

Part IB Quantum Mechanics, Michaelmas 2019

Tuesday and Thursday, 11.05 a.m. to 11.55 a.m., Mill Lane lecture room 3

Prof. Adrian Kent (apak@damtp.cam.ac.uk)

I am very grateful to previous lecturers of this course, and in particular to Jonathan Evans and Eugene Lim, who very generously shared their lecture notes and thoughts on the material. These notes have been influenced by their presentations of many of the course topics, and sometimes draw directly on their discussions. I am also grateful to past students of the course, whose feedback and corrections have been very helpful. More feedback – whether noting typos or other errors or more general comments – would be 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 at the start and the end. Some footnotes also contain references to results proved in textbooks or other references but not in the notes: these proofs too are, obviously, non-examinable.

Further course material will be added from time to time on the course web page, which is linked from www.qi.damtp.cam.ac.uk (follow the link to Undergraduate and Masters Lecture Courses on the left menu).

If you wish to make additional notes during lectures you will probably find it simplest to make them on separate sheets of paper, with footnote numbers to refer them to the appropriate place in the printed notes.

Highlights of this course:

- Historical development of quantum mechanics
- The one-dimensional Schrödinger equation; solution for particles in various potentials; probabilistic interpretation; beam scattering and tunneling.
- The basic formalism of quantum mechanics – states, operators, observables, measurement, the uncertainty principle: a new way of treating familiar dynamical quantities (position, momentum, energy, angular momentum).
- Quantum mechanics in three dimensions: the 3D Schrödinger equation, angular momentum, the hydrogen atom and other solutions of the 3D SE.

Version dated 31.12.19. Any updated versions will be placed on the course web page linked from www.qi.damtp.cam.ac.uk.

Recommended books

(*) R. Feynman, R. Leighton and M. Sands, *Feynman lectures on physics*, volume 3, chapters 1-3 (Addison-Wesley, 1989).

A beautifully written and profoundly thoughtful introduction to some of the basic ideas of quantum theory. Feynman was one of the twentieth century's most creative physicists. As these chapters illustrate, he also thought very deeply and carefully about fundamental questions in physics and about the scientific process itself. I really recommend these chapters very strongly as background reading for the course.

(*) A. Rae and J. Napolitano, *Quantum Mechanics*, chapters 1-5 (IOP Publishing, 2002).

A very good textbook which covers a range of elementary and more advanced topics in quantum theory, including a short discussion of the conceptual problems of quantum mechanics. The first five chapters form a good course text for IB QM.

(*) S. Gasiorowicz, *Quantum Physics* (Wiley 2003).

A very nicely and clearly written book, with good illustrations, which covers most of the course material well.

P. Landshoff, A. Metherell and G. Rees, *Essential Quantum Physics* (Cambridge University Press, 2001).

Another nicely and concisely written textbook, which also covers most of the course material for IB QM well.

P. A. M. Dirac, *The Principles of Quantum Mechanics* (Oxford University Press, 1967; reprinted 2003).

A more advanced treatment. Despite its age it is still a valuable exposition of the insights and perspectives of one of the pioneers of quantum theory. You may want to consider looking at this if you intend to go beyond this course and pursue Part II and Part III courses in quantum physics.

S. Brandt and H.D. Dahmen, *The Picture Book of Quantum Mechanics* (4th edition; Springer, 2012). An excellent book with accompanying plots and simulations. These give visual explanations that nicely illuminate the mathematics of solutions of the Schrödinger equation, tunnelling, reflection from barriers, atomic electron states, and other quantum phenomena.

(*) Particularly recommended.

Contents

1	Quantum Mechanics, science and technology	5
1.1	Quantum Mechanics and fundamental science	5
1.2	Quantum Mechanics and technology	6
2	Historical development of quantum mechanics	7
2.1	Planck's quantum hypothesis (1900)	7
2.2	Einstein's explanation of the photoelectric effect (1905)	8
2.3	Diffraction of single photons (1909)	9
2.4	The Rutherford atom (1911)	10
2.5	The Bohr atom (1913)	10
2.6	Compton scattering (1923)	14
2.7	Wave and particle models of electromagnetic radiation	15
2.8	De Broglie waves (1924)	15
2.9	Matter wave diffraction (1923-7)	16
2.10	Discussion of the double slit experiment	16
2.10.1	What can we conclude from double slit experiments?	19
2.11	*Ongoing tests of quantum theory	21
2.12	Closing comments	21
3	The one-dimensional Schrödinger equation	23
3.1	The 1D Schrödinger equation for a free particle	23
3.2	The momentum operator	24
3.3	The 1D Schrödinger equation for a particle in a potential	25
3.4	The wavefunction	26
3.4.1	What is the wavefunction?	26
3.4.2	What the wavefunction definitely isn't	27
3.5	The superposition principle	27
3.6	Probabilistic interpretation of the wavefunction: the Born rule	28
3.6.1	Probability density and probability current	29
4	Solutions of the 1D Schrödinger equation	31
4.1	Example I: Stationary states	31
4.2	Completeness of the energy eigenfunctions	32
4.3	Example II: Gaussian wavepackets	34
4.4	Example III: Particle in an infinite square potential well	35
4.4.1	Handling discontinuities	36
4.5	Parity	38
4.6	Example IV: Particle in a finite square well	38
4.7	Example V: The quantum harmonic oscillator	41
5	Tunnelling and Scattering	45
5.1	Scattering states	45
5.2	Interpretation of plane wave scattering solutions	46
5.3	Example I: The potential step	47
5.4	Example II: The square potential barrier	49
5.4.1	*Important examples	51
6	The basic formalism of quantum mechanics	52
6.1	Spaces of functions	52
6.2	The inner product	52
6.2.1	Properties of the inner product	52
6.3	Operators	53

6.3.1	Some examples of operators	53
6.4	Hermitian operators	53
6.4.1	Classical states and dynamical variables	53
6.4.2	Quantum states and observables	54
6.5	Some theorems about hermitian operators	55
6.6	Quantum measurement postulates	57
6.7	Expectation values	58
6.8	Commutation relations	59
6.8.1	The canonical commutation relations	60
6.9	Heisenberg's Uncertainty Principle	61
6.9.1	What does the uncertainty principle tell us?	63
6.10	Ehrenfest's theorem	64
6.10.1	Applications of Ehrenfest's theorem	64
6.11	*The harmonic oscillator revisited	66
7	The 3D Schrödinger equation	68
7.1	Quantum mechanics in three dimensions	68
7.2	Spherically symmetric potentials	69
7.3	Examples of spherically symmetric potentials	71
7.3.1	The spherical harmonic oscillator	71
7.3.2	The spherical square well	71
7.4	Spherically symmetric bound states of the hydrogen atom	72
7.5	Canonical Commutation Relations in 3D	74
7.6	Orbital Angular Momentum	74
7.6.1	Angular momentum commutation relations	75
7.6.2	Angular momentum operators in spherical polar coordinates	76
7.7	Solving the 3D Schrödinger equation for a spherically symmetric po- tential	78
7.7.1	Degeneracies	79
7.8	The Hydrogen atom	80
7.8.1	Energy level degeneracies	81
7.9	Towards the periodic table	81

1 Quantum Mechanics, science and technology

1.1 Quantum Mechanics and fundamental science

Quantum mechanics is the non-relativistic version of relativistic quantum field theory. Collectively, we refer to these theories as “quantum theory”. Most physicists would agree quantum theory is the most remarkable, interesting and surprising physical theory we have discovered.¹ Although ultimately only a non-relativistic part of a larger theory, quantum mechanics already teaches us that our universe follows laws that involve beautiful and intricate mathematics, and whose form we could not possibly have imagined had they not been illuminated by experiment.

Among other things, quantum mechanics explains the essential features of the following:

- the structure of atoms and molecules and their chemical interactions; i.e. chemistry, and biochemistry and so, in principle, biology. We will begin to discuss this towards the end of the course, when we consider quantum mechanical descriptions of the hydrogen atom and, more qualitatively, other atoms. This description is made more precise and taken further in Part II Principles of Quantum Mechanics.
- the structure and properties of solids (and so, in principle, much of classical mechanics). Conductivity (some basic theory of which is introduced in Part II Applications of Quantum Mechanics) and superconductivity.
- the thermodynamics of light and other electromagnetic radiation and also how electromagnetic radiation interacts with matter. To describe this properly requires relativistic quantum field theory, which isn’t covered until Part III.
- the physics of subatomic particles, radioactivity, nuclear fission and fusion. Again, we need relativistic quantum field theory to describe these phenomena in full detail. But, as we’ll see later, even elementary quantum mechanics gives us some useful insights into the physics of nuclear fission and fusion. For example, we can understand the random nature of these processes, and the way fusion and fission rates depend on relevant potentials, as a consequence of quantum tunnelling.

Modern cosmological models are also based on quantum theory. Since we do not have a quantum theory of gravity, and do not know for sure whether there is one, these cosmological models are at best incomplete. Nonetheless they give a good qualitative fit to observational data. Many physicists hope that this project can be completed, so that we can describe the entire universe and its evolution by quantum theory.² These topics are covered in detail in the Part II and Part III Cosmology courses.

¹Einstein’s general theory of relativity is the only other contender. It is an extraordinarily beautiful theory that transformed our understanding of space and time and their relationship to matter, and that gave us the tools to understand the cosmos. The two theories are fundamentally incompatible, and it is uncertain precisely which parts of which theory will survive in a future unification. Still, I find quantum theory more intriguing, because of the variety of deep mathematical concepts it combines and because it is so different from and so much stranger than the theories that preceded it.

²It is also true that many thoughtful physicists suspect that the project cannot be completed, because we will need something other than a quantum theory of gravity, or because quantum theory will turn out not to hold on large scales. If so, comparing quantum cosmological models to observation should eventually give us insight into new physics.

1.2 Quantum Mechanics and technology

Many of the revolutionary technological developments of the last hundred years rely on quantum mechanics:

- semiconductor physics – transistors, diodes, integrated circuits, and hence the computing and IT industries.
- laser physics
- nuclear power and the as yet unrealised dream of controllable fusion power
- tunnelling electron microscopy and atomic manipulation
- More recent inventions relying on the distinctive properties of *quantum information*. These include quantum cryptography. Its best known application is quantum key distribution, which in principle allows perfectly secure communication, and is now practical with large data transmission rates and over long distances, including between satellite and ground stations. There are many other applications, including quantum authentication, quantum position verification, quantum bit string commitment, quantum multi-party computation and quantum digital signatures. Quantum information allows perfect security for some of these tasks, and security advantages for others.

Another major development has been the invention of quantum algorithms and various types of quantum computers. We now know that quantum computers are substantially more efficient than classical computers for some important applications, including factorisation (at least compared to the best known classical algorithms) and the simulation of quantum systems. Small scale quantum computers have been built, and many research groups are competing to build a quantum computer large enough to exploit the theoretical advantages. In September 2019, Google claimed the first demonstration of so-called “quantum supremacy”, a calculation on a quantum computer that is infeasible on classical computers.³

Other fascinating applications include quantum teleportation – which in principle gives a way of effectively deleting a physical system at A and recreating it at B without sending it along a path from A to B – and other types of quantum communication.

All of these topics are covered in the Part II Quantum Information and Computation course. Notes for that course, and for Part III courses in this area, along with some information about research work on these topics in DAMTP, can be found on <http://www.qi.damtp.cam.ac.uk>.

³There are nuances of definition here, and the community is still assessing the claim.

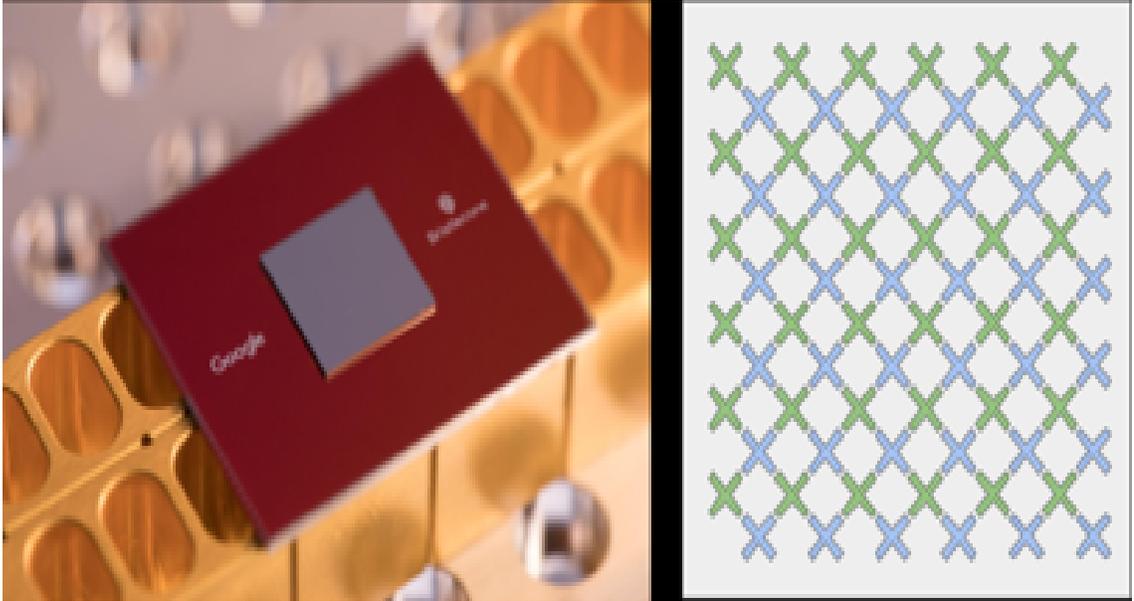


Figure 1: Google's Bristlecone quantum processor. Image on left, map of qubit connections on right.

2 Historical development of quantum mechanics

2.1 Planck's quantum hypothesis (1900)

One of the great puzzles in late 19th century physics was the inability of classical thermodynamics and electromagnetism to predict the correct spectrum — or indeed *any* sensible spectrum — for the frequency distribution of radiation from an idealized black body. Classical thermodynamics predicted an emission spectrum which suggested that the flux of emitted radiation tends to infinity as the frequency tends to infinity, and hence that an infinite amount of energy is emitted per unit time. In 1900, Max Planck showed that one could predict the experimentally observed spectrum by postulating that, because of some (at that point) unknown physics, matter can emit or absorb light of frequency ν only in discrete *quanta* which have energy

$$E = h\nu = \hbar\omega . \quad (2.1)$$

Here $\omega = 2\pi\nu$ is the light's *angular frequency* (the number of radians of oscillation per second), and

$$\hbar = \frac{h}{2\pi} \approx 1.055 \times 10^{-34} \text{ Joule sec} . \quad (2.2)$$

The constant h is a new constant of nature (Planck's constant). For most purposes it turns out to be notationally more convenient to use \hbar rather than h , and we will generally do so.

2.2 Einstein's explanation of the photoelectric effect (1905)

Experiment shows that light hitting a metal surface in a vacuum can cause electrons to be ejected with a range of energies. To emit any electrons, the incident light needs to have angular frequency ω satisfying $\omega \geq \omega_{\min}$, where ω_{\min} is a constant depending on the particular metal. When $\omega \geq \omega_{\min}$, one finds that the maximum energy of the emitted electrons, E_{\max} , obeys

$$E_{\max} = \hbar\omega - \hbar\omega_{\min} \equiv \hbar\omega - \phi, \quad (2.3)$$

where ϕ is the so-called *work function* of the metal. The average rate of electron emission is found to be proportional to the intensity of the incident light, but individual electrons appear to be emitted at random (and so in particular, measured over small enough time intervals, the emission rate fluctuates).

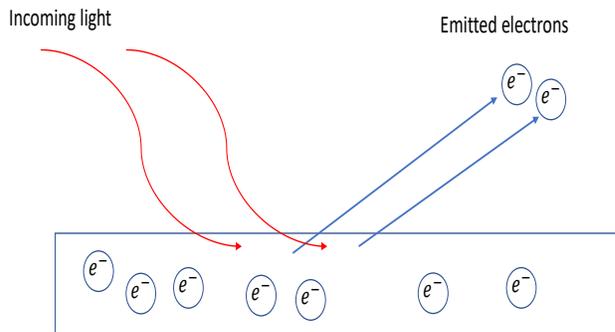


Figure 2: Schematic illustration of the photoelectric effect.

Although light had been understood as an electromagnetic wave, it was hard to explain these effects in terms of a 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 (i.e. we would not expect random fluctuations), 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.

To explain the photoelectric effect, Albert Einstein (in 1905) was led to postulate instead that light is *quantised* in small packets called *photons*, and that a photon of angular frequency ω has energy $\hbar\omega$. He reasoned that two photons are very unlikely to hit the same electron in a short enough time interval that their combined effect would knock the electron out of the metal: the energy an electron acquires from being hit by a photon is very likely to have dissipated by the time it is hit again.

Thus, one can explain the photoelectric effect as the result of single photons hitting electrons near the metal surface, if one assumes that an electron needs to acquire kinetic energy $\geq \phi$ to overcome the binding energy of the metal. An electron which acquires energy $\hbar\omega$ thus carries away energy $\leq \hbar\omega - \phi = E_{\max}$. One can also 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 photon arrival. However, individual photons arrive, and hit electrons so as to knock them out of the metal, at random – hence the randomly distributed emissions of electrons.

2.3 Diffraction of single photons (1909)

In Cambridge in 1909, J.J. Thomson suggested to G.I. Taylor (who had asked for a research project) that he investigate the interference of light waves of very low intensity. Taylor carried out an experiment in which a light source was successively filtered so that the energy flux was equivalent to the flux of a source sending no more than one photon at a time through the apparatus. He then observed the photographic image built up by diffraction of this feeble light around a needle. The characteristic diffraction pattern – the same pattern seen for strong light sources – was still observed.

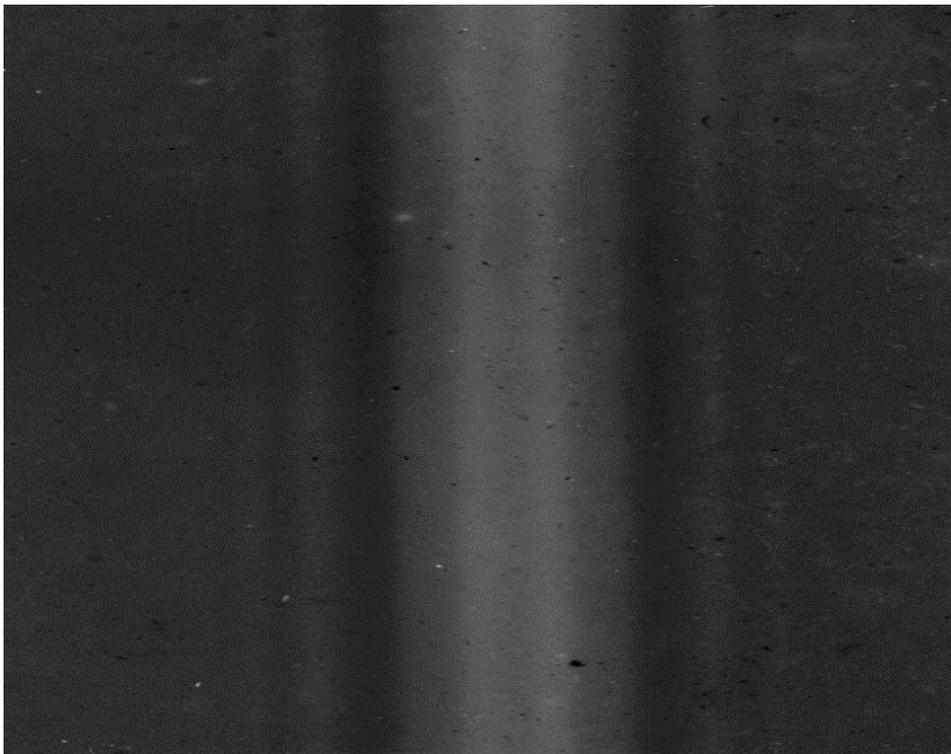


Figure 3: Photograph of diffraction bands caused by a thin wire in feeble light. See: Taylor, Proc. Camb. Phil. Soc., 15, 114, 1909. The exposure was 400 hours. Source: Cavendish Laboratory. Licensed under Creative Commons.

This seems to suggest that single photons propagate through the apparatus and nonetheless “self-interfere” in such a way that the diffraction pattern is cumulatively built up. This is indeed how Taylor’s results were interpreted for several decades after the development of quantum mechanics in 1926. Problems were later noticed

with this interpretation of this particular experiment: having the average energy flux of a single photon does not always imply that a source is emitting single photons. However, the experiment was subsequently repeated with genuine single photon sources, still showing the same diffraction pattern. Qualitatively similar patterns are seen in diffraction experiments for single electrons and other types of matter (see below).

