = \infty$, where $E = 0$, and that for bound states the energy is negative $E < 0$.

The Laplacian operator in spherical polar coordinates was given in Eq. (5.3). Using (5.23) we can write it as

$$-\hbar^2 \nabla^2 \;=\; -\frac{\hbar^2}{r}\frac{\partial^2}{\partial r^2}\, r \;+\; \frac{\hat{L}^2}{r^2}$$

which gives,

$$\hat{H} \;=\; -\frac{\hbar^2}{2mr}\frac{\partial^2}{\partial r^2}\, r \;+\; \frac{\hat{L}^2}{2mr^2} \;+\; U(\mathbf{x})$$

In particular, for the Hydrogen atom the time-independent Schrödinger equation becomes,

$$\hat{H}\chi = -\frac{\hbar^2}{2m}\left(\frac{d^2\chi}{dr^2} + \frac{2}{r}\frac{d\chi}{dr}\right) + \frac{\hat{L}^2}{2mr^2}\chi - \frac{e^2}{4\pi\varepsilon_0 r}\,\chi \;=\; E\chi, \qquad (5.30)$$

where χ are functions of (r, θ, Φ). Because of what we discussed at the end of Sect. 5.3, we now look for a simultaneous eigenstate of,

$$\left\{\; \hat{H}\;,\; \hat{L}^2,\; \hat{L}_3 \;\right\}$$

by setting,

$$\chi(r, \theta, \phi) \;=\; g(r)\, Y_{l,m}(\theta, \phi) \qquad (5.31)$$

where $Y_{l,m}$ is a spherical harmonic.

5.3.1   The radial wavefunction ($l = 0$)

We first focus on wavefunctions with spherical symmetry, i.e. $l = 0$, that resemble the Bohr model in the classical limit (although with some important differences). In this case $Y_{0,0}(\theta, \phi) = 1/\sqrt{4\pi}$, so we can substitute

$$\chi(r, \theta, \phi) \;=\; \chi(r)/\sqrt{4\pi}.$$

These are not the only eigenfunctions of the Hamiltonian, as in the next section, we will encounter eigenfunctions that also depend on (θ, Φ). The spherically symmetric eigenfunctions $\chi(r)$ obey the Schrödinger equation (5.6) with the Coulomb potential (5.29)

$$-\frac{\hbar^2}{2m_e}\left(\frac{d^2\chi(r)}{dr^2} + \frac{2}{r}\frac{d\chi(r)}{dr}\right) - \frac{e^2}{4\pi\varepsilon_0 r}\,\chi(r) \;=\; E\chi(r),$$

where m_e is the mass of the electron. To simplify this equation, we define the rescaled variables,

$$\nu^2 = -\frac{2mE}{\hbar^2} \; > \; 0, \qquad \beta = \frac{e^2 m_e}{2\pi\varepsilon_0 \hbar^2}$$

in terms of which Schrödinger equation becomes,

$$\frac{d^2\chi(r)}{dr^2} \; + \; \frac{2}{r}\frac{d\chi(r)}{dr} \; + \; \left(\frac{\beta}{r} - \nu^2\right)\chi(r) \;\; = \;\; 0. \tag{5.32}$$

We notice that

i) The large-r asymptotic behaviour of the eigenfunctions is determined by the first and last terms in (5.32). In the limit $r \to \infty$ we have,

$$\frac{d^2\chi}{dr^2} - \nu^2\chi \simeq 0$$

which implies that the solutions of (5.32) have behaviour,

$$\chi(r) \;\; \sim \;\; \exp(\pm\nu r) \qquad\qquad \text{as} \;\; r \to \infty. \tag{5.33}$$

We must choose the solution with the - sign, i.e. an exponentially decaying solution, in order to have normalisable eigenfunctions.

ii) The eigenfunctions should be finite at $r = 0$. This will guide us in the solution of the problem.

As in the analysis of the harmonic oscillator, it is convenient to separate out the exponential dependence of the eigenfunction and look for a solution of the form,

$$\chi(r) \;\; = \;\; f(r)\exp(-\nu r). \tag{5.34}$$

If we plug it into the Schrödinger equation (5.32), it becomes,

$$\frac{d^2 f}{dr^2} \; + \; \frac{2}{r}(1 - \nu r)\frac{df}{dr} \; + \; \frac{1}{r}(\beta - 2\nu)f \;\; = \;\; 0, \tag{5.35}$$

where we omitted the explicit dependence of f on r. Eq. (5.35) is a homogeneous, linear ODE with a regular singular point at $r = 0$. We now apply the standard method and look for a solution in the form of a power series around $r = 0$,

$$f(r) \;\; = \;\; r^c \sum_{n=0}^{\infty} a_n r^n. \tag{5.36}$$

89

By substituting the series (5.36) for $f(r)$ in (5.35) we get

$$\sum_{n=0}^{\infty} \left[a_n(c+n)(c+n-1)r^{c+n-2} + \frac{2}{r}(1-\nu r)a_n(c+n)r^{c+n-1} + (\beta - 2\nu)a_n r^{c+n-1} \right] = 0$$