2.4 The Rutherford atom (1911)

After discovering the electron in 1897, J.J. Thomson proposed a model of the atom as a sort of “plum pudding” with Z pointlike electrons of charge $-e$ embedded in a sphere of positive charge $+Ze$.

Geiger and Marsden’s famous experiment, carried out at 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, and Rutherford thought it unlikely anything interesting would be observed. But, in fact, some α -particles *were* observed to be scattered through 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.”

The scattering suggests a high density positive charge within the atom. Rutherford thus postulated a new model of the atom, with a heavy nearly pointlike nucleus, of charge $+Ze$, surrounded by Z electrons in orbit.

(A short popular account can be found at <http://physicsopenlab.org/2017/04/11/the-rutherford-geiger-marsden-experiment/>.)

2.5 The Bohr atom (1913)

Although the Rutherford atom was more compatible with the Geiger-Marsden 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. 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):

$$\omega_{mn} = 2\pi c R_0 \left(\frac{1}{n^2} - \frac{1}{m^2} \right) \quad \text{for } m > n, \quad (2.4)$$

where the *Rydberg constant*

$$R_0 \approx 1.097 \times 10^7 \text{ m}^{-1}. \quad (2.5)$$

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.

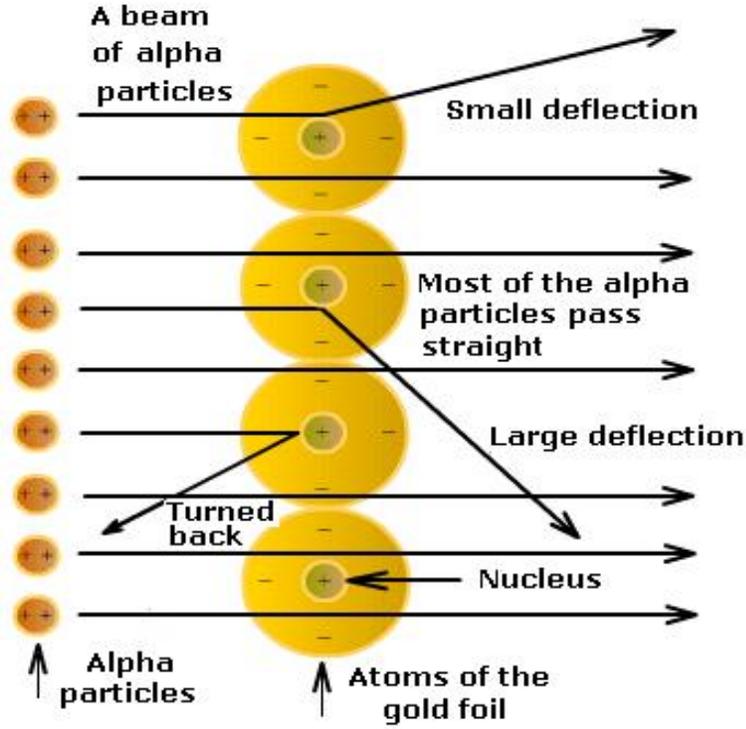


Figure 4: Schematic illustration of the Geiger-Marsden experiment, from the web page cited above.

Niels Bohr, in 1913, 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

$$L = n\hbar, \quad (2.6)$$

where n is a positive integer.

Thus, if we take an electron e moving with velocity v in a circular orbit of radius r about a proton p , $F = m_e a$ gives us that the Coulomb force

$$\frac{e^2}{4\pi\epsilon_0 r^2} = \frac{m_e v^2}{r}. \quad (2.7)$$

If we also have

$$L = m_e v r = n\hbar \quad (2.8)$$

then

$$\frac{n^2 \hbar^2}{r^3 m_e} = \frac{e^2}{4\pi\epsilon_0 r^2} \quad (2.9)$$

and hence

$$r = n^2 \left(\frac{\hbar^2 4\pi\epsilon_0}{m_e e^2} \right) = n^2 a_0, \quad (2.10)$$

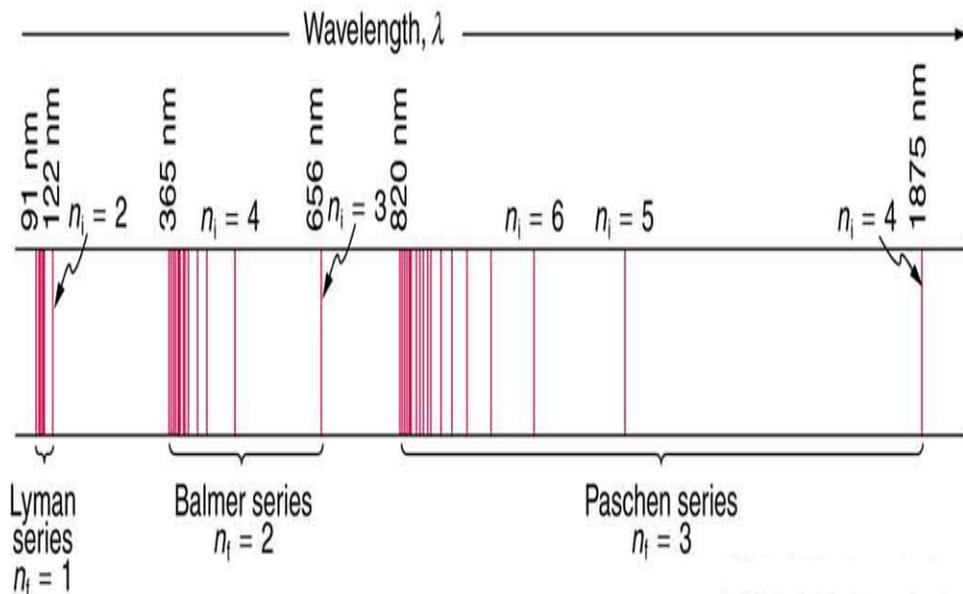


Figure 5: Spectra for the hydrogen atom. The figure shows three horizontal lines at small distances from each other. Between the two lower lines, the Lyman series, with four vertical red bands in compact form, is shown. This has $n_f = 1$ and $n_i \geq 2$, and wavelengths in the range 91 – 100 nanometers. The Balmer series is shown to the right side of this series. This has $n_f = 2$ and $n_i \geq 3$, and wavelengths in the range 365 – 656 nanometers. At the right side of this, the Paschen series bands are shown. This has $n_f = 3$ and $n_i \geq 4$, and wavelengths in the range 820 – 1875 nanometers. The Rydberg formula is obtained by taking $n_f = n$, $n_i = m$.

Source for figure and caption: <https://opentextbc.ca/physicstestbook2>
Image licenced under Creative Commons.

where

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} \approx 0.53 \times 10^{-10} \text{ m} \quad (2.11)$$

is the *Bohr radius*.

We can then obtain the energy of the n -th Bohr orbit from (2.7) and the Coulomb potential:

$$E_n = \frac{1}{2}m_e v^2 - \frac{e^2}{4\pi\epsilon_0 r} = -\frac{e^2}{8\pi\epsilon_0 r} = -\frac{e^2}{8\pi\epsilon_0 n^2 a_0} = -\frac{e^4 m_e}{32\pi^2 \epsilon_0^2 \hbar^2} \frac{1}{n^2} = \frac{E_1}{n^2}, \quad (2.12)$$

where

$$E_1 = -\frac{e^4 m_e}{32\pi^2 \epsilon_0^2 \hbar^2} \approx -13.6 \text{ eV}. \quad (2.13)$$

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 $E_{mn} = E_m - E_n$. Using $E_{mn} = \hbar\omega_{mn}$, where ω_{mn} is the angular frequency of the emitted photon, we have

$$\omega_{mn} = 2\pi R_0 c \left(\frac{1}{n^2} - \frac{1}{m^2} \right), \quad (2.14)$$

where

$$R_0 = \frac{m_e c}{2\hbar} \left(\frac{e^2}{4\pi\epsilon_0 \hbar c} \right)^2, \quad (2.15)$$

which agrees well with the experimentally determined value of the Rydberg constant.

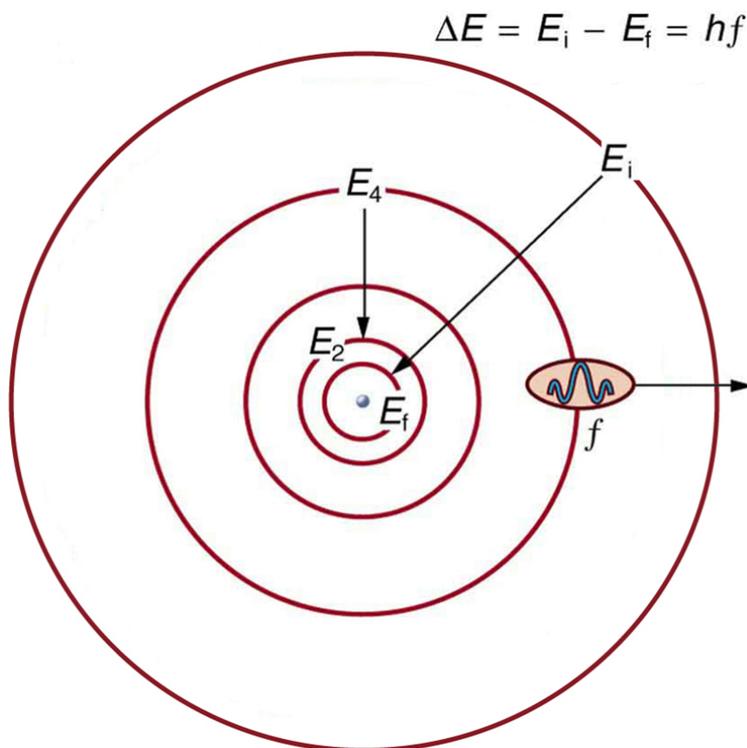


Figure 6: The orbits of Bohr's planetary model of an atom; five concentric circles are shown. The radii of the circles increase from innermost to outermost circles. On the circles, labels E_1 , E_2 , up to E_i are marked. Source for figure and caption: <https://opentextbc.ca/physicstestbook2>
Licenced under Creative Commons.

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. (We now understand that these are the spectra produced by the innermost electrons, which can be modelled in a way qualitatively similar to the electron in the hydrogen atom.)

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 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.

2.6 Compton scattering (1923)

In 1923, Arthur Holly Compton observed the scattering of X-rays by electrons associated with atoms in a crystal. Because the X-ray energies were much larger than the electron binding energies, the electrons can effectively be modelled by free electrons. Indeed, we also directly observe that if an electron beam and an X-ray beam converge, some electrons and some X-rays are deflected. This is very difficult to reconcile with a pure wave model of electromagnetic radiation, because the energy and momentum transfer for individual scatterings does not depend on the intensity of the X-ray beam.

A simple alternative explanation is that 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.) 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

$$\underline{p} = \hbar \underline{k}, \tag{2.16}$$

where \underline{k} is the wave vector of the corresponding electromagnetic wave, so that

$$|\underline{p}| = \frac{\hbar\omega}{c} = \hbar|\underline{k}|. \tag{2.17}$$

2.7 Wave and particle models of electromagnetic radiation

We thus see the emergence of two different models of light and other electromagnetic radiation.

Sometimes (classical electromagnetism, diffraction experiments with a strong light source, ...) it is useful to model light in terms of waves:

$$e^{i(\underline{k}\cdot\underline{x}-\omega t)}, \quad (2.18)$$

where \underline{k} is the wave vector, ω the angular frequency, $c = \frac{\omega}{|\underline{k}|}$ the speed of light in a vacuum, and the wavelength $\lambda = \frac{2\pi c}{\omega} = \frac{2\pi}{|\underline{k}|}$.

Sometimes (photoelectric effect, spectroscopy in the Bohr model of the hydrogen atom, Compton scattering, ...) it is useful to adopt a particle model, in which light is made up of photons with energy and momentum

$$E = \hbar\omega, \quad \underline{p} = \hbar\underline{k}. \quad (2.19)$$

The word “model” is chosen deliberately here. A model can be useful (as the wave and particle models of light are, in the appropriate contexts) without being completely correct. Indeed, G.I. Taylor’s 1909 demonstration of single photon diffraction already gave an example of a single experiment for which neither the wave nor the particle model of light appeared to be adequate. A simple particle model would not predict the observed diffraction pattern, while a simple wave model cannot explain the observation of single photons recorded on the photographic film.⁴

2.8 De Broglie waves (1924)

Louis de Broglie reexamined and extended Einstein’s photon hypothesis. If, he argued, Einstein was right that light waves can be considered as composed of particles – photons – might it not equally be the case that objects like electrons, which were thought of as particles, could exhibit wave-like behaviour?

As he pointed out in his 1924 PhD thesis, this would make the Bohr angular momentum quantization condition

$$L = pr = n\hbar \quad (2.20)$$

at least somewhat less mysterious. If we suppose that an electron of momentum p can (somehow) be thought of as a wave with *de Broglie wavelength*

$$\lambda = \frac{2\pi\hbar}{p}, \quad (2.21)$$

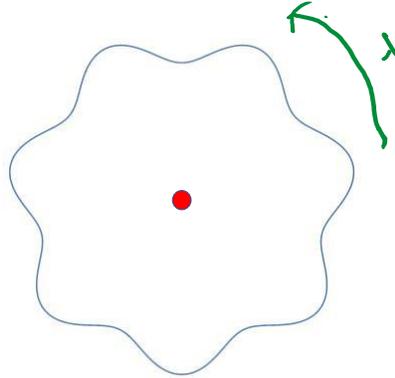
then the electron in the n -th Bohr orbit would be in a standing wave pattern with n wavelengths:

$$2\pi r = \frac{2\pi n\hbar}{p} = n\lambda. \quad (2.22)$$

De Broglie hypothesised that the frequency and wavevectors of the relevant wave were related to the energy and momentum of the particle as for photons: that is, we have $E = \hbar\omega$, $\underline{p} = \hbar\underline{k}$ even for particles with nonzero mass. (In fact, he speculated – incorrectly, according to our current understanding – that photons might also have a small nonzero mass.)

Einstein wrote, on learning of de Broglie’s hypothesis, that

⁴See earlier comment: to justify this fully we need to consider versions of the experiment with a guaranteed single photon source.



De Broglie's model of the atom. Electrons occupy orbits corresponding to an integer multiple of their de Broglie wavelength.

Figure 7: Schematic illustration of the de Broglie model of electron orbits in an atom.

“I believe it is the first feeble ray of light on this worst of our physics enigmas.”

It was.

2.9 Matter wave diffraction (1923-7)

We know that the wave model of light predicts, correctly, that light should form interference and diffraction patterns. De Broglie's hypothesis suggests the same should be true of electrons and other massive particles. This was first confirmed in experiments carried out (from 1923-7) by Davisson and Germer and (independently in 1927) by G.P. Thomson, who observed diffractive scattering of electrons from metallic crystals, with diffraction patterns consistent with the de Broglie wavelength $\lambda = \frac{2\pi\hbar}{p}$.

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.

Many diffraction experiments with electrons, neutrons and other particles have since been carried out, all confirming de Broglie's prediction.

2.10 Discussion of the double slit experiment

(Cf Feynman volume III chapters I-III)

A nice version of the double slit experiment with electrons was carried out by Akira Tonomura. It is described at <http://www.hitachi.com/rd/portal/highlight/quantum/doubleslit/index.html>

Like the other diffraction experiments mentioned above, this shows that electrons and other massive particles can produce interference patterns in the same sort of way as light and water waves. At the same time, it gives very vivid evidence of electrons being detected as individual particles. We stress again that *both* the wave description *and* the particle description are just models that are sometimes useful but, as this experiment again illustrates, not fundamentally correct. To analyse this

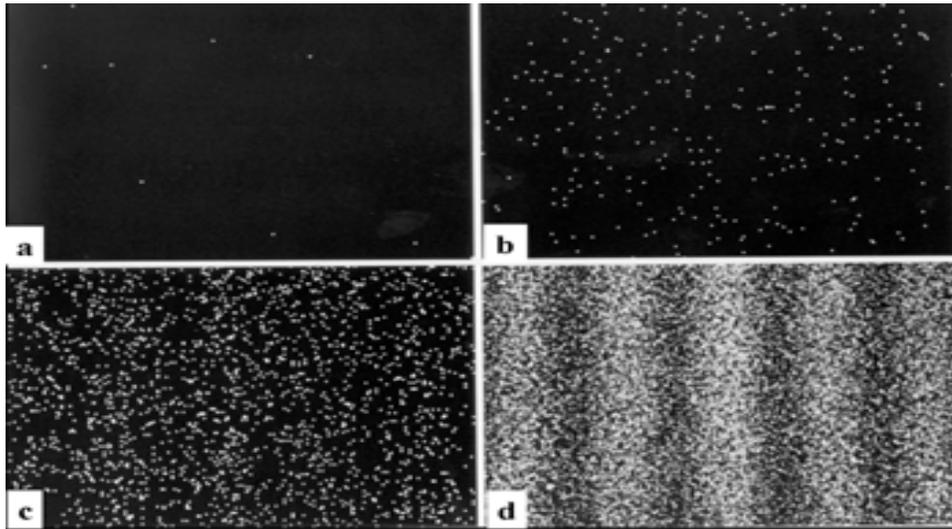


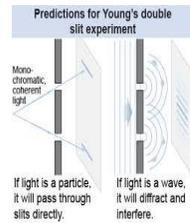
Figure 8: Single electron detection events building up an interference pattern in Tonomura's experiment. Copyright in these documents published on Hitachi World-Wide Web Server is owned by Hitachi, Ltd.

conclusion in more detail, let's follow Feynman in considering an idealised version of the double slit experiment, in which we assume we have perfect detectors that can register the passage of an electron without affecting it.

The observed interference pattern in a double slit experiment agrees with that predicted by a wave model (and disagrees with that predicted by a particle model). *But* the electrons arrive individually at the detector, which registers each time one arrives — as a particle model (but not a wave model) would suggest. The same is true of other massive particles, and also of photons. Even if we reduce the intensity of the source so that *only one* electron on average is between the source and the screen at any given time, we *still* see individual electrons detected in a pattern that cumulatively reproduces the distribution predicted by the wave model.

It's tempting to think that, when electrons leave the source, they behave like bullets from a gun — i.e. like particles coming from a well-defined small region. Certainly if we place a detector near the source this is what the detections suggest (although if we do this the electrons don't continue into the rest of the apparatus). It's also tempting to think that, since the electrons (etc.) arrive at the detector and are detected there as particles, with a definite or nearly definite position, they must have behaved like classical particles throughout, following some definite path from the source, through one hole or the other, to the detector. But, tempting though this last intuition may be, it's hard to reconcile with the observed interference pattern. If the electrons behaved like bullets throughout, we'd expect something like a superposition of two Gaussian distributions from the two slits, instead of the pattern of minima and maxima we observe.

Quantum weirdness: the double slit experiment



The observed patterns for one and two slits fit very well with a wave model of light, and seem to refute a model in which light is made up of particles.

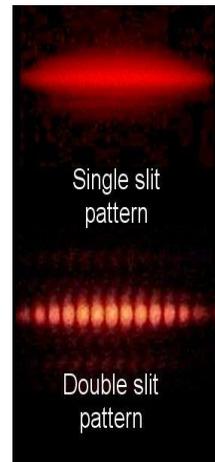


Figure 9: The double slit experiment.

There seems, in any case, a simple way to investigate more closely. Suppose that we have ideal detectors, which register that an electron has passed through a region, but don't obstruct its path. We can set up a double slit experiment with one of these ideal detectors adjacent to each slit (for definiteness, let us say they are on the far side, between the slit and the screen). Now we don't need to speculate: we can *observe* which slit the electron goes through. **But** when we do this, the interference pattern *changes*: we see a superposition of two Gaussians, as predicted by a particle model, rather than the pattern of maxima and minima predicted by the wave model and observed in the previous experiment.

What is really going on in the double slit experiment?

Why not just look and see which slit the photon goes through?

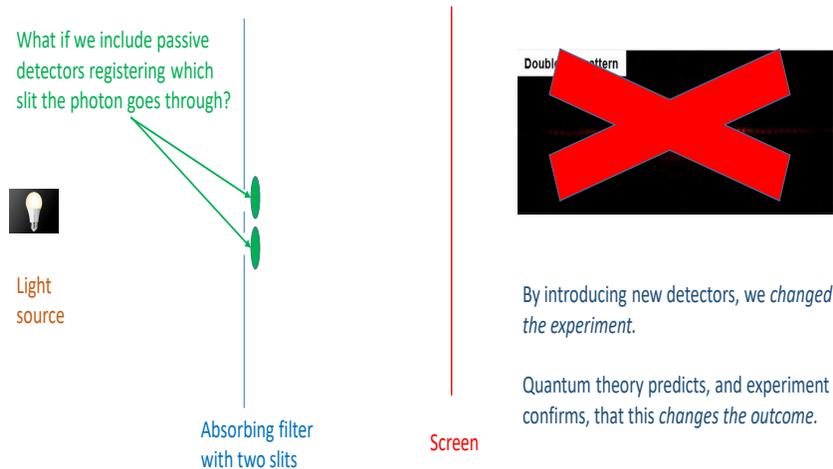


Figure 10: Observing which slit the electron goes through in the double slit experiment.

2.10.1 What can we conclude from double slit experiments?

1. As we already stressed, the wave model and the particle model are just that — models. *Neither* of them is adequate to explain the behaviour of electrons, photons, or other objects. Each of them can sometimes give a partial explanation of our observations, but that explanation *is not consistent with all the data*.

2. In particular, the type of reasoning about electron paths that would apply within a particle model does not generally apply in Nature. We can't assume that the electron follows a definite path through one slit or the other, and we can't assume (as we could with a classical particle) that observing which slit it goes through makes no difference to the interference pattern.

3. Some textbooks summarize the state of affairs described in point 1 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 rather than just a catchy name for it.

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 senses of either of those words. The electron is something different again, though it has some features in common with both. To go further, we need a *new* physical model: quantum mechanics.

4. We saw that the electron interference pattern builds up over time, yet the points at which individual electrons hit the screen do not appear to be precisely predictable: they seem to arrive at random. It might seem natural to speculate that this apparent randomness might be explained by the fact that we don't have complete data about the experiment. Perhaps, for example, the electrons leave the

source in slightly different directions, or perhaps they have some sort of internal structure that we haven't yet discussed (and perhaps hasn't yet been discovered), and perhaps *this* determines precisely where they hit the screen.

As we'll see, according to quantum mechanics this is not the case. Quantum mechanics, unlike classical mechanics, is an *intrinsically probabilistic* theory, and it tells us that there is simply no way to predict precisely where the electron will hit the screen, *even if* we have a precise and complete description of its physical state when it leaves the source.

Now, of course, quantum mechanics might not be the final theory of nature. It's possible that some as yet unknown and more complete theory could underlie quantum mechanics, and it's logically possible that this theory (if there is one) could be deterministic. However, there are very strong reasons to doubt that any theory underlying quantum mechanics can be deterministic. In particular, it can be proved (given a few very natural assumptions) that any such theory would be inconsistent with special relativity. This follows from Bell's theorem and experimental tests thereof: it is discussed further in the Part II course "Principles of Quantum Mechanics" and in Part III courses.)

2.11 *Ongoing tests of quantum theory

Although later developments are beyond the scope of this course, it would be wrong to leave the impression that the historical development of quantum mechanics ended in the first part of last century. Indeed, the basic principles of quantum mechanics are still being questioned and tested by some theorists and experimentalists today. Quantum theory is very well confirmed as a theory of microscopic physics. However, the case for believing that it applies universally to physics on all scales is much weaker.

There is a good scientific motivation for testing any scientific theory in new domains, which is that theories developed to explain phenomena in one domain may not necessarily apply in other domains. For example, quantum theory itself shows us the limits of validity of Newtonian mechanics and of classical electrodynamics. Similarly, Einstein's general theory of relativity shows us the limits of validity of Newton's theory of gravity.

There is also a specific motivation for wanting to test quantum theory for macroscopic systems. This is that the problems in making sense of quantum theory as an *explanation* of natural phenomena seem to stem from the fact that the classical physics we see on macroscopic scales appears to emerge from quantum theory in a way that, despite many attempts, remains hard to pin down. Many theorists believe it remains fundamentally unexplained. Many others believe it is explained or explainable, but there is no real consensus among them about the right explanation.

Interestingly, we know there are consistent (non-relativistic) theories that agree very precisely with quantum mechanical predictions for microscopic (small mass) particles, but disagree for macroscopic (large mass) ones.

In the past few years, experimental technology has advanced far enough to demonstrate diffraction of quite complex molecules. (Some descriptions of these experiments can be found at

<https://vcq.quantum.at/>; see in particular the work of the Arndt and Aspelmayer research groups.) Attempts continue to demonstrate interferometry for larger and larger objects, motivated not only by the technological challenge but also by the possibility of testing the validity of quantum mechanics in new domains. In October 2019, Fein et al. reported interference experiments for molecules of weight larger than 25kDa (See <https://www.nature.com/articles/s41567-019-0663-9> and the rather amusing Q and A summary

<https://www.quantumnano.at/detailview-news/news/facts-fiction-in-reports-on-high-mass-interference/>).

It should be stressed that there is to date (October 2019) *no* experimental evidence for any deviation from quantum mechanics, which has been confirmed in an impressive array of experiments investigating many different physical regimes. *

2.12 Closing comments

1. As we've seen, the photon hypothesis played a key role in the development of quantum mechanics. We've also seen that photons (which are massless) and massive particles (such as electrons) produce qualitatively similar interference and diffraction patterns. However, we won't have much more to say about photons in this course. It turns out that we can develop quantum mechanics for the electron and other massive particles using relatively simple equations. We can build up a good intuition about how quantum systems behave in experiments and in nature from these equations. A fully satisfactory quantum treatment of photons or other massless particles requires a relativistic quantum theory of fields, which requires more sophisticated concepts (and is much harder to make mathematically rigorous). Quantum electrodynamics, which is a relativistic quantum field theory incorporating photons, is discussed in

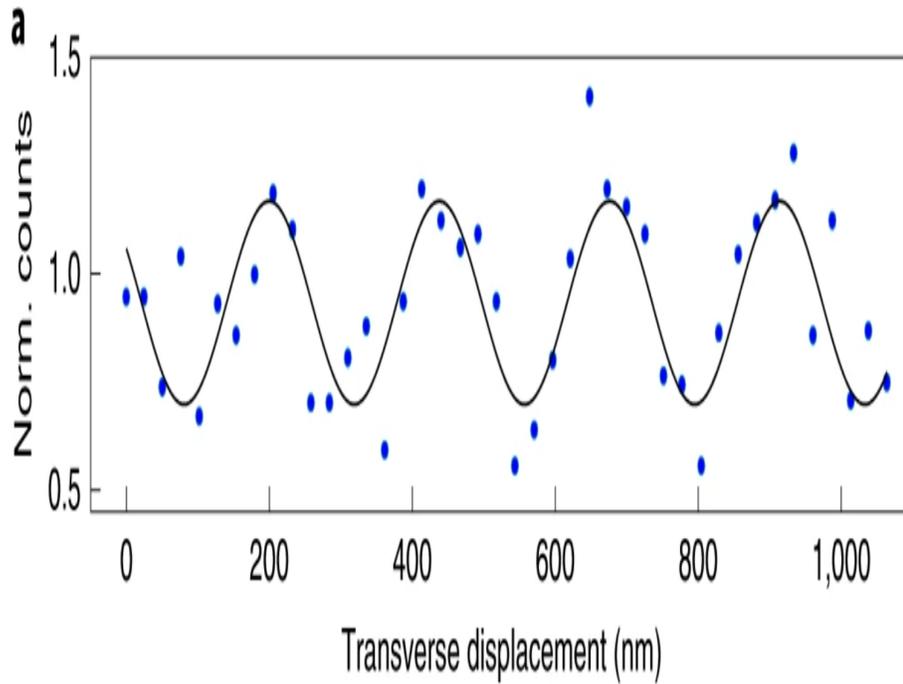


Figure 11: Interference of molecules larger than 25KDa, from the Fein et al. paper cited above.

Part III theoretical physics courses, along with other relevant quantum field theories.

2. Although we've already seen the classical particle model is inadequate, we still need some collective name for electrons, protons, neutrons and so on. Perhaps physicists should have invented another term, but we still call these "particles". We will follow this tradition, so that we might say that the electron is an elementary particle, talk about quantum mechanics applied to an abstract particle of mass m , and so forth — *always keeping in mind that the classical particle model does not actually apply.*

3 The one-dimensional Schrödinger equation

3.1 The 1D Schrödinger equation for a free particle

We are first going to develop quantum mechanics in one space (and one time) dimension. We can solve the equations for simple physical models more easily in 1D than in 3D and, happily, it turns out that 1D solutions give a good qualitative feel for a range of interesting 3D physical phenomena.

In 1924, Schrödinger developed de Broglie's ideas further, into what became a standard way of framing the laws of quantum mechanics.⁵

Recall that de Broglie postulated that matter is described by waves, and that the energy and momentum are related to the angular frequency and wave vector by $E = \hbar\omega$ and $\underline{p} = \hbar\mathbf{k}$, or in one dimension $p = \hbar k$. We can express this by associating to a particle of energy E and momentum p a wave of the form

$$\psi(x, t) = \exp(i(kx - \omega t)). \quad (3.1)$$

Now, for a mass m particle, we have $E = \frac{1}{2}mv^2 = \frac{p^2}{2m}$, and so

$$\psi(x, t) = \exp(i(kx - \omega t)) = \exp\left(\frac{i}{\hbar}\left(px - \frac{p^2}{2m}t\right)\right). \quad (3.2)$$

Notice that we have taken $\psi(x, t)$ to be complex. Using complex numbers to represent waves is familiar in classical electromagnetism, where it allows us to combine the electric and magnetic fields in a single equation. In that context, it's mathematically convenient, but the real and imaginary parts each have a simple physical interpretation. We will see that complex-valued solutions to (generalised) wave equations play a more essential role in quantum mechanics.

The simplest wave equation to which the de Broglie wave is the general solution is the *time-dependent 1D Schrödinger equation for a free particle*:

$$\frac{1}{2m}(-i\hbar\frac{\partial}{\partial x})^2\psi(x, t) = i\hbar\frac{\partial}{\partial t}\psi(x, t). \quad (3.3)$$

(By a *free particle* we mean a particle not subject to external forces, i.e. one moving in a potential $V(x) = 0$.)

⁵There is an equivalent alternative way of describing quantum mechanics, developed by Heisenberg, Born and Jordan. But Schrödinger's formulation is easier to work with and gives a more intuitive physical picture – so we will follow his approach. Note that “more intuitive” here is a relative statement. As we will see, many of the predictions of quantum mechanics are counter-intuitive. Also, some of the intuitions Schrödinger's picture suggests may be helpful to us in some contexts but are not fundamentally justifiable.

3.2 The momentum operator

We define the *momentum operator*

$$\hat{p} = -i\hbar \frac{\partial}{\partial x}. \quad (3.4)$$

So, for the de Broglie wave $\psi(x, t) = \exp(\frac{i}{\hbar}(px - \frac{p^2}{2m}t))$, we have $\hat{p}\psi = p\psi$. In other words, acting with the momentum operator on the de Broglie wave is equivalent to multiplying by the wave momentum.⁶ We can rewrite (3.3) as

$$\frac{1}{2m}\hat{p}^2\psi = i\hbar \frac{\partial \psi}{\partial t}. \quad (3.5)$$

This is our first example of a general feature of quantum mechanics. Physically significant quantities (in this case momentum) are represented by operators. These operators act on functions that represent physical states (in this case the idealized state defining the de Broglie wave).

Formally, we define an *operator* \hat{O} to be a *linear* map from a space of functions⁷ to itself, i.e. any map such that

$$\hat{O}(a_1\psi_1 + a_2\psi_2) = a_1\hat{O}\psi_1 + a_2\hat{O}\psi_2 \quad (3.6)$$

for all complex numbers a_1, a_2 and all functions ψ_1, ψ_2 in the relevant space.

⁶This tells us that the de Broglie wave is an *eigenfunction* of the momentum operator with *eigenvalue* p : we will define these terms more generally later.

⁷We will not be too precise about which space of functions we are working with, but will assume that they are suitably “well behaved”. For example, and depending on the context, we might want to consider the space of infinitely differentiable functions, $C^\infty(R)$, or the space of “square integrable” functions, i.e. those satisfying Eqn. (3.11).

3.3 The 1D Schrödinger equation for a particle in a potential

We want to consider particles subject to a potential $V(x)$ as well as free particles.

Examples: alpha rays scattering from a nucleus, electrons diffracting from a crystal, buckyballs going through an interferometer, neutrons or larger massive particles moving in a gravitational field.

To do this, we replace the kinetic energy term in (3.5) by an operator corresponding to the *hamiltonian* or *total energy*:

$$H = \frac{p^2}{2m} + V(x), \quad (3.7)$$

namely

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(x), \quad (3.8)$$

where the second operator corresponds to multiplication by $V(x)$.

This gives the general form of the *time-dependent 1D Schrödinger equation* for a single particle:

$$\hat{H}\psi(x, t) = i\hbar \frac{\partial \psi}{\partial t}. \quad (3.9)$$

Or, more explicitly:

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + V(x)\psi = i\hbar \frac{\partial \psi}{\partial t}. \quad (3.10)$$

Note that there is no way to *prove* mathematically that Eqns. (3.9, 3.10) are physically relevant, although we have given some motivation for them in the light of previous physical models and experimental results. As with any new physical theory, the only real test is experiment. Since it is not yet obvious what the complex-valued solutions to (3.9, 3.10) have to do with experimentally observable quantities, we will first need to give rules for interpreting them physically, and then test these predictions.

3.4 The wavefunction

We call a complex valued function $\psi(x, t)$ that is a solution to (3.9) or (3.10) a *wavefunction*. We say the wavefunction $\psi(x, t)$ is *normalisable* (at time t) if

$$0 < \int_{-\infty}^{\infty} |\psi(x, t)|^2 dx < \infty. \quad (3.11)$$

Note that for **any** complex valued $\psi(x, t)$ the integral is real and non-negative.

We say $\psi(x, t)$ is *normalised* if

$$\int_{-\infty}^{\infty} |\psi(x, t)|^2 dx = 1. \quad (3.12)$$

So, given a normalisable $\psi(x, t)$ with $\int_{-\infty}^{\infty} |\psi(x, t)|^2 dx = C$, so that $0 < C < \infty$, we can define a normalised wavefunction $C^{-1/2}\psi(x, t)$.

3.4.1 What is the wavefunction?

As we will explain in following lectures, the wavefunction $\psi(x, t)$ of a particle is a mathematical object that allows us to calculate the probability of detecting the particle at any given position if we set up a detector there. More generally, it allows us to calculate the probability of any given outcome for the measurement of any observable quantity (for example, energy or momentum) associated with the particle.

Sometimes in the course of your studies you may suspect that lecturers are temporarily keeping the full truth from you.⁸ Sometimes you would be right, but not here. We really don't have a better and more intuitively comprehensible story about the wavefunction.⁹

⁸If charitable, you may also suspect there may be good pedagogical reasons for this.

⁹At least, not one that is generally agreed.

3.4.2 What the wavefunction definitely isn't

Schrödinger initially hoped to interpret the wave function as describing a dispersed cloud of physical material that somehow corresponds to a “smeared-out” particle. This looks a natural interpretation at first sight, but proved untenable and was abandoned.

One problem with this interpretation is that if a charged particle is really a dispersed cloud of charge, we would expect to be able to detect bits of the cloud carrying fractions of the charge of the electron. However, we always find that charged objects carry a charge that is some integer multiple of the electron charge. Classical electrodynamics also suggests that a dispersed charge of cloud should interact repulsively with itself via the Coulomb force, and thus tend to be additionally dispersed, in a way that we do not observe.

Even if these objections could somehow be overcome, another problem remains. No matter how widely the electron's wavefunction is spread out in space, when we look for it by setting up detectors we always find an apparently pointlike particle in a definite location. If the wavefunction really represented a dispersed cloud, this cloud would have to suddenly coalesce into a particle at a single point when we carry out a measurement. This would mean that parts of the cloud would have to travel extremely fast — often much faster than light speed. This is inconsistent with special relativity.

3.5 The superposition principle

Exercises 1. The Schrödinger equation

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

is linear in $\psi(x, t)$: if ψ_1 and ψ_2 are solutions then $a_1\psi_1 + a_2\psi_2$ is a solution too, for any complex a_1, a_2 .

2. If ψ_1 and ψ_2 are normalisable and $a_1\psi_1 + a_2\psi_2$ is nonzero¹⁰, then it is also normalisable.

3. Show that it is not generally true that if ψ_1 and ψ_2 are normalised then $a_1\psi_1 + a_2\psi_2$ is, even if $|a_1|^2 + |a_2|^2 = 1$.

The linearity of the Schrödinger equation implies the so-called *superposition principle*: there is a physical solution corresponding to any linear combination of two (or more) physical solutions.

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.

We will see shortly that we need to normalise a wavefunction to obtain a sensible probability distribution from it and make physical predictions. So to make physical predictions from a superposition, we generally need to normalise the sum $\psi = a_1\psi_1 + a_2\psi_2$. As we've just seen, this is always possible if ψ_1, ψ_2 are normalisable and ψ is nonzero.

¹⁰I.e. not the zero *function*; it does not vanish everywhere.

3.6 Probabilistic interpretation of the wavefunction: the Born rule

Max Born in 1926 explained the essential connection between the wavefunction and experiment, via the so-called *Born rule*:

If we carry out an experiment to detect the position of a particle described by a normalised wavefunction $\psi(x, t)$, the probability of finding the particle in the interval $[x, x + dx]$ at time t is

$$\begin{aligned} \int_x^{x+dx} |\psi(y, t)|^2 dy &\approx dx |\psi(x, t)|^2 \\ &= dx \rho(x, t), \end{aligned} \tag{3.13}$$

where we write $\rho(x, t) = |\psi(x, t)|^2$ (see below).

More generally, the probability of finding the particle in any interval $[a, b]$ is

$$\int_a^b |\psi(y, t)|^2 dy. \tag{3.14}$$

Intuitively, it may seem that (3.14) should follow from (3.13). Certainly, it would be peculiar if the probability of finding a particle in a given interval depended on how the interval was sub-divided (i.e. on how precise our position measurements were). But we have already seen some apparently peculiar predictions of quantum mechanics, which show it is not safe to rely on intuition. We should rather understand (3.14) as a general postulate from which (3.13) follows as a special case. We will see later (section 6.6) that (3.14) itself is a special case of the general quantum measurement postulates, which apply to measurements of any physical quantity (not only position).

3.6.1 Probability density and probability current

The following mathematical quantities give very useful insights into the behaviour of solutions to the Schrödinger equation:

The *probability density*

$$\rho(x, t) = |\psi(x, t)|^2 \quad . \quad (3.15)$$

We see that the Born rule justifies the interpretation of $\rho(x, t)$ as a probability density. If we measure the position of the particle at time t , the probability of finding the particle in the interval $[x, x + dx]$ is $\rho(x, t)dx$.

The *probability current*

$$J(x, t) = -\frac{i\hbar}{2m} \left\{ \psi^*(x, t) \frac{\partial}{\partial x} \psi(x, t) - \left(\frac{\partial}{\partial x} \psi^*(x, t) \right) \psi(x, t) \right\}. \quad (3.16)$$

It is easy to verify from (3.10) that

$$\frac{\partial J}{\partial x} + \frac{\partial \rho}{\partial t} = 0. \quad (3.17)$$

Thus ρ and J do indeed satisfy a conservation equation, with ρ behaving as a density whose total integral is conserved, and J as a current describing the density flux.

The key point here is that $\frac{\partial \rho}{\partial t}$ can be written as a spatial derivative of some quantity. This means that we can calculate the time derivative of the probability of finding the particle in a region $[a, b]$:

$$\frac{d}{dt} \int_a^b |\psi(x, t)|^2 dx = \int_a^b -\frac{\partial}{\partial x} J(x, t) dx = J(a, t) - J(b, t). \quad (3.18)$$

The last term describes the probability density flux across the endpoints of the interval – the “net flow of probability out of (or in to) the interval”.

Now if ψ is normalised, i.e. Eqn (3.12) holds, then $\psi(x) \rightarrow 0$ as $x \rightarrow \pm\infty$. Thus $J(x) \rightarrow 0$ as $x \rightarrow \pm\infty$, assuming (as we will here) that $\frac{\partial}{\partial x} \psi(x)$ is bounded as $x \rightarrow \infty$. Thus

$$\frac{d}{dt} \int_{-\infty}^{\infty} |\psi(x, t)|^2 dx = \lim_{a \rightarrow -\infty} J(a, t) - \lim_{b \rightarrow \infty} J(b, t) = 0. \quad (3.19)$$

The total probability of finding the particle in $-\infty < x < \infty$ is thus constant over time: $\int_{-\infty}^{\infty} |\psi(y, t)|^2 dy = 1$ for any time t . So, the Born probabilistic interpretation is consistent: whenever we look for the particle, we will definitely find it somewhere, and only in one place.

Notes:

- We will consider measurements of quantities other than position later.
- The Born rule says nothing about where the particle is if we do not measure its position. According to the standard understanding of quantum mechanics, this is a question with no well-defined answer: the particle's position is not defined unless we measure it.