The lowest power of r which occurs on the LHS is $a_0 r^{c-2}$ with coefficient $c(c-1) + 2c = c(c+1)$. Equating this to zero yields the indicial equation

$$c(c+1) \;\; = \;\; 0$$

with roots $c = 0$ and $c = -1$. However, the root c=-1 would imply $\chi(r) \sim 1/r$ near $r = 0$. This yields a singular eigenfunction which violates the boundary condition at the origin. Thus we choose root $c = 0$ and our series solution simplifies to

$$f(r) \;\; = \;\; \sum_{n=0}^{\infty} a_n r^n. \tag{5.37}$$

If we now collect all terms of order r^{n-2} on the LHS of (5.35) and equate them to zero to get,

$$n(n-1)a_n + 2na_n - 2\nu(n-1)a_{n-1} + (\beta - 2\nu)a_{n-1} \;\; = \;\; 0 \tag{5.38}$$

or more simply,

$$a_n \;\; = \;\; \frac{(2\nu n - \beta)}{n(n+1)} \, a_{n-1} \tag{5.39}$$

The above recurrence relation determines all the coefficients a_n in the series (5.37) in terms of the first coefficent a_0. As in our analysis of the harmonic oscillator, there are two possibilities,

- The series (5.37) terminates. In other words $\exists \; N > 0$ such that $a_n = 0 \; \forall \; n \geq N$.

- The series (5.37) does not terminate. In other words $\nexists \; N > 0$ such that $a_n = 0 \; \forall \; n \geq N$.

Proposition: If the series in Eq. (5.37) does not terminate, then the function $\chi(r) = f(r)e^{-\nu r}$ is not normalisable.

   Proof: The asymptotic (large-n) behaviour of the coefficients a_n is given by

$$\frac{a_n}{a_{n-1}} \;\; \rightarrow \;\; \frac{2\nu}{n} \qquad \text{as} \;\; n \rightarrow \infty. \tag{5.40}$$

We can now compare this with the power series for the function,

$$g(r) = \exp(+2\nu r) \;\; = \;\; \sum_{n=0}^{\infty} b_n r^n \qquad \text{with} \;\; b_n = \frac{(2\nu)^n}{n!}$$

whose coefficients obey,

$$\frac{b_n}{b_{n-1}} \;=\; \frac{(2\nu)^n}{(2\nu)^{n-1}} \frac{(n-1)!}{n!} \;=\; \frac{2\nu}{n}$$

We deduce that (5.40) is consistent with the asymptotic behaviour

$$f(r) \sim g(r) = \exp(+2\nu r) \;\;\Rightarrow\;\; \chi(r) = f(r)\exp(-\nu r) \sim \exp(+\nu r)$$

as $r \to \infty$ which is consistent with the expected exponential growth (5.33) of generic solutions of (5.32). This corresponds to non-normalisable eigenfunctions which we reject. $\square$

As in the solution for the harmonic oscillator in Chapt. 3, if the series does not terminate, the Hamiltonian does not have normalisable eigenfunctions. To give a normalisable eigenfunction therefore, the series (5.37) must terminate. There must be an integer $N > 0$ such that $a_N = 0$ with $a_{N-1} \neq 0$. From the recurrence relation (5.39) we can see that this happens if and only if,

$$2\nu N - \beta = 0 \qquad\qquad \Rightarrow \qquad\qquad \nu = \frac{\beta}{2N}.$$

Recalling the definitions $\nu^2 = -\frac{2mE}{\hbar^2}$ and $\beta = \frac{e^2 m_e}{2\pi\varepsilon_0\hbar^2}$ the equation above yields the spectrum of energy levels,

$$E = E_N \;\;=\;\; -\frac{e^4 m_e}{32\pi^2\epsilon_0^2\hbar^2}\;\frac{1}{N^2} \tag{5.41}$$

for $N = 1, 2, \ldots$. The resulting spectrum is identical to that of the Bohr atom. Thus, the Schrödinger equation predicts the same set of spectral lines for Hydrogen which are in good agreement with experiment, although - as we will see in the next sections - the degeneracies (ie number of eigenstates with the same energy) are still wrong. An important difference is that Bohr's spectrum was based on quantisation of angular momentum corresponding to $L = N\hbar$. In contrast the eigenfunctions we have constructed are spherically symmetric and therefore have zero angular momentum. We will see in the next two sections that, including the eigenfunctions that are not spherically symmetric, there are more eigenfunctions corresponding to the eigenvalues in (**??**), so the degeneracy of the excited levels is larger than one.