As we'll see, according to quantum mechanics, we can generally only calculate the probabilities of possible measurement results: we can't predict with certainty which result will occur. Moreover, the theory only allows us to predict probabilities for the possible results of measurements that actually take place in a given experiment. We cannot consistently combine these predictions with those for other measurements that could have been made had we done a different experiment instead.¹¹

We'll see when we discuss the general measurement postulates of quantum mechanics in section 6.6 that measuring the position of a particle generally changes its wavefunction. Recall the earlier discussion of the 2-slit experiment. We found no definite answer to the question "which slit did the particle go through?" – unless we put detectors beside the slits, and this changed the experiment *and changed the interference pattern*.

¹¹See again the analysis of the double slit experiment above, and (for example) the relevant chapters of Feynman's lecture notes, for further discussion of this.

4 Solutions of the 1D Schrödinger equation

We now look at various examples of solutions to the 1D SE for one particle:

4.1 Example I: Stationary states

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

Assume V is time-independent. We can then use the method of *separation of variables* to find solutions, by trying the ansatz

$$\psi(x, t) = \psi(x)T(t). \quad (4.2)$$

This gives us

$$(\hat{H}\psi(x))T(t) = \psi(x)(i\hbar \frac{\partial}{\partial t} T(t)), \quad (4.3)$$

and so

$$\frac{1}{\psi(x)} \hat{H}\psi(x) = \frac{1}{T(t)} (i\hbar \frac{\partial}{\partial t} T(t)). \quad (4.4)$$

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 . We thus have

$$T(t) = \exp\left(\frac{-i}{\hbar} Et\right). \quad (4.5)$$

and the *time-independent Schrödinger equation*

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) + V(x)\psi(x) = E\psi(x). \quad (4.6)$$

Solutions to the SE of the form

$$\psi(x) \exp\left(\frac{-i}{\hbar} Et\right),$$

where $\psi(x)$ is a solution to (4.6) are called *stationary states*.

We say $\psi(x)$ is an *eigenfunction* of the hamiltonian operator \hat{H} which corresponds to a physical state whose energy is given by the *eigenvalue* E . When $V = 0$, (4.6) reduces to

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) = E\psi(x). \quad (4.7)$$

More generally, if

$$A\psi(x) = a\psi(x) \quad (4.8)$$

for some operator A and complex number a , then we say $\psi(x)$ is an *eigenfunction* of A with *eigenvalue* a . The terminology is a natural extension to infinite-dimensional spaces (of functions) of the definitions of eigenvector and eigenvalue for finite-dimensional matrices: compare the eigenequation $Av = av$ for matrix A , vector v and complex number a .

As we will see, the eigenvalues of physically significant operators are observable physical quantities. They turn out always to be real – as we would expect, since we know of no natural way to directly observe a complex number in nature. We will see the mathematical reason for this when we discuss the general formalism of quantum mechanics.

Notes

- Equation (4.7) has solution $e^{i(kx-\omega t)}$, where $E = \frac{\hbar^2 k^2}{2m} = \hbar\omega$. This is the de Broglie wave we originally considered, so our discussion is at least self-consistent.
- This de Broglie wave solution to the 1D SE for a free particle is *not* normalisable, and so does not have a well-defined probabilistic interpretation via the Born rule. The quantum mechanical free particle is a very useful idealisation, but not a physical solution. We never actually have a uniformly zero potential throughout space, nor a single particle whose wavefunction is spread uniformly over all of space. There could not be any consistent way of assigning probabilities to finding such a particle in finite regions, since there is no well-defined uniform probability distribution on the 1D real line. (The same is true in 3D, of course.)
- Any stationary state $\psi(x, t) = \psi(x)e^{-\frac{iEt}{\hbar}}$ has probability density

$$\rho(x, t) = |\psi(x, t)|^2 = |\psi(x)|^2.$$

In other words, its probability density is time-independent. More generally, we will see the probability of measuring any given outcome for any dynamical quantity is time-independent. Hence the name *stationary* state.

4.2 Completeness of the energy eigenfunctions

An important fact is that the general solution to the SE with t -independent potential $V(x)$ is a *superposition* of stationary states:

$$\psi(x, t) = \sum_{n=1}^{\infty} a_n \psi_n(x) e^{-\frac{iE_n t}{\hbar}}. \quad (4.9)$$

This follows from the fact that any wavefunction at a given time, say $t = 0$, can be written as a superposition of energy eigenfunctions, so that,

$$\psi(x, 0) = \sum_{n=1}^{\infty} a_n \psi_n(x), \quad (4.10)$$

and from the linearity of the SE.

We will see later (from theorem 6.3) that Eqn. (4.10) is a special case of the more general result that any wavefunction can be written as a superposition of eigenfunctions for any operator corresponding to a physically observable quantity.

If the set of energy eigenvalues is continuous, we need to replace the sum by an integral. If there are continuous and discrete subsets, we need both sum and integral. Thus if there is a discrete set of normalised energy eigenfunctions $\{\psi_i\}_{i=1}^N$ (we include $N = \infty$ as a possibility) and a continuous set of eigenfunctions $\{\psi_\alpha\}_{\alpha \in \Delta}$, where the indexing set Δ is some sub-interval of the real numbers, we can write

$$\psi(x, t) = \sum_{n=1}^N a_n \psi_n(x) e^{-\frac{iE_n t}{\hbar}} + \int_{\Delta} a_\alpha \psi_\alpha(x) e^{-\frac{iE_\alpha t}{\hbar}} d\alpha, \quad (4.11)$$

4.3 Example II: Gaussian wavepackets

As an example of a continuous superposition of free particle stationary states, consider

$$\psi(x, t) = C \int_{-\infty}^{\infty} e^{-\frac{\sigma}{2}(k-k_0)^2} e^{i(kx - \frac{\hbar k^2}{2m}t)} dk. \quad (4.12)$$

The exponent here is $-\frac{1}{2}(\sigma + \frac{i\hbar t}{m})k^2 + (k_0\sigma + ix)k - \frac{\sigma}{2}k_0^2$, which we can write in the form $-\frac{1}{2}a(k - \frac{b}{a})^2 + (\frac{b^2}{2a} - \frac{\sigma}{2}k_0^2)$, where $a = \sigma + \frac{i\hbar t}{m}$, $b = k_0\sigma + ix$.

Now we have

$$\int e^{-\frac{1}{2}a(k - \frac{b}{a})^2} dk = \sqrt{\frac{2\pi}{a}}. \quad (4.13)$$

This gives us

$$\psi(x, t) = C' \sqrt{\frac{2\pi}{\sigma + \frac{i\hbar t}{m}}} \exp\left(\frac{(k_0\sigma + ix)^2}{2(\sigma + \frac{i\hbar t}{m})}\right) \quad (4.14)$$

$$\begin{aligned} |\psi(x, t)|^2 &= C'' \sqrt{\frac{1}{\sigma^2 + \frac{\hbar^2 t^2}{m^2}}} \exp\left(\frac{(k_0\sigma + ix)^2}{2(\sigma + \frac{i\hbar t}{m})} + \frac{(k_0\sigma - ix)^2}{2(\sigma - \frac{i\hbar t}{m})}\right) \\ &= C'' \sqrt{\frac{1}{\sigma^2 + \frac{\hbar^2 t^2}{m^2}}} \exp\left(\frac{\sigma^3 k_0^2 - \sigma x^2 + \frac{2k_0\sigma x \hbar t}{m}}{\sigma^2 + \frac{\hbar^2 t^2}{m^2}}\right) \\ &= C''' \sqrt{\frac{1}{\sigma^2 + \frac{\hbar^2 t^2}{m^2}}} \exp\left(\frac{-\sigma(x - \frac{k_0 \hbar t}{m})^2}{\sigma^2 + \frac{\hbar^2 t^2}{m^2}}\right), \quad (4.15) \end{aligned}$$

where C' , C'' and C''' are constants.

Since

$$\begin{aligned} \int_{-\infty}^{\infty} |\psi(x, t)|^2 dx &= C''' \sqrt{\frac{2(\sigma^2 + \frac{\hbar^2 t^2}{m^2})\pi}{2\sigma}} \sqrt{\frac{1}{(\sigma^2 + \frac{\hbar^2 t^2}{m^2})}} \\ &= C''' \sqrt{\frac{\pi}{\sigma}} \quad (4.16) \end{aligned}$$

the normalisation condition gives us

$$C''' = \sqrt{\frac{\sigma}{\pi}}. \quad (4.17)$$

Notes:

- The Gaussian wavepacket describes an *approximately localised* particle, almost all of whose probability density is within a finite region of size $\approx N(\sigma^2 +$

$\frac{\hbar^2 t^2}{m^2})^{\frac{1}{2}} \sigma^{-\frac{1}{2}}$ at time t , for some smallish positive number N .¹² This follows since the probability density function is a Gaussian curve with standard deviation $(\sigma^2 + \frac{\hbar^2 t^2}{m^2})^{\frac{1}{2}} \sigma^{-\frac{1}{2}}$.

- Recall that we obtained the Gaussian wavepacket as a superposition of stationary states. Considering its evolution, we thus see that a superposition of stationary states is *not* necessarily stationary, since the Gaussian wavepacket's probability density function varies with time.
- From Eqn. (4.15), we see the speed of the Gaussian wavepacket's centre is $\approx \frac{\hbar k_0}{m}$. Since its standard deviation is

$$(\sigma^2 + \frac{\hbar^2 t^2}{m^2})^{\frac{1}{2}} \sigma^{-\frac{1}{2}} \approx \frac{\hbar t}{m \sigma^{\frac{1}{2}}} \quad (4.18)$$

for large t , the speed at which it spreads in width is $\approx \frac{\hbar}{m \sigma^{\frac{1}{2}}}$. Note that, mathematically speaking, neither is necessarily bounded by the speed of light c . This illustrates that the Schrödinger equation is *non-relativistic* – i.e. incompatible with special relativity. If it was possible to produce Gaussian wavepackets for any values of k_0 and σ , and if the Schrödinger equation precisely described their evolution, we could use the wavepackets to send signals faster than light.¹³

As this suggests, the SE is only approximately valid.¹⁴ We cannot use Gaussian wavepackets (or, unless we find experimental evidence against special relativity, anything else) to send signals faster than light. However, the SE is a good enough approximation to allow us to understand a great deal of physics and chemistry.

- Pictures of an evolving Gaussian wave packet can be found at (for example) <http://demonstrations.wolfram.com/EvolutionOfAGaussianWavePacket/>

4.4 Example III: Particle in an infinite square potential well

Consider a particle confined by a 1D box with impermeable walls. We can model this by a potential

$$V(x) = \begin{cases} 0 & |x| < a, \\ \infty & |x| > a. \end{cases} \quad (4.19)$$

We look for stationary states obeying

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) + V(x) \psi(x) = E \psi(x), \quad (4.20)$$

¹²For example, 99.7% of the probability density is within the region defined by $N = 3$.

¹³We can also see that the SE is non-relativistic by noting that it is second order in x and first order in t . A relativistically invariant equation must be of the same order in x and t , because Lorentz transformations map x and t to linear combinations.

¹⁴Quantum mechanics, which is the topic of this course, is a non-relativistic theory, which is approximately correct when relativistic corrections are small. Physicists tend to think of quantum theory as defining a mathematical framework, and non-relativistic quantum mechanics as a particular theory within this framework. The framework of quantum theory also includes relativistic quantum field theories, which (as the name suggests) *are* consistent with special relativity. In particular, there is a relativistically invariant equation for the evolution of an electron's wavefunction, from which the SE emerges as a non-relativistic limit. We won't worry about relativistic corrections or extensions of quantum mechanics in this course; they are discussed in Part III.

which gives us

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) = \begin{cases} E\psi(x) & |x| < a, \\ (E - \infty)\psi(x) & |x| > a. \end{cases} \quad (4.21)$$

Here the second alternative is a useful but unrigorous shorthand, and should perhaps be written in quotation marks to emphasize this. The next paragraph explains how we handle the infinity.

4.4.1 Handling discontinuities

We need a consistent treatment of discontinuities in $V(x)$ and (hence perhaps) in $\psi(x)$, $\frac{d\psi}{dx}$, \dots . One method is to take a discontinuous $V(x)$ as the limiting case of a set of continuous potentials, and consider the limits of the solutions $\psi(x)$. But we can address the SE with discontinuous $V(x)$ directly using the following rules:

A step function discontinuity in ψ would create a delta function derivative in $\frac{d^2\psi}{dx^2}$ on the LHS of the SE, which would not be cancelled by other terms. So we expect ψ to be continuous everywhere.

A step function discontinuity in $\frac{d\psi}{dx}$ would create a delta function in $\frac{d^2\psi}{dx^2}$ on the LHS of the SE; we can allow this only at points where $V(x)$ is infinitely discontinuous.

A step function discontinuity in $\frac{d^2\psi}{dx^2}$ gives a step function discontinuity on the LHS of the SE; we can allow this only at points where $V(x)$ has a finite step function discontinuity so that the two discontinuities on the LHS cancel.

So, for the infinite square well we expect ψ to be continuous everywhere, and $\frac{d\psi}{dx}$ to be continuous except at $|x| = a$. Since $V = \infty$ for $|x| > a$, we need $\psi = 0$ in this region for consistency of the SE. We have

$$\psi(x) = A \cos(kx) + B \sin(kx), \quad k = \sqrt{\frac{2mE}{\hbar^2}} \quad \text{in } |x| < a. \quad (4.22)$$

Now $\psi = 0$ at $x = \pm a$ gives

$$A \cos(ka) \pm B \sin(ka) = 0 \quad (4.23)$$

and hence

$$A \cos(ka) = 0 \text{ and } B \sin(ka) = 0. \quad (4.24)$$

Hence

$$\begin{aligned}
 \text{either } A = 0, & & \sin(ka) = 0, \\
 \psi(x) = B \sin(kx), & & k = \frac{n\pi}{2a} \text{ for } n \geq 2 \text{ even.} \\
 \text{or } B = 0, & & \cos(ka) = 0, \\
 \psi(x) = A \cos(kx), & & k = \frac{n\pi}{2a} \text{ for } n \geq 1 \text{ odd.}
 \end{aligned}
 \tag{4.25}$$

This gives us solutions for

$$E = E_n = \frac{\hbar^2 \pi^2 n^2}{8ma^2} \text{ for } n = 1, 2, 3, \dots
 \tag{4.26}$$

Notes:

- We have *quantised* (i.e. discrete) energies for the stationary states of the infinite square well. Cf. the Bohr atom – it begins to look plausible that quantum mechanics could explain the apparent quantisation of physical states. Perhaps the allowed electron orbits in the Bohr atom correspond to the stationary states in a 3D Coulomb potential?¹⁵
- The solutions fall into two classes:

$$\psi(x) = \begin{cases} \psi(-x) & \text{even parity,} \\ -\psi(-x) & \text{odd parity.} \end{cases}
 \tag{4.27}$$

¹⁵We will see later that the relationship is a bit more complicated, but the intuition is along the right lines.

4.5 Parity

Let us consider this last phenomenon more generally. Suppose that $V(x) = V(-x)$ and that $\psi(x)$ is a solution of the time-independent SE. We have

$$\begin{aligned} -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) + V(x)\psi(x) &= E\psi(x), \\ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(-x) + V(x)\psi(-x) &= E\psi(-x), \end{aligned}$$

obtaining the second equation by substituting $y = -x$ in the first. Hence $\psi(-x)$ is *also* a solution of the time-independent SE.

There are two possibilities:

- (i) $\psi(x)$ and $\psi(-x)$ are linearly independent. Then $\psi(x) + \psi(-x)$ and $\psi(x) - \psi(-x)$ are also linearly independent, and have parity +1 and -1 respectively.
- (ii) $\psi(x)$ and $\psi(-x)$ are linearly dependent: $\psi(-x) = A\psi(x)$. Then

$$\psi(x) = A\psi(-x) = A^2\psi(x) \quad (4.28)$$

so we have

$$A = \pm 1. \quad (4.29)$$

$A = 1$ implies $\psi(x) = \psi(-x)$, so ψ has even parity. $A = -1$ implies $\psi(x) = -\psi(-x)$, so ψ has odd parity.

Conclusion Any solution of the t-independent SE with a symmetric potential is a linear combination of solutions of definite parity.¹⁶

4.6 Example IV: Particle in a finite square well

A slightly more realistic model of a confined particle is given by the finite square well potential

$$V(x) = \begin{cases} -U & |x| < a, \\ 0 & |x| > a. \end{cases} \quad (4.30)$$

– a crude model of the Coulomb potential of a charged nucleus. We will look for *bound state* solutions – normalisable solutions – as opposed to unnormalisable *scattering solutions*, whose physical interpretation we will discuss later.

A solution of energy E obeys

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) = \begin{cases} (E + U)\psi(x) & |x| < a, \\ E\psi(x) & |x| > a. \end{cases} \quad (4.31)$$

If $E > 0$, the solution for $|x| > a$ takes the form $\psi = Ae^{ikx} + Be^{-ikx}$, where $k^2 = \frac{2mE}{\hbar}$. This is unnormalisable.

¹⁶One can also prove that the ground state (lowest energy solution), if one exists, is unique and has even parity. In fact, one can prove the stronger result that the ground state wavefunction is unique, real and strictly positive, which implies that it must have even parity. See for example, volume 4 of Reed and Simon, “Methods of Modern Mathematical Physics”, Theorem XIII.4.6. A proof is also given in Eugene Lim’s lecture notes for an earlier version of this course, available online.

So we require $E < 0$. We will assume that $E > -U$, i.e. that $|E| < U$, since we do not expect there to be solutions with energy lower than the minimum potential.

Exercise: Verify that indeed there are no such solutions.

Let

$$k^2 = \frac{2m|E|}{\hbar^2}, \quad (4.32)$$

$$l^2 = \frac{2m(U - |E|)}{\hbar^2}, \quad (4.33)$$

$$\text{so that } k^2 + l^2 = \frac{2mU}{\hbar^2}. \quad (4.34)$$

For $|x| > a$ we have

$$\psi(x) = Ae^{-kx} + Be^{kx}, \quad (4.35)$$

and normalisability implies that

$$\psi(x) = \begin{cases} Ae^{-kx} & x > a, \\ Be^{kx} & x < -a. \end{cases} \quad (4.36)$$

For $|x| < a$ we have

$$\psi(x) = C \cos(lx) + D \sin(lx). \quad (4.37)$$

Since we know that the states of definite parity span the space of all bound states, we can solve separately for even and odd parity states to get a complete spanning set of solutions.

Consider the even parity solutions. If $\psi(x) = \psi(-x)$ we have:

$$\psi(x) = \begin{cases} Ae^{-kx} & x > a, \\ Ae^{kx} & x < -a, \\ C \cos(lx) & |x| < a. \end{cases} \quad (4.38)$$

$$\psi'(x) = \begin{cases} -kAe^{-kx} & x > a, \\ kAe^{kx} & x < -a, \\ -lC \sin(lx) & |x| < a. \end{cases} \quad (4.39)$$

As ψ and ψ' are continuous at $x = \pm a$, we find

$$C \cos(la) = A \exp(-ka), \quad (4.40)$$

$$-lC \sin(la) = -kA \exp(-ka). \quad (4.41)$$

This gives

$$k = l \tan(la)$$

or

$$(ka) = (la) \tan(la) \quad (4.42)$$

and from (4.34) we have

$$(ka)^2 + (la)^2 = \frac{2mUa^2}{\hbar^2}. \quad (4.43)$$

We can solve these last two equations graphically.

The odd parity bound states (if any exist) can be similarly obtained.

The finite square well illustrates a general feature of quantum potentials $V(x)$ such that $V(x) \rightarrow 0$ as $x \rightarrow \pm\infty$. The time-independent SE has two linearly independent solutions.

For $E > 0$, these take the form of scattering solutions

$$\begin{aligned} \psi(x) &= A \exp(ikx) + B \exp(-ikx) \text{ as } x \rightarrow -\infty \\ &C(A, B) \exp(ikx) + D(A, B) \exp(-ikx) \text{ as } x \rightarrow \infty. \end{aligned}$$

For $E < 0$, they have the form

$$\begin{aligned} \psi(x) &= A \exp(kx) + B \exp(-kx) \text{ as } x \rightarrow -\infty \\ &C(A, B) \exp(kx) + D(A, B) \exp(-kx) \text{ as } x \rightarrow \infty. \end{aligned}$$

Here $C(A, B)$ and $D(A, B)$ are linear in A and B , and depend also on k and on the parameters defining $V(x)$.

Only for special values of $E < 0$ can we find solutions such that $B = 0$ and $C(A, B) = 0$, as required for normalisability. For the remaining values of $E < 0$, the solutions blow up exponentially as $x \rightarrow -\infty$ or as $x \rightarrow \infty$ (or both) and are unphysical: neither bound states nor scattering solutions.

[These asymptotically exponential functions are not physically meaningful: they are not normalisable and do not define eigenfunctions of \hat{H} .]

4.7 Example V: The quantum harmonic oscillator

has potential

$$V(x) = \frac{1}{2}m\omega^2x^2. \quad (4.44)$$

This is a particularly important example for several reasons.

First, it is not only solvable but (as we will see later) has a very elegant solution that explains some important features of quantum theory.

Second, generic symmetric potentials in 1D are at least approximately described by Eqn. (4.44). To see this, consider a potential $V(x)$ taking a minimum value V_{\min} at x_{\min} and symmetric about x_{\min} . Writing $y = x - x_{\min}$, we have

$$\begin{aligned} V(y) &= V(0) + yV'(0) + \frac{y^2}{2!}V''(0) + \frac{y^3}{3!}V'''(0) + \dots \\ &\approx V_{\min} + \frac{y^2}{2!}V''(0) + O(y^4). \end{aligned}$$

Renormalising the potential so that $V_{\min} = 0$, and taking $V''(0) = m\omega^2$, we recover (4.44). (If $V''(0) = 0$ this argument does not hold, but even in this case the harmonic oscillator potential can be a good qualitative fit. See for example comments below on approximating a finite square well potential by a harmonic oscillator potential.)

In fact, we can show something more general. Suppose we have a system with n degrees of freedom that has a unique stable equilibrium. Then the small oscillations about equilibrium can be approximated by n independent harmonic oscillators. This means that solving the harmonic oscillator allows us to give a quantum description of molecules excited by radiation (and hence understand the greenhouse effect and many other phenomena) and of the behaviour of crystals and other solids.

Third, the harmonic oscillator plays an even more fundamental role in quantum field theory, where we understand particles as essentially harmonic oscillator excitations of fields. Our entire understanding of matter and radiation is thus based on the quantum harmonic oscillator!

To obtain the stationary states of the quantum harmonic oscillator we need to solve

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) + \frac{1}{2}m\omega^2x^2\psi(x) = E\psi(x). \quad (4.45)$$

We define

$$y = \sqrt{\frac{m\omega}{\hbar}}x \text{ and } \alpha = \frac{2E}{\hbar\omega}. \quad (4.46)$$

We have

$$\frac{d^2\psi}{dy^2} - y^2\psi = -\alpha\psi. \quad (4.47)$$

If $\psi \sim \exp(-\frac{1}{2}y^2)$, then $\psi' \sim -y \exp(-\frac{1}{2}y^2)$ and $\psi'' \sim y^2 \exp(-\frac{1}{2}y^2)$ to leading order, which means that $\psi'' - y^2\psi \sim (y^2 - y^2) \exp(-\frac{1}{2}y^2)$ satisfies (4.47) to leading order. Similarly $\psi \sim y^n \exp(-\frac{1}{2}y^2)$ satisfies (4.47) to leading order. This suggests considering power series solutions of the form

$$\psi(y) = H(y)e^{-\frac{y^2}{2}}, \quad (4.48)$$

$$H(y) = \sum_{m=0}^{\infty} a_m y^m, \quad (4.49)$$

and trying to match the coefficients a_m to produce an exact solution.

From (4.47) we get

$$\psi'' + (\alpha - y^2)\psi = 0 \quad (4.50)$$

and substituting from (4.48, 4.50) we get

$$\frac{d^2}{dy^2}(H \exp(-\frac{y^2}{2})) + (\alpha - y^2)H \exp(-\frac{y^2}{2}) = 0 \quad (4.51)$$

$$(H'' - 2yH' + Hy^2 - H + (\alpha - y^2)H) \exp(-\frac{y^2}{2}) = 0 \quad (4.52)$$

$$(H'' - 2yH' + (\alpha - 1)H) \exp(-\frac{y^2}{2}) = 0 \quad (4.53)$$

We have

$$H'(y) = \sum_{m=0}^{\infty} m a_m y^{m-1}, \quad (4.54)$$

$$\begin{aligned} H''(y) &= \sum_{m=0}^{\infty} m(m-1) a_m y^{m-2} \\ &= \sum_{m=0}^{\infty} a_{m+2} (m+2)(m+1) y^m. \end{aligned} \quad (4.55)$$

So from (4.53) we get

$$\sum_{m=0}^{\infty} (a_{m+2} (m+2)(m+1) - a_m (2m+1-\alpha)) y^m = 0. \quad (4.56)$$

This must be true for all y , so each coefficient of y^m must vanish:

$$a_{m+2} = \frac{(2m+1-\alpha)}{(m+2)(m+1)} a_m \text{ for } m = 0, 1, 2, \dots \quad (4.57)$$

which determines the a_m for $m \geq 2$ in terms of a_0 and a_1 .

For large m , (4.57) gives $a_{m+2} \approx \frac{2}{m} a_m$, which asymptotically describes the coefficients of e^{y^2} :

$$e^{y^2} = \sum_{n=0}^{\infty} \frac{1}{n!} y^{2n} = \sum_{n=0}^{\infty} b_n y^n \quad (4.58)$$

where for n odd we have $b_n = 0$ and for n even we have

$$b_{n+2} = \frac{1}{\frac{n}{2} + 1} b_n = \frac{2}{n + 2} b_n \approx \frac{2}{n} b_n. \quad (4.59)$$

So, if the power series is infinite, we expect

$$H(y) \approx C e^{y^2}$$

for some nonzero constant C : i.e. that $\frac{H(y)}{C e^{y^2}} \rightarrow 1$ as $y \rightarrow \infty$. This implies that

$$\psi(y) \approx H(y) e^{-\frac{y^2}{2}} \approx C e^{\frac{1}{2} y^2}, \quad (4.60)$$

which is not normalisable.

We can justify this rigorously by noting that, for any $\epsilon > 0$, there exists some integer m_0 such that for all $m \geq m_0$ we have

$$\frac{a_{m+2}}{a_m} > \frac{2(1 - \epsilon)}{m}. \quad (4.61)$$

Hence we have that

$$H(y) > C e^{(1-\epsilon)y^2} - P(y), \quad (4.62)$$

where $C = a_{m_0} \neq 0$ and $P(y)$ is a polynomial of degree $\leq m_0$. This implies that there exists a y_0 such that for all $y \geq y_0$ we have

$$|H(y)| > \frac{C}{2} e^{(1-\epsilon)y^2}, \quad (4.63)$$

and hence

$$|\psi(y)| > \frac{C}{2} e^{(\frac{1}{2}-\epsilon)y^2}. \quad (4.64)$$

So indeed $\psi(y)$ diverges as $y \rightarrow \infty$, and in particular ψ is not normalisable.

We thus need the power series to be truncated to a polynomial in order to find normalisable physical solutions. Again, we can consider fixed (even and odd) parity solutions separately:

- Even parity: $a_0 \neq 0, a_1 = 0, \alpha = 2m + 1$ for some even m .
- Odd parity: $a_0 = 0, a_1 \neq 0, \alpha = 2m + 1$ for some odd m .

So the physical solutions are given by

$$\alpha = \frac{2E}{\hbar\omega} = 2m + 1 \quad (m = 0, 1, 2, \dots) \quad (4.65)$$

i.e.

$$E = \left(m + \frac{1}{2}\right)\hbar\omega \quad (m = 0, 1, 2, \dots). \quad (4.66)$$

The harmonic oscillator has energy levels equally spaced, separated by $\hbar\omega$, with minimum (or *zero-point*) energy $\frac{1}{2}\hbar\omega$.

The polynomials $H_n(y)$ corresponding to the physical solutions with energy $(n + \frac{1}{2})\hbar\omega$ are the *Hermite polynomials* of degree n . We can obtain them explicitly¹⁷ from (4.65), using conventional normalisations:

$$H_0 = 1, H_1 = 2y, H_2 = 4y^2 - 2, H_3 = 8y^3 - 12y, \dots \quad (4.67)$$

The corresponding wavefunctions are thus (up to normalisation)

$$\psi_0 = e^{-\frac{1}{2}y^2}, \psi_1 = 2ye^{-\frac{1}{2}y^2}, \dots \quad (4.68)$$

¹⁷In fact one can derive a general expression (see e.g. Schiff, “Quantum Mechanics”, 3rd edition):

$$H_n(y) = (-1)^n e^{y^2} \frac{d^n}{dy^n} e^{-y^2}.$$

5 Tunnelling and Scattering

Let us reconsider the bound state solutions for the finite square well potential:

$$\psi(x) \sim \begin{cases} e^{-kx} & x > a, \\ e^{kx} & x < -a, \\ \cos(lx) \text{ or } \sin(lx) & |x| < a. \end{cases} \quad (5.1)$$

Although ψ tends to zero rapidly outside the well, it *is* non-zero there. In particular

$$\int_{|x|>a} |\psi(x)|^2 dx > 0 \quad (5.2)$$

for a bound state solution, although the energy $E = -\frac{\hbar^2 k^2}{2m} < 0$.

The Born rule (3.13) thus tells us that the probability of finding the particle outside the well is *non-zero*, even though the particle's energy is less than the potential height in this region ($V = 0$ in $|x| > a$ by our convention).

Clearly, a classical particle with $E < 0$ would never make its way outside the well. We have here a suggestive argument for the existence of an intrinsically quantum phenomenon – the possibility of *tunnelling* through a potential barrier into a classically forbidden state.

However, the physical interpretation of this calculation is complicated by the fact that we can only assign meaning in quantum mechanics to things we can detect, and we do not have any way to detect negative energy particles. So, we cannot actually observe a bound state particle outside the region $|x| < a$ unless we alter the potential.

What we can and do observe are quantum particles tunnelling through *finite* width potential barriers through which a classical particle of the same energy could not pass.

5.1 Scattering states

To understand tunnelling we need a general treatment of *scattering*: the transmission or reflection of particles by potential barriers (or potential wells).

In principle, we could take as initial state an incoming wavepacket (e.g. an approximately Gaussian packet) far from the potential barrier, solve the time-dependent Schrödinger equation, and obtain a solution that at asymptotic late times takes the form of a superposition of two wavepackets, one reflected by the barrier and one transmitted through it:

$$\psi \sim \psi_r + \psi_t. \quad (5.3)$$

We could then calculate the *reflection probability*

$$|\psi_r|^2 \approx \int_{-\infty}^0 |\psi(x, t)|^2 dx \text{ for large } t \quad (5.4)$$

and *transmission probability*

$$|\psi_t|^2 \approx \int_0^{\infty} |\psi(x, t)|^2 dx \text{ for large } t. \quad (5.5)$$

Some pictures of this process can be found in Schiff (3rd edition, pp 106-9), and in Brandt and Dahmen, “Picture Book of Quantum Mechanics”.

However, studying wavepacket scattering directly is mathematically **complicated**. We can model it computationally, and this is often helpful and illuminating, but it is not so easy to prove simple analytic results.

A simpler way to obtain scattering probabilities is to consider unnormalised stationary state solutions that asymptotically take the form of superpositions of plane waves as $x \rightarrow \pm\infty$. Thus:

$$\begin{aligned}\psi &= \exp(ikx) + R \exp(-ikx) \text{ as } x \rightarrow -\infty \\ \psi &= T \exp(ik'x) \text{ as } x \rightarrow \infty.\end{aligned}$$

If $V = 0$ as $x \rightarrow \pm\infty$ then $k = k'$. We are also interested in examples where V has different limits as $x \rightarrow \pm\infty$, so allow $k' \neq k$.

Here we follow the convention that incoming plane waves arrive from $-\infty$ but not ∞ . Now $\exp(ikx - \omega t)$ is rightward travelling, and $\exp(-ikx - \omega t)$ is leftward travelling. We consider a stationary solution at (for simplicity) $t = 0$. We thus allow a component of the rightward travelling incoming wave $\exp(ikx)$ as $x \rightarrow -\infty$ but do not allow any component of a leftward travelling wave $\exp(-ik'x)$ as $x \rightarrow \infty$, which would represent a wave incoming from ∞ .

We expect in general a *reflected plane wave* of form

$$R \exp(-ikx) \text{ as } x \rightarrow -\infty,$$

, and a *transmitted plane wave* of form

$$T \exp(ik'x) \text{ as } x \rightarrow \infty.$$

5.2 Interpretation of plane wave scattering solutions

We can justify taking these unnormalised stationary state solutions as representing:

- (i) an approximation to the behaviour of a 1-particle incoming wavepacket with wave vector sharply peaked about k .
- (ii) an approximation to the behaviour of a **beam** of particles (whose interactions with one another are negligible) with approximate wave vector k .
- (iii) a mathematical calculation of the behaviour of the wave vector k Fourier component of (i) or (ii).

The beam picture is perhaps the most intuitive, and we will use it, considering beam scattering from various potentials. We can interpret $|R|^2$ as the density of particles in the reflected beam, $|T|^2$ as the density in the transmitted beam. Recall that we have $p = \hbar k$ and so the speed $v = \frac{p}{m} = \frac{\hbar k}{m}$. The particles in the incoming, reflected and transmitted beams thus have speeds

$$\frac{\hbar k}{m}, \quad -\frac{\hbar k}{m}, \quad \frac{\hbar k'}{m} \tag{5.6}$$

respectively.

We define the *particle flux* in the beams to be the number of particles per second in the beam passing a fixed point. We have

$$\text{flux} = \text{velocity} \times \text{density} = \begin{cases} \frac{\hbar k}{m} & \text{incoming,} \\ \frac{\hbar k}{m} |R|^2 & \text{reflected,} \\ \frac{\hbar k'}{m} |T|^2 & \text{transmitted.} \end{cases} \quad (5.7)$$

Particle conservation – no particles are destroyed or created in the scattering process – thus implies that

$$\frac{\hbar k}{m} = \frac{\hbar k}{m} |R|^2 + \frac{\hbar k'}{m} |T|^2. \quad (5.8)$$

We will verify this in particular examples.

5.3 Example I: The potential step

Consider the potential

$$V(x) = \begin{cases} 0 & x < 0, \\ U & x > 0. \end{cases} \quad (5.9)$$

A solution of energy E obeys

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) = \begin{cases} (E - U)\psi(x) & x > 0, \\ E\psi(x) & x < 0. \end{cases} \quad (5.10)$$

For the moment we consider $U > 0$ (a step rather than a drop).

We want solutions of the form $\exp(ikx) + R \exp(-ikx)$ for $x < 0$; these have $E = \frac{\hbar^2 k^2}{2m} > 0$.

Case 1: $E < U$ Define

$$k = \frac{\sqrt{2mE}}{\hbar}, \quad l = \frac{\sqrt{2m(U - E)}}{\hbar}. \quad (5.11)$$

We have

$$\frac{d^2}{dx^2} \psi(x) = \begin{cases} l^2 \psi(x) & x > 0, \\ -k^2 \psi(x) & x < 0. \end{cases} \quad (5.12)$$

So

$$\psi(x) = \begin{cases} A \exp(-lx) & x > 0, \\ \exp(ikx) + R \exp(-ikx) & x < 0. \end{cases} \quad (5.13)$$

We see there is no transmitted plane wave in this case.

Continuity of ψ and ψ' at $x = 0$ gives

$$1 + R = A \quad (5.14)$$

$$ik(1 - R) = -lA \quad (5.15)$$

Hence

$$A = \frac{2k}{k + il} \quad (5.16)$$

$$R = \frac{k - il}{k + il}. \quad (5.17)$$

In particular $|R|^2 = 1$: the reflected flux equals the incoming flux, and thus the reflection probability for any given incoming particle is one.

Case 2: $E > U$ Now we have

$$l = \frac{\sqrt{2m(E - U)}}{\hbar} \quad (5.18)$$

and

$$\psi(x) = \begin{cases} T \exp(ilx) & x > 0, \\ \exp(ikx) + R \exp(-ikx) & x < 0. \end{cases} \quad (5.19)$$

Continuity of ψ, ψ' at $x = 0$ gives

$$T = 1 + R \quad (5.20)$$

$$\frac{l}{k}T = 1 - R. \quad (5.21)$$

Hence

$$T = \frac{2k}{k + l}, \quad (5.22)$$

$$R = \frac{k - l}{k + l}. \quad (5.23)$$

The incoming, reflected and transmitted fluxes are respectively

$$F_I = \frac{\hbar k}{m} \quad (5.24)$$

$$F_R = \frac{\hbar k}{m} |R|^2 = \frac{\hbar k}{m} \left(\frac{k - l}{k + l} \right)^2 \quad (5.25)$$

$$F_T = \frac{\hbar l}{m} |T|^2 = \frac{\hbar l}{m} \left(\frac{4k^2}{(k + l)^2} \right). \quad (5.26)$$

We see that

$$\begin{aligned}
F_R + F_T &= \frac{\hbar k(k-l)^2 + 4k^2l}{m(k+l)^2} \\
&= \frac{\hbar k(k+l)^2}{m(k+l)^2} \\
&= \frac{\hbar k}{m} \\
&= F_I
\end{aligned} \tag{5.27}$$

as particle conservation requires.

Comments 1. Case 1 ($E < U$) accords at least roughly with classical intuition: a particle with insufficient energy to climb the step is (eventually) reflected back.

2. The calculations for case 2 ($E > U$) apply for $U < 0$ – a downward step – as well as for $U > 0$. In the classical limit in which the step is negligible compared to the particle energy, $E \gg |U|$, we have

$$k \approx l, F_R \approx 0, F_T \approx 1.$$

This implies near-perfect transmission, again as classical intuition would suggest.

3. But the results in general do *not* accord with classical intuition. Consider for example the case $U < 0$ and $E \ll |U|$. Here we have

$$l \gg k, |R|^2 \approx 1, |T|^2 \approx 0.$$

In other words, we find near perfect **reflection** from a **downward** step, precisely the opposite result to that indicated by classical intuition.

5.4 Example II: The square potential barrier

Consider an incoming beam of particles with $E < U$ scattering from the potential

$$V(x) = \begin{cases} 0 & \text{for } x < 0, \\ U & \text{for } 0 < x < a \\ 0 & \text{for } x > a. \end{cases} \tag{5.28}$$

We have

$$\psi(x) = \begin{cases} T \exp(ikx) & \text{for } x > a, \\ A \exp(-lx) + B \exp(lx) & \text{for } 0 < x < a, \\ \exp(ikx) + R \exp(-ikx) & \text{for } x < 0, \end{cases} \tag{5.29}$$

where

$$E = \frac{\hbar^2 k^2}{2m}, \quad U - E = \frac{\hbar^2 l^2}{2m}. \quad (5.30)$$

Continuity of ψ and ψ' at $x = 0$ gives

$$1 + R = A + B, \quad (5.31)$$

$$ik(1 - R) = l(B - A), \quad (5.32)$$

Continuity of ψ and ψ' at $x = a$ gives

$$A \exp(-la) + B \exp(la) = T \exp(ika), \quad (5.33)$$

$$-lA \exp(-la) + lB \exp(la) = ikT \exp(ika). \quad (5.34)$$

Exercise Solve the algebra, to obtain

$$T = \frac{4ikle^{-ika}}{(2ikl + l^2 - k^2)e^{-la} + (2ikl - l^2 + k^2)e^{la}}. \quad (5.35)$$

This gives

$$|T|^2 = \frac{16k^2l^2}{4k^2l^2(e^{la} + e^{-la})^2 + (l^2 - k^2)^2(e^{la} - e^{-la})^2}. \quad (5.36)$$

In the case $la \gg 1$ we can neglect the $\exp(-la)$ terms, obtaining

$$|T|^2 \approx \frac{16k^2l^2}{4k^2l^2 \exp(2la) + (l^2 - k^2)^2 \exp(2la)}. \quad (5.37)$$

$$\approx \frac{16k^2l^2}{(k^2 + l^2)^2} \exp(-2la). \quad (5.38)$$

Since $l = \frac{\sqrt{2m(U-E)}}{\hbar}$, we find

$$|T|^2 \approx \exp\left(-\frac{2a}{\hbar} \sqrt{2m(U-E)}\right). \quad (5.39)$$

Comments 1. The most immediately striking result is the fact that T is *nonzero*: quantum particles *can* tunnel through classically impenetrable potentials.

2. If $la \gg 1$ the tunnelling probability is small; we approach the classical limit as $la \rightarrow \infty$. The tunnelling probability in this regime depends inverse exponentially on the barrier width a and on $\sqrt{U-E} \approx \sqrt{U}$ (if $E \ll U$). This is a general feature of tunnelling, true for a wide class of barrier potentials.

5.4.1 *Important examples

- Nuclear fission.
- Nuclear fusion.
- Muon-catalysed nuclear fusion.
- Scanning tunnelling electron microscopy. *

6 The basic formalism of quantum mechanics

6.1 Spaces of functions

We will not be too precise about the space of functions to which physically relevant wavefunctions belong in any given context. One possible example is the set of infinitely differentiable complex-valued functions $\psi(x)$ of one real variable x . This forms a *complex vector space* which we denote by $C^\infty(R)$. Another is the space of normalisable wavefunctions, which again forms a complex vector space, $L^2(R)$.