To conclude this section, we now want to determine the eigenfunctions corresponding to the eigenvalues of Eq. (5.41). We do it by setting $\nu = \beta/2N$, the recurrence relation (5.39), which becomes,

$$\begin{aligned}
\frac{a_n}{a_{n-1}} \;&=\; \frac{2\nu n - \beta}{n(n+1)} \\
&=\; -2\nu\left(\frac{N-n}{n(n+1)}\right)
\end{aligned}$$

This formula can be used to give expliciot results for the first few levels (see Figure 5.5),

$$
\begin{aligned}
\chi_1(r) &= \exp(-\nu r) \\
\chi_2(r) &= (1 - \nu r) exp(-\nu r) \\
\chi_3(r) &= \left(1 - 2\nu r + \frac{2}{3}(\nu r)^2\right) \exp(-\nu r).
\end{aligned}
$$

The eigenfunction for the N-th level can be written as,

$$
\chi_N(r) = L_N(\nu r) \ \exp(-\nu r)
$$

where L_N is a polynomial of order $N - 1$ known as the N-th **Laguerre polynomial**. The wavefunction $\chi_N(r)$ thus has $N - 1$ nodes or zeros. Finally, we want to normalise the eigenfunctions to 1, by defining $\tilde{\chi}_1(r) = A_1\chi_1(r) = A_1 \exp(-\nu r)$. The constant A_1 is fixed by normalisation condition,

$$
\int_{\mathbb{R}^3} |\tilde{\chi}_1(r)|^2 \ dV = 1
$$

Evaluating the integral we find,

$$
|A_1|^2 \int_0^{2\pi} d\phi \int_{-1}^{+1} d(\cos\theta) \int_0^{\infty} dr \ r^2 \exp(-2\nu r) = 1
$$

Thus $|A_1|^2 = 1/\mathcal{I}_2$ where,

$$
\mathcal{I}_2 = 4\pi \int_0^{\infty} r^2 \exp(-2\nu r) = \frac{\pi}{\nu^3}
$$

Finally we can choose A_1 to be real by choosing the complex phase to be zero and

$$
A_1 = \frac{1}{\sqrt{\pi}} \left(\frac{e^2 m_e}{4\pi\varepsilon_0\hbar^2 N}\right)^{\frac{3}{2}} \tag{5.42}
$$

<u>Exercise</u> Prove that in the spherically-symmetric ground state,

$$
\langle r \rangle = \frac{3}{2} \ r_1
$$

where $r_1 = 2/\beta = 4\pi\varepsilon_0\hbar^2/m_e e^2$ is the Bohr radius as defined in Sect. 1.

5.3.2   The full wavefunction

We now look for a simultaneous eigenstate of,

$$\left\{ \ \hat{H} \ , \ \hat{L}^2, \ \hat{L}_3 \ \right\}$$

for generic values of (l, m), so we set

$$\chi(r, \theta, \phi) \ = \ g(r) \, Y_{l,m}(\theta, \phi) \tag{5.43}$$

where $Y_{l,m}$ is a spherical harmonic. In particular, as above, we have,

$$\hat{L}^2 Y_{l,m}(\theta, \phi) \ = \ l(l+1)\hbar^2 Y_{l,m}(\theta, \phi).$$

Substituting (5.43) into (5.30) we obtain a second order linear homogeneous ODE for $g(r)$,

$$-\frac{\hbar^2}{2m}\left(\frac{d^2g(r)}{dr^2} \ + \ \frac{2}{r}\frac{dg(r)}{dr}\right) \ + \ \frac{\hbar^2 l(l+1)}{2mr^2}g(r) \ - \ \frac{e^2}{4\pi\varepsilon_0 r} \, g(r) \ = \ Eg(r) \tag{5.44}$$

As we did in Sect. 5.2, we define the rescaled variables,

$$\nu^2 = -\frac{2mE}{\hbar^2} \ > \ 0, \qquad \beta = \frac{e^2 m_e}{2\pi\varepsilon_0 \hbar^2}$$

in terms of which the Schrödinger equation becomes,

$$\frac{d^2g}{dr^2} \ + \ \frac{2}{r}\frac{dg}{dr} \ - \ \frac{l(l+1)}{r^2}g \ + \ \frac{\beta}{r}g \ - \nu^2 g \ = \ 0, \tag{5.45}$$

where we omitted the dependence of g on r. The analysis proceeds exactly as for the spherically symmetric case that we have worked on in Sect. 5.2. In particular, the large-r asymptotic behaviour of $g(r)$ is determined by the first and last terms in (5.45). In the limit $r \to \infty$ we have,

$$g(r) \ \sim \ \exp(-\nu r) \qquad\qquad \text{as} \ \ r \to \infty \tag{5.46}$$

where we choose an exponentially decaying solution for a normalisable eigenfunction. As before, it is convenient to separate out the exponential dependence of the eigenfunctions and look for a solution of the form,

$$g(r) \ = \ f(r)\exp(-\nu r)$$

The Schrödinger equation (5.45) now becomes,

$$\frac{d^2 f}{dr^2} \ + \ \frac{2}{r}(1 - \nu r)\frac{df}{dr} \ - \ \frac{l(l+1)}{r^2}f \ + \ \frac{1}{r}(\beta - 2\nu)f \ = \ 0 \tag{5.47}$$