Exercise Check these are indeed vector spaces.

6.2 The inner product

We can define a natural inner product on the space of normalisable wavefunctions by

$$(\psi_1, \psi_2) = \int_{-\infty}^{\infty} \psi_1^*(x)\psi_2(x)dx. \quad (6.1)$$

Exercise: Show that this is well-defined whenever ψ_1 and ψ_2 are normalisable. [Hint: look ahead to Thm. 6.6 below.]

We extend the definition of the inner product $(,)$ to any ψ_1 and ψ_2 for which the integral is well-defined, whether or not ψ_1 or ψ_2 is normalisable.

6.2.1 Properties of the inner product

1. $(\psi_1, \psi_2) = (\psi_2, \psi_1)^*$
2. The inner product is *anti-linear* in the first entry and *linear* in the second:

$$\begin{aligned}(a_1\psi_1 + a_2\psi_2, \phi) &= a_1^*(\psi_1, \phi) + a_2^*(\psi_2, \phi), \\ (\psi, a_1\phi_1 + a_2\phi_2) &= a_1(\psi, \phi_1) + a_2(\psi, \phi_2).\end{aligned}$$

3. The inner product is *positive definite* on continuous wavefunctions, i.e.

$$(\psi, \psi) \geq 0 \quad \text{and} \quad (\psi, \psi) = 0 \text{ if and only if } \psi = 0.$$

To see this, note that

$$(\psi, \psi) = \int_{-\infty}^{\infty} \psi^*(x)\psi(x)dx = \int_{-\infty}^{\infty} |\psi(x)|^2 dx.$$

6.3 Operators

Recall that we defined an *operator* \hat{O} to be any *linear* map from a space of functions¹⁸ (for example $C^\infty(R)$ or $L^2(R)$) to itself, i.e. any map such that

$$\hat{O}(a_1\psi_1 + a_2\psi_2) = a_1\hat{O}\psi_1 + a_2\hat{O}\psi_2 \quad (6.2)$$

for all complex numbers a_1, a_2 and all $\psi_1, \psi_2 \in C^\infty(R)$.

6.3.1 Some examples of operators

- Finite differential operators $\sum_{i=0}^n p_i(x) \frac{\partial^i}{\partial x^i}$, where the $p_i(x)$ are polynomials in x . (This class includes \hat{p} and \hat{x} , defined below.)
- The translation operator

$$S_a : \psi(x) \rightarrow \psi_a(x) = \psi(x - a).$$

- The parity operator

$$P : \psi(x) \rightarrow \psi_P(x) = \psi(-x).$$

6.4 Hermitian operators

We define the *hermitian conjugate* A^\dagger of an operator A to be the operator such that

$$(A^\dagger\psi_1, \psi_2) = (\psi_1, A\psi_2) \quad (6.3)$$

for all normalisable wavefunctions ψ_1, ψ_2 . An operator is *hermitian* if $A = A^\dagger$.

Exercise Verify the identities

$$\begin{aligned} (a_1A_1 + a_2A_2)^\dagger &= a_1^*A_1^\dagger + a_2^*A_2^\dagger, \\ (AB)^\dagger &= B^\dagger A^\dagger. \end{aligned}$$

6.4.1 Classical states and dynamical variables

By the *state* of a physical system we mean its mathematical representation in a given theory. In classical mechanics, the states are points in a $2n$ -dimensional space, *phase space*. For example, a system of n point particles in 1D is described by $2n$ coordinates: $(x_1, \dots, x_n, p_1, \dots, p_n)$, where x_i is the position of the i th particle

¹⁸*Non-examinable comment: In our discussions below, we generally take the space of functions to be the space $L^2(R)$ of normalisable physical wavefunctions satisfying Eqn. (3.11). Strictly speaking, a completely rigorous discussion ought to take account of subtleties which arise from this choice. For example, the momentum operator is not actually defined on all functions in $L^2(R)$, but only those that are differentiable. Similarly, the position operator is not defined on all functions in $L^2(R)$, but only those normalisable functions $\psi(x)$ for which $x\psi(x)$ is also normalisable.

However, a fully rigorous treatment would go well beyond the scope of this course. We shall follow standard practice in introductory treatments of quantum mechanics and simply assume, whenever we have an equation in which the position, momentum or other operators act on physical wavefunctions, that the wavefunctions are chosen such that the action of the operators is well defined. [You should make the same assumption in tackling problem sheets or exam questions.] *

positions and p_i its momentum. Classical dynamical variables – for instance the energy

$$E = \sum_{i=1}^n \frac{p_i^2}{2m_i} + V(x_1, \dots, x_n) \quad (6.4)$$

– are defined by functions on phase space. Note that the classical state of a finite number of particles can be described by a finite number of parameters.

In principle, the classical state and (hence) the value of all classical dynamical variables can be measured with arbitrary precision.

6.4.2 Quantum states and observables

The possible quantum mechanical states of a physical system at any given time t are normalised wavefunctions. In particular, the state of a single particle in 1D is given by a wavefunction $\psi(x, t)$ such that $\int_{-\infty}^{\infty} |\psi(x, t)|^2 dx = 1$. We need infinitely many parameters to specify a normalised wavefunction, and thus to specify the state of a quantum system – even a single particle in 1D.

All quantum dynamical variables or *observables* – quantities we can measure – are represented by hermitian operators, and vice versa. Examples of hermitian operators defining quantum observables are the position, momentum and energy operators

$$\hat{p} = -i\hbar \frac{\partial}{\partial x} \quad (6.5)$$

$$\hat{x} = x \quad (\text{i.e. multiplication by } x) \quad (6.6)$$

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \quad (6.7)$$

Another example is the parity operator

$$P : \psi(x) \rightarrow \psi_P(x) = \psi(-x). \quad (6.8)$$

Exercise (important!): Verify that each of the operators (6.5-6.8) is hermitian.

6.5 Some theorems about hermitian operators

Theorem 6.1. *The eigenvalues of a hermitian operator are real.*

Proof. Let A be hermitian and ψ be a normalised eigenfunction with eigenvalue a : $A\psi = a\psi$. We have

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

Hence $a = a^*$. ■

Theorem 6.2. *Let 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 $A\psi_1 = a_1\psi_1$ and $A\psi_2 = a_2\psi_2$, and by theorem 6.1 we have that a_1, a_2 are real.

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

Hence $(a_1 - a_2)(\psi_1, \psi_2) = 0$, and since $a_1 \neq a_2$ we have $(\psi_1, \psi_2) = 0$. ■

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 momentum operator $\hat{p} = -i\hbar\frac{\partial}{\partial x}$ has eigenfunctions $\exp(\frac{i}{\hbar}px)$ with eigenvalue p .
- The position operator \hat{x} has eigenfunctions $\delta(x - q)$ with eigenvalue q : we have $\hat{x}\delta(x - q) = x\delta(x - q) = q\delta(x - q)$.
- The energy operator $\hat{H} = \frac{1}{2m}\hat{p}^2 + V(x)$, for a finite square well potential $V(x)$, has normalisable eigenfunctions (corresponding to bound states) and unnormalisable eigenfunctions (corresponding to scattering solutions).

Theorems 6.1 and 6.2 extend to these non-normalisable eigenfunctions of \hat{p}, \hat{x} and \hat{H} : these are the only operators with continuous sets of eigenfunctions that we consider in this course.

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 apply to both sets.¹⁹

¹⁹*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.*

Theorem 6.3. *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 functions $\psi(x)$ of one real variable x .*

Proof. This is quite hard to prove in complete generality. We will assume it without proof in this course. ■

Note: We say the hermitian operator A has a *degenerate* eigenvalue λ if it has more than one linearly independent eigenfunction with eigenvalue λ . If 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 .

Corollary 6.3.1. *Let A be a hermitian operator with a discrete set of normalised eigenfunctions $\{\psi_i\}_{i=1}^N$ (we include $N = \infty$ as a possibility) and a continuous set of eigenfunctions $\{\psi_\alpha\}_{\alpha \in \Delta}$, where the indexing set Δ is some sub-interval of the real numbers. Then any wavefunction ψ can be written as*

$$\psi = \sum_{i=1}^N a_i \psi_i + \int_{\Delta} a_\alpha \psi_\alpha d\alpha, \quad (6.9)$$

where $a_i = (\psi_i, \psi)$ and a_α are complex numbers.

Comment In fact, by appropriately normalising, we can also ensure that $a_\alpha = (\psi_\alpha, \psi)$.

Corollary 6.3.2. *Let A be a hermitian operator with (only) a discrete set of orthonormalised eigenfunctions $\{\psi_i\}_{i=1}^\infty$. Then any wavefunction ψ can be written as*

$$\psi = \sum_{i=1}^{\infty} (\psi_i, \psi) \psi_i. \quad (6.10)$$

Proof. Theorem 6.3 implies that

$$\psi = \sum_{i=1}^{\infty} a_i \psi_i \quad (6.11)$$

for some complex numbers a_i . Taking the inner product with ψ_j , and applying theorem 6.2, we see $(\psi_j, \psi) = a_j$. ■

6.6 Quantum measurement postulates

- **Postulate 1** Every quantum observable O is represented by a hermitian operator \hat{O} .
- **Postulate 2** The possible outcomes of a measurement of O are the eigenvalues of \hat{O} .
- **Postulate 3** If \hat{O} has (only) a discrete set of normalised eigenfunctions $\{\psi_i\}_{i=1}^{\infty}$ with corresponding distinct eigenvalues $\{\lambda_i\}$, and a measurement of O is carried out on a particle with normalised wavefunction

$$\psi = \sum_{i=1}^{\infty} a_i \psi_i,$$

then the probability of outcome λ_i is $|a_i|^2$.

- **Postulate 4** (the *projection postulate*) If a measurement of the observable O is carried out on a particle with normalised wavefunction $\psi(x, t)$ at time t and the outcome λ_i is obtained, the wavefunction instantaneously after the measurement becomes $\psi_i(x)$. [This is sometimes referred to as the “collapse of the wavefunction”.] The wavefunction then evolves according to the Schrödinger equation, with initial state $\psi_i(x)$ at time t , until the next measurement.

Notes

- It follows from postulates 2 and 3 that the total probability of all possible outcomes is

$$\sum_i |a_i|^2 = \sum_i (a_i \psi_i, a_i \psi_i) = \sum_{ij} (a_i \psi_i, a_j \psi_j) = (\psi, \psi) = 1. \quad (6.12)$$

So 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 position measurements.)

- It follows from postulates 3 and 4 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 .

Exercise: Check this.

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.

- If the wavefunction ψ is an eigenfunction ψ_i of \hat{O} , 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 the energy is measured.
- 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.
- We can extend postulates 3 and 4 to the case when \hat{O} has degenerate eigenvalues. If $\psi = \sum_i a_i \psi_i$ is measured, where $\{\psi_i\}_{i=1}^{\infty}$ are orthonormalised eigenfunctions of \hat{O} and $\{\psi_i\}_{i \in I}$ are a complete set of orthonormalised eigenfunctions with the same eigenvalue λ , the probability of outcome λ is $\sum_{i \in I} |a_i|^2$, and the state resulting after a measurement with outcome λ is (up to normalisation) $\sum_{i \in I} a_i \psi_i$.
- The projection postulate is so called because it implies that the post-measurement wavefunction $\psi_i(x, t)$ is obtained from the pre-measurement wavefunction $\psi(x, t)$ by the action of the *projection operator* P_i defined by $P_i : \psi \rightarrow (\psi, \psi_i) \psi_i$, up to normalisation. We call 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 $(x, 0, 0)$.

6.7 Expectation values

Consider a measurement of the observable A on the state ψ . If the corresponding hermitian operator \hat{A} has only a discrete set of normalisable eigenfunctions $\{\psi_i\}$, the possible outcomes are the corresponding eigenvalues λ_i , and the 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

$$\begin{aligned}
 \sum_i p_i \lambda_i &= \sum_i |(\psi, \psi_i)|^2 \lambda_i \\
 &= \left(\sum_i (\psi, \psi_i) \psi_i, \lambda_j \sum_j (\psi, \psi_j) \psi_j \right) \\
 &= (\psi, \hat{A} \psi).
 \end{aligned} \tag{6.13}$$

We write $\langle \hat{A} \rangle_\psi$ for the expectation value of a measurement of A on the state ψ . Thus we have

$$\langle \hat{A} \rangle_\psi = (\psi, \hat{A} \psi). \tag{6.14}$$

We can similarly justify this definition of expectation value for the position operator \hat{x} from the Born rule. Recall that the probability of obtaining a position measurement outcome in the interval $[x, x + dx]$ is given by $|\psi(x)|^2 dx$. The expectation value of a position measurement is thus

$$\int_{-\infty}^{\infty} x |\psi(x)|^2 dx = \int_{-\infty}^{\infty} \psi^*(x) x \psi(x) dx = (\psi, \hat{x}\psi).$$

This encourages us to take Eqn. (6.14) as a general definition of expectation value for any observable, whether its eigenvalues are discrete, continuous or a combination. This general definition too can be similarly justified, given a careful treatment of the eigenfunctions of general observables and their normalisation: this is beyond our scope here, but it is important to note that the definition turns out to agree with the statistical definition of expectation in all cases. In particular, the definitions

$$\langle \hat{p} \rangle_{\psi} = (\psi, \hat{p}\psi), \quad \langle \hat{H} \rangle_{\psi} = (\psi, \hat{H}\psi),$$

often give simple ways of calculating the statistical expectation of measurements of momentum and energy for general wavefunctions ψ , since the right hand sides are often given by simple integrals.

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

$$\langle a\hat{A} + b\hat{B} \rangle_{\psi} = a\langle \hat{A} \rangle_{\psi} + b\langle \hat{B} \rangle_{\psi}, \quad (6.15)$$

for any hermitian operators A, 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 A is hermitian, since observables are always represented by hermitian operators. A complex multiple of a hermitian operator is not generally hermitian: if A is hermitian then $(aA)^{\dagger} = a^*A^{\dagger}$.

6.8 Commutation relations

We define the *commutator* of two operators by

$$[A, B] = AB - BA. \quad (6.16)$$

It is easy to verify the identities

$$\begin{aligned} [A, A] &= 0, \\ [A, B] &= -[B, A], \\ [A, BC] &= [A, B]C + B[A, C] \\ [AB, C] &= A[B, C] + [A, C]B. \end{aligned} \quad (6.17)$$

Note that the commutator $[A, B]$ is a sum of products of operators, and thus itself an operator. Note also that it depends linearly on both entries.

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: see the following note and the later discussion of the uncertainty principle.

Definition We say two hermitian operators A and 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 $A\psi_i = a_i\psi_i$ and $B\psi_i = b_i\psi_i$ for some real numbers a_i, b_i .

Theorem 6.4. *Two hermitian operators A and B are simultaneously diagonalisable if and only if $[A, B] = 0$.*

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

$$[A, B]\psi_i = AB\psi_i - BA\psi_i = (a_i b_i - b_i a_i)\psi_i = 0. \quad (6.18)$$

If the basis $\{\psi_i\}$ is complete, any wavefunction ψ can be written as $\psi = \sum_i c_i \psi_i$, and we have that

$$[A, B]\psi = [A, B] \sum_i c_i \psi_i = \sum_i c_i [A, B]\psi_i = 0. \quad (6.19)$$

So we have the operator equation $[A, B] = 0$, where 0 is the zero operator, which maps any wavefunction to the zero function.

Conversely, if $[A, B] = 0$ and $A\psi_i = a_i\psi_i$, then $0 = [A, B]\psi_i = A(B\psi_i) - a_i(B\psi_i)$, so that $B\psi_i$ is also an eigenfunction of A with eigenvalue a_i . Thus B maps the eigenspace of A with eigenvalue a_i to itself. If we write E for this eigenspace, and $B|_E$ for the operator B restricted to E , then clearly $B|_E$ is a hermitian operator on E . Thus, by theorem 6.3, we can find a basis of E in which B acts diagonally. Since this is true for all eigenspaces of A , we can find a complete basis of simultaneous eigenfunctions of A and B . ■

6.8.1 The canonical commutation relations

Recall that the position and momentum operators are $\hat{p} = -i\hbar \frac{\partial}{\partial x}$ and $\hat{x} = x$ (multiplication by x). We can work out their commutator by considering their action on a general wavefunction ψ :

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

and as this is true for all ψ we have

$$[\hat{x}, \hat{p}] = i\hbar. \quad (6.20)$$

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.

6.9 Heisenberg's Uncertainty Principle

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

$$\begin{aligned} (\Delta_\psi A)^2 &= \langle (A - \langle A \rangle_\psi)^2 \rangle_\psi \\ &= \langle A^2 \rangle_\psi - (\langle A \rangle_\psi)^2. \end{aligned} \quad (6.21)$$

Note that Theorem 6.1 implies that the expectation value and the uncertainty are always real, as we would expect if they are physically meaningful.

Exercise Verify that $(\Delta_\psi A)^2$ is the statistical variance of the probability distribution for the possible outcomes of the measurement of A on ψ , and $\Delta_\psi A$ is the distribution's standard deviation.

Lemma 6.5. *The uncertainty $\Delta_\psi A \geq 0$ and $\Delta_\psi A = 0$ if and only if ψ is an eigenfunction of A .*

Proof. We can write

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

Write $\phi = (A - \langle A \rangle_\psi)\psi$. We have $(\phi, \phi) \geq 0$, with equality only if $\phi = 0$, i.e.

$$A\psi = \langle A \rangle_\psi \psi, \quad (6.22)$$

which implies that ψ is an eigenfunction of A . Conversely, if ψ is an eigenfunction of A then $(\psi, A\psi) = \langle A \rangle_\psi$, and so $A\psi = \langle A \rangle_\psi \psi$. In other words A has eigenvalue $\langle A \rangle_\psi$, and so $\Delta_\psi A = 0$. ■

Theorem 6.6 (Schwarz's inequality). *If ϕ, ψ are any two normalisable wavefunctions, then $|\langle \phi, \psi \rangle|^2 \leq (\phi, \phi)(\psi, \psi)$. We have equality if and only if $\phi = a\psi$ for some complex number a .*

Proof. For any a , $0 \leq (\phi - a\psi, \phi - a\psi)$. If we take $a = \frac{(\psi, \phi)}{(\psi, \psi)}$, we see

$$\begin{aligned} 0 &\leq (\phi, \phi) - \frac{|(\psi, \phi)|^2}{(\psi, \psi)} - \frac{|(\psi, \phi)|^2}{(\psi, \psi)} + \frac{|(\psi, \phi)|^2}{(\psi, \psi)} \\ &= (\phi, \phi) - \frac{|(\psi, \phi)|^2}{(\psi, \psi)}. \end{aligned}$$

Multiplying through by (ψ, ψ) we obtain the inequality. We have equality only if $\phi - a\psi = 0$, for the value of a specified above. ■

Theorem 6.7 (the generalised uncertainty relations). *If A and B are any two observables, and ψ is any state, then*

$$\Delta_\psi A \Delta_\psi B \geq \frac{1}{2} |(\psi, [A, B]\psi)|. \quad (6.23)$$

Proof. We have

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

So, writing $A' = A - \langle A \rangle_\psi$, $B' = B - \langle B \rangle_\psi$, we have

$$\begin{aligned} (\Delta_\psi A)^2 (\Delta_\psi B)^2 &= (A'\psi, A'\psi)(B'\psi, B'\psi) \\ &\geq |(A'\psi, B'\psi)|^2 \quad (\text{from Thm. 6.6}) \\ &= |(\psi, A'B'\psi)|^2 \quad (\text{as } A' \text{ is hermitian}). \end{aligned} \quad (6.24)$$

Now: (1) $A'B' = \frac{1}{2}([A', B'] + \{A', B'\})$, where the *anti-commutator* $\{A', B'\} = A'B' + B'A'$;

(2) $[A', B'] = [A, B]$;

(3) $(\psi, \{A', B'\}\psi)$ is real, since

$$(\psi, \{A', B'\}\psi) = ((\{A', B'\})^\dagger \psi, \psi) = ((\{A', B'\})\psi, \psi) = (\psi, \{A', B'\}\psi)^*. \quad (6.25)$$

A similar argument shows that $(\psi, [A', B']\psi)$ is imaginary.

So

$$|(\psi, A'B'\psi)|^2 = \frac{1}{4} (|(\psi, \{A', B'\}\psi)|^2 + |(\psi, [A', B']\psi)|^2). \quad (6.26)$$

Combining (6.24) and (6.26), we have that

$$(\Delta_\psi A)^2 (\Delta_\psi B)^2 \geq \frac{1}{4} |(\psi, [A, B]\psi)|^2. \quad (6.27)$$

Taking the square root gives (6.23). ■

Corollary 6.7.1. *(the Heisenberg uncertainty principle for position and momentum).*

$$(\Delta_\psi x)(\Delta_\psi p) \geq \frac{1}{2} \hbar \quad (6.28)$$

Proof. Taking $A = \hat{x}$ and $B = \hat{p}$, we have $[A, B] = i\hbar$, and the result follows from Thm. 6.23. ■

Thus, the smaller the uncertainty in position, $\Delta_\psi x$, the greater the minimum possible uncertainty in momentum, $\Delta_\psi p$, and vice versa.

Lemma 6.8. *If*

$$\hat{x}\psi = ia\hat{p}\psi \tag{6.29}$$

for some real parameter a , then $(\Delta_\psi x)(\Delta_\psi p) = \frac{1}{2}\hbar$.

Proof. If $\hat{x}\psi = ia\hat{p}\psi$ then we have

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

which is the condition for the first term on the RHS of (6.26) 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 $\langle \hat{x} \rangle_\psi = \langle \hat{p} \rangle_\psi = 0$. Hence

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

which means we have equality in the Schwarz's inequality (6.24) used to derive (6.28). ■

Lemma 6.9. *The condition (6.29) holds if and only if $\psi(x) = C \exp(-bx^2)$ for some constants b, C .*

Proof. If $\hat{x}\psi = ia\hat{p}\psi$ for some real a , we have that $x\psi = a\hbar \frac{\partial}{\partial x}\psi$ and so $\psi(x) = C \exp(-bx^2)$ for some real $b = -\frac{1}{2a\hbar}$, and because we have equality in (6.28) 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 . ■

Note: For the wavefunction to be normalisable, we require $b > 0$ and $C \neq 0$. We can take $C = |C| > 0$ by multiplying $\psi(x)$ by a phase factor (which does not alter any physical quantity: the probabilities of outcomes for any measurement are unaffected).

Exercise: Show that the condition (6.29) for minimum uncertainty is necessary as well as sufficient. Hence show that the normalisable minimum uncertainty states are precisely the wavefunctions defined by Gaussian functions.

6.9.1 What does the uncertainty principle tell us?

The uncertainty principle is a mathematical statement relating the uncertainties of x and p (or more generally A and B), which are quantities defined for a given state ψ . We can say, for example, that for any state ψ with $\Delta_\psi x = \delta$, we have $\Delta_\psi p \geq \frac{\hbar}{2\delta}$.

Heisenberg originally suggested that the uncertainty principle can be understood simply by observing that a measurement of A creates uncertainty by disturbing the value of any observable B that does not commute with A . This is not a valid argument! There are two problems with it.

First, according to quantum mechanics, it is not generally the case that there is a definite fixed pre-measurement value of either A or B , which is somehow disturbed in the process of our attempting to measure it. Unless the wavefunction is an eigenfunction of A or B , the result of any measurement is indeterminate.

Second, the mathematical derivation of the uncertainty principle does not require us to consider measurements of A or B actually taking place. The quantity $\Delta_\psi A$ is mathematically defined whether or not we choose to carry out a measurement of A . Even if we choose to interpret it as referring to a possible measurement, we cannot interpret the mathematics as referring to *successive* measurements of A

and B . $(\Delta_\psi A)$ and $(\Delta_\psi B)$ are the standard deviations for the outcomes of measurements of A and B , but these hypothetical measurements are *alternative* possible measurements on the same state ψ , *not* actual measurements carried out one after the other. If we measured, say, first A and then B , the first measurement would collapse the wavefunction onto an eigenfunction of A , and the second measurement would hence not generally be a measurement on the original state ψ .

6.10 Ehrenfest's theorem

Theorem 6.10 (Ehrenfest's theorem). *The expectation value $\langle A \rangle_\psi$ of an operator A in the state ψ evolves by*

$$\frac{d}{dt}\langle A \rangle_\psi = \frac{i}{\hbar}\langle [\hat{H}, A] \rangle_\psi + \left\langle \frac{\partial A}{\partial t} \right\rangle_\psi. \quad (6.30)$$

Proof. We have

$$\begin{aligned} \frac{d\langle A \rangle_\psi}{dt} &= \frac{d}{dt} \int_{-\infty}^{\infty} \psi^* A \psi dx \\ &= \int_{-\infty}^{\infty} \left(\frac{\partial \psi^*}{\partial t} A \psi + \psi^* \frac{\partial A}{\partial t} \psi + \psi^* A \frac{\partial \psi}{\partial t} \right) dx \\ &= \left\langle \frac{\partial A}{\partial t} \right\rangle_\psi + \frac{i}{\hbar} \int_{-\infty}^{\infty} ((\hat{H}\psi)^* A \psi - \psi^* A (\hat{H}\psi)) dx \\ &= \left\langle \frac{\partial A}{\partial t} \right\rangle_\psi + \frac{i}{\hbar} \int_{-\infty}^{\infty} (\psi^* \hat{H} A \psi - \psi^* A (\hat{H}\psi)) dx \\ &= \frac{i}{\hbar} \langle [\hat{H}, A] \rangle_\psi + \left\langle \frac{\partial A}{\partial t} \right\rangle_\psi. \end{aligned} \quad (6.31)$$

■

6.10.1 Applications of Ehrenfest's theorem

For $\hat{H} = \frac{\hat{p}^2}{2m} + V(x)$, we have

$$\begin{aligned} [\hat{H}, \hat{p}] &= [V(x), \hat{p}] \\ &= [V(x), -i\hbar \frac{\partial}{\partial x}] \\ &= i\hbar \frac{dV}{dx} \end{aligned} \quad (6.32)$$

$$\begin{aligned} [\hat{H}, \hat{x}] &= \left[\frac{\hat{p}^2}{2m}, \hat{x} \right] \\ &= \frac{1}{2m} 2[\hat{p}, \hat{x}]\hat{p} = \frac{-i\hbar}{m} \hat{p} \end{aligned} \quad (6.33)$$

$$[\hat{H}, \hat{H}] = 0. \quad (6.34)$$

Since none of these operators is explicitly time-dependent, we have that $\frac{\partial \hat{H}}{\partial t} = \frac{\partial \hat{x}}{\partial t} = \frac{\partial \hat{p}}{\partial t} = 0$ and so the $\langle \frac{\partial A}{\partial t} \rangle_\psi$ term on the RHS of (6.30) vanishes in each case, giving

$$\begin{aligned}\frac{d}{dt}\langle \hat{p} \rangle_\psi &= -\left\langle \frac{dV}{dx} \right\rangle_\psi, \\ \frac{d}{dt}\langle \hat{x} \rangle_\psi &= \frac{1}{m}\langle \hat{p} \rangle_\psi, \\ \frac{d}{dt}\langle \hat{H} \rangle_\psi &= 0.\end{aligned}\tag{6.35}$$

These are quantum versions of the classical laws $\frac{d}{dt}x = \frac{1}{m}p$ (which follows from $p = mv$), $\frac{d}{dt}p = -\frac{dV}{dx}$ (which follows from $F = ma$), and $\frac{d}{dt}E = 0$ (conservation of total energy).

* Ehrenfest's theorem and the harmonic oscillator. *

6.11 *The harmonic oscillator revisited

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

$$\begin{aligned}\hat{H} &= \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 \\ &= \frac{1}{2m}(\hat{p} + im\omega\hat{x})(\hat{p} - im\omega\hat{x}) + \frac{i\omega}{2}[\hat{p}, \hat{x}] \\ &= \frac{1}{2m}(\hat{p} + im\omega\hat{x})(\hat{p} - im\omega\hat{x}) + \frac{\hbar\omega}{2}\end{aligned}\quad (6.36)$$

Define the operator $a = \frac{1}{\sqrt{2m}}(\hat{p} - im\omega\hat{x})$. Since \hat{p} and \hat{x} are hermitian, we have $a^\dagger = \frac{1}{\sqrt{2m}}(\hat{p} + im\omega\hat{x})$, and

$$\hat{H} = a^\dagger a + \frac{1}{2}\hbar\omega. \quad (6.37)$$

We have the following commutation relations:

$$[a, a^\dagger] = \frac{1}{2m}(-im\omega)2[\hat{x}, \hat{p}] = \hbar\omega, \quad (6.38)$$

$$[\hat{H}, a] = [a^\dagger a, a] = [a^\dagger, a]a = -\hbar\omega a, \quad (6.39)$$

$$[\hat{H}, a^\dagger] = [a^\dagger a, a^\dagger] = a^\dagger[a, a^\dagger] = \hbar\omega a^\dagger. \quad (6.40)$$

Suppose now that ψ is a harmonic oscillator eigenfunction of energy E :

$$\hat{H}\psi = E\psi.$$

We then have

$$\hat{H}a\psi = [\hat{H}, a]\psi + a\hat{H}\psi = (E - \hbar\omega)a\psi \quad (6.41)$$

$$\hat{H}a^\dagger\psi = [\hat{H}, a^\dagger]\psi + a^\dagger\hat{H}\psi = (E + \hbar\omega)a^\dagger\psi, \quad (6.42)$$

so that $a\psi$ and $a^\dagger\psi$ are eigenfunctions of energy $(E - \hbar\omega)$ and $(E + \hbar\omega)$.

We can use this to prove by induction that $a^n\psi$ and $(a^\dagger)^n\psi$ are eigenfunctions of energy $(E - n\hbar\omega)$ and $(E + n\hbar\omega)$. For example,

$$\hat{H}a^n\psi = \hat{H}a(a^{n-1}\psi) = (E_{n-1} - \hbar\omega)a^n\psi, \quad (6.43)$$

where E_r is the energy eigenvalue of $a^r\psi$. Since $E_0 = E$, it follows by induction that $E_n = (E - n\hbar\omega)$.

In particular, if it were true that $a^n \psi \neq 0$ 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

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

since both terms are non-negative. (Important note: this argument can obviously be generalised to show that, if we have any potential V such that $V(x) \geq 0$ for all x , then $\langle \hat{H} \rangle_\psi \geq 0$ for all states ψ .)

So there cannot be negative energy eigenfunctions. Thus there must be a lowest energy (i.e. ground state) eigenfunction ψ_0 such that

$$0 = a\psi_0 = \frac{1}{\sqrt{2m}} (\hat{p} - im\omega\hat{x})\psi_0, \quad (6.44)$$

which implies

$$-i\hbar \frac{d\psi_0}{dx} = im\omega x \psi_0 \quad (6.45)$$

and hence

$$\psi_0(x) = C \exp\left(-\frac{m\omega x^2}{2\hbar}\right), \quad (6.46)$$

which is indeed the ground state wavefunction we previously obtained. (This time, though, we derived it much more simply, by introducing the operators a , a^\dagger and considering their commutation relations.)

Since $\hat{H} = a^\dagger a + \frac{\hbar\omega}{2}$ and $a\psi_0 = 0$, we have $\hat{H}\psi_0 = \frac{\hbar\omega}{2}\psi_0$, giving us the previously obtained value of $\frac{\hbar\omega}{2}$ for the ground state energy. We have also obtained a closed form expression (6.46) for the ground state and hence for the excited states,

$$(a^\dagger)^n \psi_0 = C \left(\frac{1}{\sqrt{2m}} (\hat{p} + im\omega\hat{x}) \right)^n \exp\left(-\frac{m\omega x^2}{2\hbar}\right), \quad (6.47)$$

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

We can similarly show that there cannot be eigenfunctions ψ' with energies taking values other than $(n + \frac{1}{2})\hbar\omega$. If there were, then $a^m \psi'$ cannot vanish for any m , since ψ_0 is the unique wavefunction annihilated by 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} , a and a^\dagger). *

7 The 3D Schrödinger equation

7.1 Quantum mechanics in three dimensions

We can develop quantum mechanics in three dimensions following the analogy with classical mechanics that we used to obtain the 1D Schrödinger equation. The classical state of a single particle is described by six dynamical variables: its position $\underline{x} = (x_1, x_2, x_3) = \sum_{i=1}^3 x_i \underline{e}_i$ and momentum $\underline{p} = (p_1, p_2, p_3) = \sum_{i=1}^3 p_i \underline{e}_i$, where \underline{e}_i are the standard orthonormal basis vectors. The particle's energy is

$$H = \frac{\underline{p} \cdot \underline{p}}{2M} + V(\underline{x}). \quad (7.1)$$

(Throughout this section, we use M to denote mass, to avoid confusion with the L_3 angular momentum eigenvalue denoted by m which we introduce below.)

Proceeding by analogy with the 1D case, we can introduce operators

$$\begin{aligned} \hat{x}_i &= x_i && \text{(i.e. multiplication by } x_i \text{),} \\ \hat{p}_i &= -i\hbar \frac{\partial}{\partial x_i}, \end{aligned}$$

$$\text{or in vector form } \hat{\underline{x}} = \underline{x} \quad (7.2)$$

$$\hat{\underline{p}} = -i\hbar \underline{\nabla} \quad (7.3)$$

$$\begin{aligned} \text{and } \hat{H} &= \frac{\hat{\underline{p}} \cdot \hat{\underline{p}}}{2M} + V(\underline{x}) \\ &= -\frac{\hbar^2}{2M} \nabla^2 + V(\underline{x}). \end{aligned} \quad (7.4)$$

Here we define a *vector operator* \underline{A} to be a triple of operators (A_1, A_2, A_3) such that $\underline{A}\psi(\underline{x}) = (A_1\psi(\underline{x}), A_2\psi(\underline{x}), A_3\psi(\underline{x}))$ is a vector for all wavefunctions $\psi(\underline{x})$ and all \underline{x} .

We also take the wavefunction ψ to depend on 3 space and 1 time coordinates: $\psi \equiv \psi(\underline{x}, t)$. The 3D normalisation condition is

$$\int |\psi(\underline{x}, t)|^2 d^3x = 1. \quad (7.5)$$

Following the 1D analogy gives us the *3D time-dependent Schrödinger equation*

$$\hat{H}\psi(\underline{x}, t) = i\hbar \frac{\partial}{\partial t} \psi(\underline{x}, t), \quad (7.6)$$

or more explicitly, for a time-independent potential V ,

$$-\frac{\hbar^2}{2M} \nabla^2 \psi(\underline{x}, t) + V(\underline{x})\psi(\underline{x}, t) = i\hbar \frac{\partial}{\partial t} \psi(\underline{x}, t). \quad (7.7)$$

Using the method of separation of variables, as before, we can derive the 3D time-independent Schrödinger equation

$$-\frac{\hbar^2}{2M} \nabla^2 \psi(\underline{x}) + V(\underline{x})\psi(\underline{x}) = E\psi(\underline{x}). \quad (7.8)$$

We can define the probability density and current

$$\rho(\underline{x}, t) = |\psi(\underline{x}, t)|^2, \quad (7.9)$$

$$\underline{J}(\underline{x}, t) = \frac{-i\hbar}{2m}(\psi^*(\underline{x}, t)\underline{\nabla}\psi(\underline{x}, t) - (\underline{\nabla}\psi^*(\underline{x}, t))\psi(\underline{x}, t)), \quad (7.10)$$

and as in the 1D case (cf (3.17)) we can show that they obey a conservation equation

$$\frac{\partial\rho}{\partial t} + \underline{\nabla}\cdot\underline{J} = 0. \quad (7.11)$$

Notice that the 3D Schrödinger equation, like the 1D SE, is linear and the superposition principle thus applies to its solutions: there is a physical solution corresponding to any linear combination of two (or more) physical solutions.

The Born rule naturally extends to the 3D case: the probability of finding a particle in a small 3D volume V which contains a point \underline{x}_0 is

$$\begin{aligned} \int_V |\psi(\underline{x}, t)|^2 d^3x &\approx V |\psi(\underline{x}_0, t)|^2 \\ &= V \rho(\underline{x}_0, t). \end{aligned} \quad (7.12)$$

We define the inner product of 3D wavefunctions by

$$(\psi_1, \psi_2) = \int \psi_1^*(\underline{x})\psi_2(\underline{x})d^3x.$$

The definition of the expectation value of an observable \hat{A} in terms of the corresponding hermitian operator A also naturally extends to 3D wavefunctions:

$$\langle \hat{A} \rangle_\psi = \int \psi^*(\underline{x}, t)A\psi(\underline{x}, t)d^3x = (\psi, A\psi). \quad (7.13)$$

We can thus define the uncertainty $\Delta_\psi A$ as in (6.21), using the definition (7.13) for expectation values.

The proofs of theorems 6.1 and 6.2 apply to hermitian operators on 3D wavefunctions just as to hermitian operators on 1D wavefunctions.

Theorem 6.3 also extends to hermitian operators on 3D wavefunctions: the discrete and continuous sets of eigenfunctions of a hermitian operator form a complete orthogonal basis of the normalisable complex-valued functions $\psi(\underline{x})$ of 3D vectors \underline{x} .

7.2 Spherically symmetric potentials

The 3D time-independent Schrödinger equation simplifies considerably when the potential $V(\underline{x})$ is spherically symmetric about the origin.²⁰ We call this a *central* potential.

²⁰Of course, if V is spherically symmetric about any given point, we can redefine our coordinates by translation to make that point the origin.

It is convenient to use spherical polar coordinates

$$x_1 = r \sin \theta \cos \phi, \quad x_2 = r \sin \theta \sin \phi, \quad x_3 = r \cos \theta, \quad (7.14)$$

which have ranges $0 \leq r < \infty$, $0 \leq \phi < 2\pi$, $0 \leq \theta \leq \pi$.

In spherical polars, a central potential $V(\underline{x}) \equiv V(r, \theta, \phi)$ depends only on $r = |\underline{x}|$. Recall that in Cartesian coordinates we have $\nabla^2 = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}$. In spherical polars this becomes²¹

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \quad (7.15)$$

For spherically symmetric stationary states $\psi(r)$, the time-independent SE simplifies. We have $\frac{\partial \psi}{\partial \theta} = \frac{\partial \psi}{\partial \phi} = 0$ and so

$$\begin{aligned} \nabla^2 \psi &= \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d}{dr} \right) \psi \\ &= \frac{d^2 \psi}{dr^2} + \frac{2}{r} \frac{d\psi}{dr} \\ &= \frac{1}{r} \frac{d^2}{dr^2} (r\psi). \end{aligned} \quad (7.16)$$

so we have

$$-\frac{\hbar^2}{2M} \frac{1}{r} \frac{d^2}{dr^2} (r\psi(r)) + V(r)\psi(r) = E\psi(r), \quad (7.17)$$

which we can rewrite as

$$-\frac{\hbar^2}{2M} \frac{d^2}{dr^2} (r\psi(r)) + V(r)(r\psi(r)) = E(r\psi(r)). \quad (7.18)$$

Notice that (7.18) is the 1D time-independent SE for $\phi(r) = r\psi(r)$, on the interval $0 \leq r < \infty$.

We require $\phi(r) \rightarrow 0$ as $r \rightarrow 0$, otherwise $\psi(r) \approx O(1/r)$ as $r \rightarrow 0$ and so is singular at $r = 0$. It can be shown that this means that the 3D Schrödinger equation (7.7) fails to hold there.²²

Any solution to (7.18) with $\phi(r) \rightarrow 0$ as $r \rightarrow 0$ can be extended to an odd parity solution $\tilde{\phi}(r)$ of the 1D SE in $-\infty < r < \infty$ of the same energy, with $\tilde{\phi}(r)$ and $\frac{d\tilde{\phi}}{dr}$ continuously defined at $r = 0$, by defining

$$\tilde{\phi}(r) = \begin{cases} \phi(r) & r \geq 0, \\ -\phi(-r) & r < 0. \end{cases} \quad (7.19)$$

Conversely, any odd parity solution of the 1D SE for $-\infty < r < \infty$ defines a solution to (7.18) with $\phi(r) \rightarrow 0$ as $r \rightarrow 0$ and $\frac{d\phi}{dr}$ finite at $r = 0$. Provided that

²¹See IA Vector Calculus, or e.g. Collinson "Introductory Vector Analysis", Chap. 12.

²²* See Dirac, "The Principles of Quantum Mechanics" (4th edition), Chap. VI for a full discussion. *

$V(r)$ is finite and continuous at $r = 0$, these are the correct continuity conditions for solutions of the 3D SE: they imply that ψ and ψ' are continuous at the origin.

Solving (7.18) thus becomes equivalent to finding *odd parity solutions* to the 1D SE for $-\infty < r < \infty$.

Comments We will show later (see Thm. 7.2) that the ground state (the lowest energy bound state, if there is one) of a 3D quantum system with spherically symmetric potential is itself spherically symmetric. (Cf. the 1D result that the ground state of a symmetric potential always has even parity.) Hence we can always use the method above to obtain the ground state.