Equation (5.47) is a homogeneous, linear ODE with a regular singular point at $r = 0$. Apply standard method and look for a solution in the form of a power series around $r = 0$,

$$f(r) \quad = \quad r^\sigma \sum_{n=0}^{\infty} a_n r^n \tag{5.48}$$

Substitute series (5.48) for $f(r)$ in (5.47). We observe that the lowest power of r which occurs on the LHS is $a_0 r^{\sigma-2}$ with coefficient $\sigma(\sigma-1) + 2\sigma - l(l+1) = \sigma(\sigma+1)) - l(l+1)$. Equating this to zero yields the indicial equation,

$$\sigma(\sigma+1) \quad = \quad l(l+1)$$

with roots $\sigma = l$ and $\sigma = -l - 1$. The root $\sigma = -l - 1$ yields $g(r) \sim 1/r^{l+1}$ near $r = 0$, which is singular at the origin. Thus we choose root $\sigma = l$ and our series solution simplifies to,

$$f(r) \quad = \quad r^l \sum_{n=0}^{\infty} a_n r^n \tag{5.49}$$

We now collect all terms of order r^{l+n-2} on the LHS of (5.47) and equate them to zero to get,

$$(n+l)(n+l-1)a_n + 2(n+l)a_n - l(l+1)a_n \quad -$$
$$2\nu(n+l-1)a_{n-1} + (\beta - 2\nu)a_{n-1} \quad = \quad 0$$

or more simply,

$$a_n \quad = \quad \frac{(2\nu(n+l) - \beta)}{n(n+2l+1)} \, a_{n-1} \tag{5.50}$$

This recurrence relation determines all the coefficients a_n in the series (5.49) in terms of the first coefficient a_0. As above, there are two possibilities,

- The series (5.49) terminates. In other words $\exists \, n_{\max} > 0$ such that $a_n = 0 \, \forall \, n \geq n_{\max}$.

- The series (5.49) does not terminate. In other words $\nexists \, n_{\max} > 0$ such that $a_n = 0 \, \forall \, n \geq n_{\max}$.

As before, the second possibility does not yield normalisable wave functions and the proof is exactly the same as the one provided in Lemma 5.1. To give a normalisable wave-function therefore, the series (5.49) must therefore terminate. There must be an integer $n_{\max} > 0$ such that $a_{n_{\max}} = 0$ with $a_{n_{\max}-1} \neq 0$. From the recurrence relation (5.50) we can see that this happens if and only if,

$$2\nu N - \beta = 0 \qquad \Rightarrow \qquad \nu = \frac{\beta}{2N}$$

for an integer $N = n_{\max} + l > l$. Recalling the definitions,

$$\nu^2 = -\frac{2mE}{\hbar^2} > 0, \qquad \beta = \frac{e^2 m_e}{2\pi\varepsilon_0\hbar^2}$$

This yields the spectrum of energy levels,

$$E = E_N = -\frac{e^4 m_e}{32\pi^2\varepsilon_0^2\hbar^2}\frac{1}{N^2} \qquad (5.51)$$

for $N = 1, 2, \ldots$. The resulting spectrum of energy eigenvalues is identical to that of the Bohr atom and to our analysis of the spherically-symmetric wavefunctions.

The new feature is that there is a large **degeneracy** at each level. To see this note that the energy E_N given in (5.51) only depends on N and not on the angular momentum quantum numbers,

$$0 \le l \le N - 1 \qquad\qquad -l \le m \le +l$$

Thus the total degeneracy at each level is,

$$\begin{aligned}
D(N) &= \sum_{l=0}^{N-1}\sum_{-l}^{+l} 1 \\
&= \sum_{l=0}^{N-1}(2l+1) \\
&= 2\left(\frac{1}{2}N(N-1)\right) + N = N^2
\end{aligned}$$

We now put everything together to find the full spectrum of eigenstates of the hydrogen atom:

$$\chi_{N,l,m}(r,\theta,\phi) = \xi^l\, L_{N,l}(\xi)\,\exp(-\xi)\, Y_{l,m}(\theta,\phi)$$

where,

$$\xi = \frac{\beta r}{2N} = \frac{e^2 m r}{4N\pi\varepsilon_0\hbar^2}$$

$L_{N,l}(\xi)$ is a *generalised Laguerre polynomial* and $Y_{l,m}(\theta,\phi)$ is a spherical harmonic. The quantum numbers are,

- $N = 1, 2, 3, \ldots$ is the principal quantum number.

- $l = 0, 1, \ldots, N - 1$ is the total angular momentum quantum number.

- The integer m with $-l \leq m \leq +l$ is the quantum number for the z-component of angular momentum.

Now, comparing the full result to the Bohr model of atom, we see that the latter emerges for states with $m = l \simeq N >> 1$. In this case the z-component of angular momentum $L_z = m\hbar \simeq N\hbar$ and the total angular momentum $L = \sqrt{l(l+1)\hbar^2} \simeq N\hbar$.

The radial probability distribution,

$$
\begin{aligned}
r^2 g(r)^2 \quad &\sim \quad r^{2(l+1)} \, \exp\left(-\frac{\beta r}{2(l+1)}\right) \\
&\sim \quad r^{2N} \, \exp\left(-\frac{\beta r}{N}\right)
\end{aligned}
$$

Attains a maximum where,

$$
\frac{2N}{r} - \frac{\beta}{N} \;=\; 0 \tag{5.52}
$$

Thus the peak value is at,

$$
r_{\text{peak}} \quad \simeq \quad \frac{2N^2}{\beta} \;=\; N^2 r_1
$$

where $r_1 = 2/\beta$ is the Bohr radius. Thus the radial probability distribution is therefore peaked around the radius of the N'th Bohr orbit.

5.4   Towards the periodic table

We could try to generalise this discussion to atoms other than hydrogen. These have a nucleus with charge $+Ze$, orbited by Z independent electrons, where the atomic number Z is an integer greater than one. If we take the nucleus to be fixed, as we did with hydrogen, this means we need to solve the Schrödinger equation for Z independent electrons in a central Coulomb potential. This is not so simple, since the electrons also interact with each other. If we ignore this temporarily, we can obtain solutions of the form

$$
\chi(\mathbf{x}_1, ..., \mathbf{x}_Z) = \chi(\mathbf{x_1})...\chi(\mathbf{x}_Z), \tag{5.53}
$$

where the χ_j are the rescaled solutions for the hydrogen atom (the rescaling is because the nucleus has charge $+Ze$ instead of $+e$). The energy is just the sum

$$
E \;=\; \sum_{i=1}^{Z} E_i, \tag{5.54}
$$

It turns out that for relatively small atoms this gives qualitatively the right form, with corrections arising from the electron-electron interactions that can be calculated using perturbation theory (which you will do in Principles of Quantum Mechanics). However, we also need to allow for the Pauli exclusion principle, which implies that no two electrons in the same atom can be in the same state. So the lowest overall energy state is given by filling up the energy levels in order of increasing energy, starting with the lowest. Allowing for the twofold degeneracy arising from spin (which you will discuss in Principles of Quantum Mechanics) we have $2 \times N^2$ states in the N-th energy level. This gives us an atom with a full energy level with $Z = 2$, $10 = 8{+}2$, . . . for $N = 1, 2, \cdots$; these are the elements helium, neon...The elements with outer electrons in the 1st and 2nd energy levels fill out the corresponding first two rows of the periodic table. The analysis gets more complicated as atoms get larger, because electron-electron interactions become more important, and this qualitative picture is not adequate for the third and higher rows of the periodic table. We can understand that helium and neon are chemically inert (i.e. un-reactive) as a consequence of the fact that they have full energy levels, which turns out to be a very stable state that does not easily undergo transitions by capturing or losing electrons.
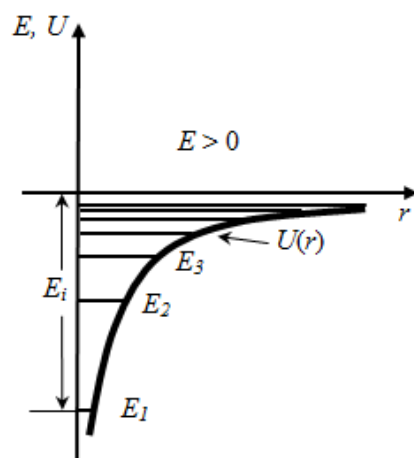
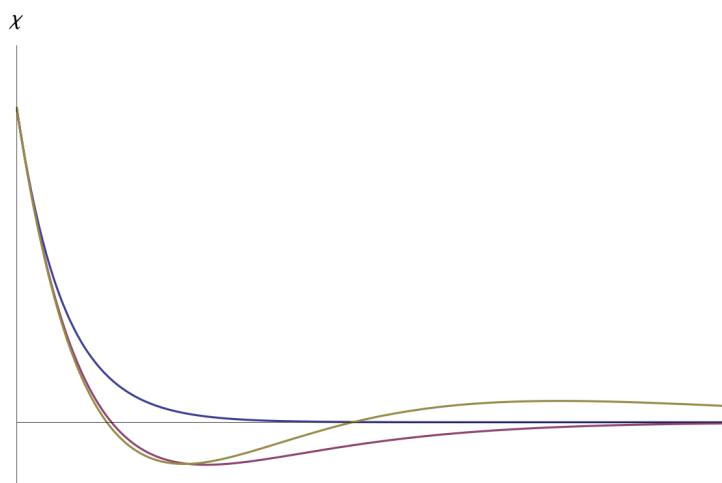Figure 5.4: The Coulomb potential and the energy levels of the hydrogen atom

Figure 5.5: Spherically symmetric eigenfunctions $\chi_1(r)$, $\chi_2(r)$, $\chi_3(r)$ corresponding to the the three lowest energy levels of the hydrogen atom.

Appendix A

Appendix

A.1 Fundamental constants

- Planck's constant, $\hbar = 1.05 \times 10^{-34} \; J\,s$

- Speed of light, $c = 3 \times 10^8 \; ms^{-1}$

- Charge of the electron, $e = 1.60 \times 10^{-19} \; C$

- Mass of the electron, $m_e = 9.11 \times 10^{-31} \; kg$

- Mass of proton, $m_p = 1.67 \times 10^{-27} \; kg$

- The vacuum permittivity constant, $\varepsilon_0 = 8.854 \times 10^{-12} \; F\,m^{-1}$.

A.2 Basic facts about waves

Complex wave-form,

$$A \exp\left(i\mathbf{k} \cdot \mathbf{x} - i\omega t\right)$$

It is useful to define the following quantities:

- Wave-vector $\mathbf{k}$.

- Wavelength λ is given as $\lambda = 2\pi/|\mathbf{k}|$.

- Angular frequency ω.

- Frequency $\nu = \omega/2\pi$.

The velocity of a wave v_{wave} is given as,

$$v_{\text{wave}} \quad = \quad \frac{\omega}{|\mathbf{k}|} = \nu\lambda.$$

For electromagnetic waves this is equal to c.

A.3   Derivation of the Planck formula for the black-body radiation*

In the last decades of the 19th century, physicists were very keen to understand the nature of black-body radiation, *i.e.* of the radiation emitted by a totally absorbing heated body that had come into thermal equilibrium at a given temperature T. The energy density $\rho(\nu, T)$ of the light radiated at the equilibrium temperature as a function of the frequency radiation ν had been measured and it was known on thermodynamic grounds to be an universal function of ν and T. How to calculate this function? The first one to attempt it was Lord Rayleigh in 1900. He modelled the black body as a cubical box of width L. *[1] Due to boundary conditions, the wavenumber $\mathbf{k}$ of the radiation field in the box must change by an integer multiple of 2π in a distance L. As a result, the radiation field is a Fourier sum over normal modes proportional to

$$\exp\left(i\mathbf{k}\cdot\mathbf{x}\right), \qquad \text{with} \qquad \mathbf{k} = \frac{2\pi\mathbf{n}}{L}, \tag{A.1}$$

where $\mathbf{n}$ is characterised by a vector of integers $(n_x,\ n_y,\ n_z) \in \mathbb{Z}$ and a polarisation state (either left- or right-circular polarisation). The frequency of each normal mode is given by

$$\nu = \frac{c}{\lambda} = \frac{|\mathbf{k}|c}{2\pi} = \frac{|\mathbf{n}|c}{L}. \tag{A.2}$$

Each normal mode occupies a cell of unit volume in the space of the vectors $\mathbf{n}$, so the number of normal modes $N(\nu)d\nu$ in the range of frequencies between ν and $\nu + d\nu$ is twice the volume of the corresponding shell in the $\mathbf{n}$ vector space.

$$N(\nu)d\nu = 2 \cdot 4\pi|\mathbf{n}|^2\,d|\mathbf{n}| \; = \; 8\pi\left(\frac{L}{c}\right)^3\nu^2 d\nu, \tag{A.3}$$

where the factor of 2 accounts for the two possible polarisation states. Rayleigh noted that in classical statistical mechanics, in any system that can be regarded as a collection of harmonic oscillators, the mean energy of each oscillator $\langle E(T)\rangle$ is simply proportional to the temperature, a relation written as $\langle E(T)\rangle = k_B T$, where k_B is a fundamental constant, known as Boltzmann's constant. If this is applied to the radiation inside the black body,*

---

[1]From here the content is non examinable until the next asterisk.

the energy density in the radiation between frequencies ν and $\nu + d\nu$ would then be given by what has come to be called the Rayleigh–Jeans formula[2]

$$\rho(\nu, T) \;=\; \frac{\langle E(T)\rangle N(\nu)}{L^3} \tag{A.4}$$

$$=\; \frac{8\pi k_B T \nu^2}{c^3}. \tag{A.5}$$

The prediction that $\rho(\nu, T)$ is proportional to $T\nu^2$ was actually in agreement with observation for small values of ν/T, but failed badly for larger values. Indeed, if it held for all frequencies at a given temperature, then the total energy density $\rho(\nu, T)d\nu$ would be infinite! This became known as the *ultraviolet catastrophe*.

The correct value was published shortly later by Max Planck. The initial formula was just guesswork, as he noted that the black-body radiation data could all be fitted by the formula

$$\rho(\nu, T) \;=\; \frac{8\pi h}{c^3} \frac{\nu^3}{\exp(h\nu/k_B T) - 1}, \tag{A.6}$$

where h was a constant, which was fitted to experimental data along with k_B to give

$$h \sim 6.610^{-27} \text{Joule} \times \sec \qquad \text{and} \qquad k_B \sim 1.410^{-16} \text{erg}/K. \tag{A.7}$$

The dimension of the new constant h (now universally knowns as the Planck constant), are

$$[h] = M\,L^2\,T^{-1} \;=\; [\text{energy}] \times [\text{time}] = [\text{position}] \times [\text{momentum}]. \tag{A.8}$$

Although at the beginning it appeared simply as a constant in a fit of a set of data, we now think of this constant as representing the "strength" of quantum effects. Despite having these new profound features, early 20$^{\text{th}}$ century physicists were guided by the expectation to recover classical physics in limit $h \to 0$. Indeed, in this limit, Eq. (A.6) reduces exactly to Eq. (A.5). After the initial formula to fit the data, Planck later gave a derivation of the above formula, based on the assumption that the radiation inside the back-body was quantised, *i.e.* the energy of the radiation comes in integer multiples of $h\nu$. A simpler derivation compared to the original by Planck is given by Weinberg in [3], following the derivation of the black-body formula by Hendrik Lorentz in 1910. The assumption by Rayleigh that the mean energy of each normal mode is simply proportional to the temperature T stemmed from the fact that in a system containing a large number of identical systems at thermal equilibrium at a given temperature T (like light modes in a black-body cavity), the probability that one of these systems has an energy E is proportional to $\exp(-E/k_B T)$. If the energies of light quanta were continuously distributed, this would give a mean energy

$$\langle E(T)\rangle \;=\; \frac{\int_0^\infty E \exp\left(-E/k_B T\right) dE}{\int_0^\infty \exp(-E/k_B T) dE} = k_B T. \tag{A.9}$$

---

[2] A numerical error in Rayleigh's derivation was corrected by J. Jeans in 1905.

The energy density in radiation between frequencies ν and $\nu+d\nu$ is again given by Eq. (A.4), and the above mean energy yields the Rayleigh-Jeans formula, Eq. (A.5). But if the energies are instead quantised, *i.e.* if they comes only as integer multiples of $h\nu$, then the mean energy is given by

$$\langle E(T) \rangle = \frac{\sum_{n=0}^{\infty} nh\nu \exp(-nh\nu/k_BT)}{\sum_0^{\infty} \exp(-nh\nu/k_BT)} = \frac{h\nu}{\exp(h\nu/k_BT) - 1}, \tag{A.10}$$

which now yields the Planck formula.

A.4   Basic facts about atoms

An atom has a positively charged <u>nucleus</u> surrounded by negatively charged electrons.
Nucleus: Z protons, each of positive charge $+e$. Z is known as the atomic number. Also has $A - Z$ neutrons, each of mass $m_n \simeq m_p$, which carry no electric charge. Total mass of the nucleus,

$$M = Zm_p + (A - Z)m_n \simeq Am_n$$

A is known as the atomic weight.
A neutral atom has Z electrons each of negative charge $-e$. If some are removed the atom becomes a positive ion.
The electrons are much lighter than the protons and neutrons of the nucleus: $m_e/m_p \simeq 1/1837$. Nearly all the mass of the atom resides in the nucleus.
Electrons are held in the atom by the electrostatic attraction between each electron and the nucleus.
Protons and neutrons are bound in the nucleus by the <u>strong nuclear force</u>. Though short-ranged, this is much stronger than the electrostatic repulsion between protons. The electrons do not feel the strong force.
Diameter of nucleus $\sim 10^{-15}$ m. Diameter of whole atom $\sim 10^{-10}$ m. Because the size of the nucleus is so much smaller than that of the whole atom, for the purpose of understanding atomic structure the nucleus can be treated as a point charge.
Chemical properties of atoms are determined only by the number of electrons Z.
Isotopes are atoms with the same value of Z but different A. They have the same chemistry but different radioactive properties.

A.5   Useful Dynamics and Relativity formulae

- **Non-relativistic mechanics** Free particle of mass m, moving at velocity $\mathbf{v}$. Speed $v = |\mathbf{v}| << c$. Momentum and energy of the particle are given as,

$$\mathbf{p} = m\mathbf{v} \qquad E = \frac{1}{2}mv^2$$

Thus we have

$$E = \frac{|\mathbf{p}|^2}{2m}$$

.

- **Relativistic mechanics** Free particle of mass m, moving at velocity $\mathbf{v}$. Speed $v = |\mathbf{v}|$. Momentum and energy of the particle are given as,

$$\mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1 - v^2/c^2}} \qquad\qquad E = \frac{mc^2}{\sqrt{1 - v^2/c^2}}$$

Thus we have,

$$E = \sqrt{m^2 c^4 + |\mathbf{p}|^2 c^2}.$$

For the special case of a massless particle this reduces to $E = c|\mathbf{p}|$.

A.6   Useful Integrals

$$\mathcal{I}(a) = \int_{-\infty}^{+\infty} dx \ \exp(-a\,x^2) = \sqrt{\frac{\pi}{a}} \tag{A.11}$$

The integral exists for complex a provided $\Re[a] > 0$. The integral,

$$\mathcal{I}_2(a) = \int_{-\infty}^{+\infty} dx \ x^2 \ exp(-a\,x^2) = \frac{1}{2}\sqrt{\frac{\pi}{a^3}} \tag{A.12}$$

is obtained by differentiating (A.11) with respect to the parameter a.

Another useful integral is,

$$\mathcal{J}(a,b) = \int_{-\infty}^{+\infty} dx \ \cos(bx) \ \exp(-a\,x^2) = \sqrt{\frac{\pi}{a}} \ \exp\left(-\frac{b^2}{4a}\right) \tag{A.13}$$

Again the integral exists for complex a and b provided $\Re[a] > 0$.

To prove (A.13), first note that

$$\mathcal{J}(a,0) = \mathcal{I}(a) = \sqrt{\frac{\pi}{a}}$$

Differentiating $\mathcal{J}(a, b)$ wrt to b yields,

$$\frac{\partial \mathcal{J}}{\partial b} = -\int_{\infty}^{+\infty} dx \; x \; \sin(bx) \; \exp(-a\, x^2)$$

Integrating by parts on the RHS then gives

$$\frac{\partial \mathcal{J}}{\partial b} = \left[\frac{\sin(bx)}{2a} \; \exp(-ax^2)\right]_{-\infty}^{+\infty} - \int_{\infty}^{+\infty} dx \; \frac{b}{2a} \; \cos(bx) \; \exp(-a\, x^2)$$

$$= -\frac{b}{2a} \; \mathcal{J}$$

Integrating this relation we obtain,

$$\mathcal{J}(a, b) = \mathcal{J}(a, 0) \exp\left(-\frac{b^2}{4a}\right) = \sqrt{\frac{\pi}{a}} \; \exp\left(-\frac{b^2}{4a}\right)$$

We can now use (A.11) and (A.13) to do the Gaussian integral discussed below equation (3.61) in the text,

$$\psi(x, t) = \int_{-\infty}^{+\infty} dk \; \exp\left(F(k)\right)$$

with,

$$F(k) = -\frac{1}{2}\alpha k^2 + \beta k + \delta$$

where the complex constants α, β and γ are defined in the text. Noting that $\exp(-iz) = \cos(z) - i\sin(z)$ for any complex number z we find,

$$\psi(x, t) = e^\delta \int_{-\infty}^{+\infty} dk \; (\cos(i\beta k) - i\sin(i\beta k)) \; \exp\left(-\frac{1}{2}\alpha k^2\right)$$

The second term in brackets vanishes as the integrand is an odd function of k. The remaining integral can be evaluated using (A.13) with $b = i\beta$ and $a = \alpha/2$, to get,

$$\psi(x, t) = \sqrt{\frac{2\pi}{\alpha}} \; \exp\left(\frac{\beta^2}{2\alpha} + \delta\right)$$

as claimed in the text.

A.7 Angular momentum operators in spherical polar coordinates

Relation between cartesian and spherical polar coordinates,

$$x_1 = r\sin(\theta)\cos(\phi) \quad x_2 = r\sin(\theta)\sin(\phi) \quad x_3 = r\cos(\theta)$$

Using the chain rule,

$$
\begin{aligned}
\frac{\partial}{\partial x_1} &= \left(\frac{\partial r}{\partial x_1}\right)\frac{\partial}{\partial r} + \left(\frac{\partial \theta}{\partial x_1}\right)\frac{\partial}{\partial \theta} + \left(\frac{\partial \phi}{\partial x_1}\right)\frac{\partial}{\partial \phi} \\
&= \sin(\theta)\cos(\phi)\frac{\partial}{\partial r} + \cos(\theta)\cos(\phi)\frac{1}{r}\frac{\partial}{\partial \theta} - \frac{\sin(\phi)}{r\sin(\theta)}\frac{\partial}{\partial \phi}
\end{aligned}
$$

and similarly for $\partial/\partial x_2$ and $\partial/\partial x_3$.

Proceeding in this way, we obtain

$$
\begin{aligned}
\hat{L}_1 &= -i\hbar\left(x_2\frac{\partial}{\partial x_3} - x_3\frac{\partial}{\partial x_2}\right) \\
&= i\hbar\left(\cos(\phi)\cot(\theta)\frac{\partial}{\partial \phi} + \sin(\phi)\frac{\partial}{\partial \theta}\right)
\end{aligned}
$$

Similarly we find

$$
\hat{L}_2 = i\hbar\left(\sin(\phi)\cot(\theta)\frac{\partial}{\partial \phi} - \cos(\phi)\frac{\partial}{\partial \theta}\right)
$$

and,

$$
\hat{L}_3 = -i\hbar\frac{\partial}{\partial \phi}
$$

Finally can check that,

$$
\hat{L}^2 = \hat{L}_1^2 + \hat{L}_2^2 + \hat{L}_3^2 = -\frac{\hbar^2}{\sin^2(\theta)}\left[\sin(\theta)\frac{\partial}{\partial \theta}\left(\sin(\theta)\frac{\partial}{\partial \theta}\right) + \frac{\partial^2}{\partial \phi^2}\right]
$$

The alternating tensor Indices i, j, $k = 1, 2, 3$. $\varepsilon_{ijk} = 0$ unless all indices are different, $i \neq j \neq k \neq i$. If all indices are different, then $\varepsilon_{ijk} = +1$ if (ijk) is a cyclic permutation of (123) and -1 otherwise.

Bibliography

[1] Richard Phillips Feynman, Robert Benjamin Leighton, and Matthew Sands. The Feynman lectures on physics; New millennium ed. Basic Books, New York, NY, 2010. Originally published 1963-1965.

[2] S. Gasiorowicz. Quantum Physics. Wiley, 2003.

[3] D. J. Griffiths. Introduction to Quantum Mechanics. Prentice Hall Inc, 1995.

[4] J. Napolitano J. J. Sakurai. Modern quantum mechanics. Addison-Wesley, 2011.

[5] L. D. Landau and E. .M. Lifshitz. Quantum mechanics (non-relativistic theory). 2003.

[6] A.I.M. Rae and J. Napolitano. Quantum Mechanics. CRC Press, 2015.

[7] Steven Weinberg. Lectures on Quantum Mechanics. Cambridge University Press, 2 edition, 2015.