One might wonder whether there might not exist even parity solutions $\phi_+(r)$ of the 1D SE with the property that $\phi_+(0) = \frac{d\phi_+}{dr}(0) = 0$, which would also define solutions to (7.18) for $0 \leq r < \infty$ with the appropriate properties. The following lemma rules this out.

Lemma 7.1. *There are no even parity solutions $\phi_+(r)$ of the 1D SE with the property that $\phi_+(0) = \frac{d\phi_+}{dr}(0) = 0$.*

Proof. If such a solution ϕ_+ were to exist, we could define a continuous odd parity solution $\phi_-(r)$ by

$$\phi_-(r) = \begin{cases} \phi_+(r) & r \geq 0, \\ -\phi_+(-r) & r < 0. \end{cases} \quad (7.20)$$

Then, by the superposition principle, $\phi(r) = \phi_+(r) - \phi_-(r)$ would also be a solution. But we have $\phi(r) = 0$ for $r > 0$, so that all derivatives of ϕ vanish for $r \geq 0$. The Schrödinger equation has no non-trivial solutions with this property: hence $\phi(r) = 0$ for all r . Hence $\phi_+(r) = \phi_-(r) = 0$ for all r , so in particular the hypothesised even parity solution ϕ_+ is not a physical solution, as it vanishes everywhere. ■

7.3 Examples of spherically symmetric potentials

7.3.1 The spherical harmonic oscillator

has potential

$$V(r) = \frac{1}{2}M\omega^2r^2.$$

The general method we have just given for constructing spherically symmetric stationary states shows that its spherically symmetric stationary states are related by (7.19) to the odd parity bound states of the 1D harmonic oscillator, and have the same energies. Thus the lowest energy spherically symmetric stationary state – i.e. the ground state – has energy $\frac{3}{2}\hbar\omega$, and the higher energy (excited) spherically symmetric states have energies $(2n + \frac{3}{2})\hbar\omega$ for positive integer n .

7.3.2 The spherical square well

has potential

$$V(r) = \begin{cases} -U & r < a, \\ 0 & r > a. \end{cases} \quad (7.21)$$

By the above argument, spherically symmetric stationary states correspond to odd parity bound states of the 1D square well potential

$$V(x) = \begin{cases} -U & |x| < a, \\ 0 & |x| > a. \end{cases} \quad (7.22)$$

These, if they exist, can be obtained by the graphical method used earlier to obtain 1D square well potential bound states. In particular, one can show (cf. section 4 above and example sheet I, question 10) that there exists an odd parity bound state if and only if

$$\sqrt{\frac{2MU}{\hbar^2}} \geq \frac{\pi}{2a}. \quad (7.23)$$

So, if this condition is not satisfied, the 3D spherical square well does not have a spherically symmetric stationary state: i.e. it does not have a ground state, and thus does not have any bound states.

As this illustrates, 3D potential wells (continuous potentials with $V(x) \leq 0$ for all x , $V(x) < 0$ for some x , and $V(x) = 0$ for $|x| > a$, for some finite a) do **not** necessarily have bound states. This is in contrast to the 1D case:

Exercise (important!): Show that all 1D potential wells have at least one bound state.

7.4 Spherically symmetric bound states of the hydrogen atom

We model the hydrogen atom by treating the proton as infinitely massive and at rest.²³

We seek spherically symmetric bound state wavefunctions $\psi(r)$ for the electron orbiting in a Coulomb potential $V(r) = -\frac{e^2}{4\pi\epsilon_0 r}$:

$$-\frac{\hbar^2}{2M} \left(\frac{d^2\psi}{dr^2} + \frac{2}{r} \frac{d\psi}{dr} \right) - \frac{e^2}{4\pi\epsilon_0 r} \psi(r) = E\psi(r), \quad (7.24)$$

for some $E < 0$. Writing $a = \frac{e^2 M}{2\pi\epsilon_0 \hbar^2}$, $b = \frac{\sqrt{-2ME}}{\hbar}$, we have

$$\frac{d^2\psi}{dr^2} + \frac{2}{r} \frac{d\psi}{dr} + \frac{a}{r} \psi - b^2 \psi = 0. \quad (7.25)$$

If we try the ansatz

$$\psi(r) \approx \exp(-br), \quad (7.26)$$

we see the first and fourth terms dominate the other two for large r , and cancel one another precisely. This suggests trying an ansatz of the form $\psi(r) = f(r) \exp(-br)$, with $f(r) = \sum_{n=0}^{\infty} a_n r^n$, in the hope of finding values of the coefficients a_n such that the four terms cancel precisely to all orders. (Cf. our first solution to the harmonic oscillator.)

Our previous discussion assumed that $V(r)$ is nonsingular as $r \rightarrow 0$. Since here $V(r)$ diverges as $r \rightarrow 0$, we cannot use the same justification as previously to argue that that $\phi(r) = r\psi(r) \rightarrow 0$ as $r \rightarrow 0$. However, we still require $\psi(r)$ to define a normalisable 3D wavefunction, so that

$$0 < \int_{r=0}^{\infty} r^2 |\psi(r)|^2 dr < \infty.$$

This implies that $\psi(r) = O(r^{-1})$ as $r \rightarrow 0$, i.e. that ψ can at worst have a singularity of order r^{-1} at zero. We also require that ψ should correspond to a continuous wavefunction. This excludes a singularity of order r^{-1} , so we require that ψ is regular — i.e. has a finite limit — as $r \rightarrow 0$.

²³*This can be shown to be an excellent approximation (see Part II Principles of Quantum Mechanics): it gives the correct energy levels up to an overall constant factor (the same for each energy level) of order $1 + (m_e/m_p) \approx 1.0005$.*

We have

$$\frac{d^2 f}{dr^2} + \left(\frac{2}{r} - 2b\right) \frac{df}{dr} + \frac{1}{r}(a - 2b)f(r) = 0. \quad (7.27)$$

Hence

$$\sum_{n=0}^{\infty} (a_n n(n-1)r^{n-2} + 2a_n n r^{n-2} - 2ba_n n r^{n-1} + (a-2b)a_n r^{n-1}) = 0 \quad (7.28)$$

and taking the coefficient of r^{n-2} we have

$$a_n n(n-1) + 2a_n n - 2ba_{n-1}(n-1) + (a-2b)a_{n-1} = 0 \text{ for } n \geq 1. \quad (7.29)$$

This gives

$$\begin{aligned} a_n &= a_{n-1} \frac{(2b(n-1) - (a-2b))}{n(n-1) + 2n} \\ &= a_{n-1} \frac{2bn - a}{n(n+1)}. \end{aligned} \quad (7.30)$$

We thus have that $a_n \rightarrow \frac{2b}{n}a_{n-1}$ for large n . If the coefficients do not vanish for large n , this means they have the asymptotic behaviour of the coefficients of $\exp(2br)$, i.e. $f(r) \approx C \exp(2br)$. This would give $\psi(r) \approx C \exp(2br) \exp(-br) = \exp(br)$, leading to an unnormalisable and thus unphysical wavefunction. So there must be some integer $N \geq 1$ for which $a_N = 0$, and we can take N to be the smallest such integer.

Then $a_{N-1} \neq 0$, so that $a_N = 0$ implies $2bN = a$ or $b = a/2N$, and so

$$\begin{aligned} E &= -\frac{\hbar^2 a^2}{8MN^2} \\ &= -\frac{Me^4}{32\pi^2 \epsilon_0^2 \hbar^2 N^2}, \end{aligned} \quad (7.31)$$

which is precisely the energy spectrum of the Bohr orbits, but now derived from quantum mechanics (though still with an assumption of spherical symmetry, which we will need to relax to obtain the general orbital wavefunction).

From

$$a = 2bN \quad \text{and} \quad a_n = a_{n-1} \frac{2bn - a}{n(n+1)} \quad (7.32)$$

we obtain

$$a_n = a_{n-1} 2b \frac{n - N}{n(n+1)}. \quad (7.33)$$

This gives solutions of the form

$$f(r) = \begin{cases} 1 & N = 1, \\ (1 - br) & N = 2, \\ (1 - 2br + \frac{2}{3}(br)^2) & N = 3, \end{cases} \quad (7.34)$$

and generally $f(r) = L_{N-1}^1(2br)$ where L_{N-1}^1 is one of the *associated Laguerre polynomials*.²⁴

The corresponding wavefunctions are $\psi(r) = CL_{N-1}^1(2br) \exp(-br)$, where the constant C is determined by normalisation.

7.5 Canonical Commutation Relations in 3D

From (7.2) we have

$$\hat{x}_i = x_i \quad (\text{multiplication by } x_i), \quad (7.35)$$

$$\hat{p}_i = -i\hbar \frac{\partial}{\partial x_i}, \quad (7.36)$$

and so, by calculating the action on a general wavefunction as before, we obtain

$$[\hat{x}_i, \hat{x}_j] = 0 = [\hat{p}_i, \hat{p}_j] \quad (7.37)$$

$$[\hat{x}_i, \hat{p}_j] = i\hbar \delta_{ij}. \quad (7.38)$$

(The prefactors i on the right hand sides are $\sqrt{-1}$.)

7.6 Orbital Angular Momentum

Recall that in classical mechanics we define the angular momentum vector

$$\underline{L} = \underline{x} \wedge \underline{p}, \quad L_i = \epsilon_{ijk} x_j p_k, \quad (7.39)$$

and that \underline{L} is conserved in a spherically symmetric potential $V(r)$.

We define the quantum mechanical operators

$$\underline{\hat{L}} = -i\hbar \underline{\hat{x}} \wedge \underline{\nabla}, \quad \hat{L}_i = -i\hbar \epsilon_{ijk} \hat{x}_j \frac{\partial}{\partial x_k}, \quad (7.40)$$

and the total angular momentum

$$\hat{L}^2 = \underline{\hat{L}} \cdot \underline{\hat{L}} = \hat{L}_1^2 + \hat{L}_2^2 + \hat{L}_3^2. \quad (7.41)$$

^{24*}There is a simple expression for the associated Laguerre polynomials:

$$L_N^k(x) = \frac{1}{N!} e^x x^{-k} \frac{d^N}{dx^N} (x^{N+k} e^{-x}).$$

Some plots of L_N^0 for small N and some other information about the Laguerre polynomials (L_N^k for $k = 0$) and the associated Laguerre polynomials can be found at mathworld.wolfram.com/LaguerrePolynomial.html and at mathworld.wolfram.com/AssociatedLaguerrePolynomial.html .*

7.6.1 Angular momentum commutation relations

$$\begin{aligned}
[\hat{L}_i, \hat{L}_j] &= -\hbar^2 \epsilon_{ilm} \epsilon_{jnp} \left[\hat{x}_l \frac{\partial}{\partial x_m}, \hat{x}_n \frac{\partial}{\partial x_p} \right] \\
&= -\hbar^2 \epsilon_{ilm} \epsilon_{jnp} \left(\left[\hat{x}_l \frac{\partial}{\partial x_m}, \hat{x}_n \right] \frac{\partial}{\partial x_p} + \hat{x}_n \left[\hat{x}_l \frac{\partial}{\partial x_m}, \frac{\partial}{\partial x_p} \right] \right) \\
&= -\hbar^2 \epsilon_{ilm} \epsilon_{jnp} \left(\hat{x}_l \left[\frac{\partial}{\partial x_m}, \hat{x}_n \right] \frac{\partial}{\partial x_p} + \hat{x}_n \left[\hat{x}_l, \frac{\partial}{\partial x_p} \right] \frac{\partial}{\partial x_m} \right) \\
&= -\hbar^2 \epsilon_{ilm} \epsilon_{jnp} \left(\hat{x}_l \delta_{mn} \frac{\partial}{\partial x_p} - \hat{x}_n \delta_{lp} \frac{\partial}{\partial x_m} \right) \\
&= -\hbar^2 \epsilon_{mil} \epsilon_{mpj} \hat{x}_l \frac{\partial}{\partial x_p} + \hbar^2 \epsilon_{pjn} \epsilon_{pmi} \hat{x}_n \frac{\partial}{\partial x_m} \\
&= -\hbar^2 (\delta_{ip} \delta_{lj} - \delta_{ij} \delta_{lp}) \hat{x}_l \frac{\partial}{\partial x_p} + \hbar^2 (\delta_{jm} \delta_{ni} - \delta_{ji} \delta_{nm}) \left(\hat{x}_n \frac{\partial}{\partial x_m} \right) \\
&= -\hbar^2 \left(\hat{x}_j \frac{\partial}{\partial x_i} - \delta_{ij} \hat{x}_l \frac{\partial}{\partial x_l} - \hat{x}_i \frac{\partial}{\partial x_j} + \delta_{ij} \hat{x}_l \frac{\partial}{\partial x_l} \right) \\
&= i\hbar \epsilon_{ijk} \hat{L}_k . \\
&=
\end{aligned}$$

$$\begin{aligned}
[\hat{L}^2, \hat{L}_i] &= [\hat{L}_j \hat{L}_j, \hat{L}_i] \\
&= [\hat{L}_j, \hat{L}_i] \hat{L}_j + \hat{L}_j [\hat{L}_j, \hat{L}_i] \\
&= i\hbar (\epsilon_{jik} (\hat{L}_k \hat{L}_j + \hat{L}_j \hat{L}_k)) \\
&= 0 .
\end{aligned} \tag{7.42}$$

Since the \hat{L}_i do not commute, they are not simultaneously diagonalisable. However, \hat{L}^2 and any one of the \hat{L}_i can be simultaneously diagonalised, since $[\hat{L}^2, \hat{L}_i] = 0$.

We also have

$$[\hat{L}_i, \hat{x}_j] = i\hbar \epsilon_{ijk} \hat{x}_k, \tag{7.43}$$

$$[\hat{L}_i, \hat{p}_j] = i\hbar \epsilon_{ijk} \hat{p}_k, \tag{7.44}$$

$$[\hat{L}_i, \sum_j \hat{x}_j^2] = 2i\hbar \epsilon_{ijk} \hat{x}_j \hat{x}_k = 0, \tag{7.45}$$

$$[\hat{L}_i, \sum_j \hat{p}_j^2] = 2i\hbar \epsilon_{ijk} \hat{p}_j \hat{p}_k = 0. \tag{7.46}$$

Now we have that $\hat{r} = \sqrt{\sum_j \hat{x}_j^2}$. We also have that $[\hat{L}_i, \sum_j \hat{x}_j^2] = 0$. One can show directly from this (see part II Principles of Quantum Mechanics), or check by calculation (Exercise), that $[\hat{L}_i, \hat{r}] = 0$. More generally, one can show that $[\hat{L}_i, V(r)] = 0$ for any spherically symmetric potential $V(r)$. We also have that

$$[\hat{L}_i, \frac{\hat{p} \cdot \hat{p}}{2m}] = [\hat{L}_i, \frac{1}{2m} \sum_j p_j^2] = 0. \tag{7.47}$$

So, for any spherically symmetric potential $V(r)$, we have that

$$\begin{aligned} [\hat{L}_i, \hat{H}] &= [\hat{L}_i, -\frac{\hbar^2}{2M}\nabla^2 + V(r)] \\ &= 0, \end{aligned} \tag{7.48}$$

$$[\hat{L}^2, \hat{H}] = 0. \tag{7.49}$$

In other words, \hat{H} , \hat{L}_i and \hat{L}^2 all commute with one another.

Comment This is an important and powerful result. Given any 3D quantum system, we can find a basis of simultaneous eigenfunctions of \hat{H} , \hat{L}^2 and \hat{L}_3 . (The standard convention is to use \hat{L}_3 . Of course, as the operators are related by rotational symmetry, \hat{L}_1 or \hat{L}_2 would work equally well.) As we will see, working in such a basis greatly simplifies the solution of the time-independent SE.

7.6.2 Angular momentum operators in spherical polar coordinates

We can translate the definitions of \hat{L}_i to spherical polars. We have

$$x_1 = r \sin \theta \cos \phi, \quad x_2 = r \sin \theta \sin \phi, \quad x_3 = r \cos \theta. \tag{7.50}$$

Thus

$$\begin{aligned} \frac{\partial}{\partial \theta} &= \sum_i \frac{\partial x_i}{\partial \theta} \frac{\partial}{\partial x_i} \\ &= r \cos \theta \cos \phi \frac{\partial}{\partial x_1} + r \cos \theta \sin \phi \frac{\partial}{\partial x_2} - r \sin \theta \frac{\partial}{\partial x_3}, \end{aligned} \tag{7.51}$$

$$\begin{aligned} \frac{\partial}{\partial \phi} &= \sum_i \frac{\partial x_i}{\partial \phi} \frac{\partial}{\partial x_i} \\ &= -r \sin \theta \sin \phi \frac{\partial}{\partial x_1} + r \sin \theta \cos \phi \frac{\partial}{\partial x_2}. \end{aligned} \tag{7.52}$$

We thus obtain

$$\begin{aligned} i\hbar(\cos \phi \cot \theta \frac{\partial}{\partial \phi} + \sin \phi \frac{\partial}{\partial \theta}) &= -i\hbar(x_2 \frac{\partial}{\partial x_3} - x_3 \frac{\partial}{\partial x_2}) \\ &= \hat{L}_1, \end{aligned} \tag{7.53}$$

$$\begin{aligned} i\hbar(\sin \phi \cot \theta \frac{\partial}{\partial \phi} + \cos \phi \frac{\partial}{\partial \theta}) &= -i\hbar(x_3 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_3}) \\ &= \hat{L}_2, \end{aligned} \tag{7.54}$$

$$\begin{aligned} -i\hbar \frac{\partial}{\partial \phi} &= -i\hbar(x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1}) \\ &= \hat{L}_3. \end{aligned} \tag{7.55}$$

Notes 1. We could obtain these identities more straightforwardly but long-windedly by first expressing $\frac{\partial}{\partial x_i}$ in terms of $r, \theta, \phi, \frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}, \frac{\partial}{\partial \phi}$ and thus converting the definitions of the \hat{L}_i into spherical polars.

2. Observe in particular that the \hat{L}_i are all independent of $\frac{\partial}{\partial r}$, as we expect from the commutation relation $[\hat{L}_i, r] = 0$.

We can also obtain

$$\begin{aligned}
\hat{L}^2 &= \sum_i \hat{L}_i^2 = (\hat{L}_1 + i\hat{L}_2)(\hat{L}_1 - i\hat{L}_2) + i[\hat{L}_1, \hat{L}_2] + \hat{L}_3^2 \\
&= (\hat{L}_1 + i\hat{L}_2)(\hat{L}_1 - i\hat{L}_2) - \hbar\hat{L}_3 + \hat{L}_3^2 \\
&= -\hbar^2(\cot\theta e^{i\phi} \frac{\partial}{\partial\phi} - ie^{i\phi} \frac{\partial}{\partial\theta})(\cot\theta e^{-i\phi} \frac{\partial}{\partial\phi} + ie^{-i\phi} \frac{\partial}{\partial\theta}) \\
&\quad + i\hbar^2 \frac{\partial}{\partial\phi} - \hbar^2 \frac{\partial^2}{\partial\phi^2} \\
&= -\hbar^2(\frac{\partial^2}{\partial\theta^2} + \cot\theta \frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2}) \\
&= -\hbar^2(\frac{1}{\sin\theta} \frac{\partial}{\partial\theta}(\sin\theta \frac{\partial}{\partial\theta}) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2}). \tag{7.56}
\end{aligned}$$

Recall that $[\hat{L}^2, \hat{L}_3] = 0$. We have

$$\begin{aligned}
\hat{L}^2 &= -\hbar^2(\frac{1}{\sin\theta} \frac{\partial}{\partial\theta}(\sin\theta \frac{\partial}{\partial\theta}) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2}), \\
\hat{L}_3 &= -i\hbar \frac{\partial}{\partial\phi}. \tag{7.57}
\end{aligned}$$

We can thus seek simultaneous eigenfunctions of the form $Y(\theta) \exp(im\phi)$, since $\hat{L}_3 \exp(im\phi) = \hbar m \exp(im\phi)$. As ϕ is defined modulo 2π , we need $e^{im(\phi+2\pi)} = e^{im\phi}$, so $e^{i2m\pi} = 1$ and m is an integer.

This leaves us with an eigenvalue equation for \hat{L}^2 :

$$-\hbar^2 \left(\frac{1}{\sin\theta} \frac{d}{d\theta} \left(\sin\theta \frac{d}{d\theta} \right) - \frac{m^2}{\sin^2\theta} \right) Y(\theta) = \lambda Y(\theta). \tag{7.58}$$

From a physics perspective, the key fact about this equation is that we can show it has non-singular solutions if and only if $\lambda = \hbar^2 l(l+1)$ for some integers $l \geq 0$ and m such that m is in the range $-l \leq m \leq l$.

Less crucial for now are the details, although it is interesting to see them. The solutions are called the *associated Legendre functions* $P_{l,m}(\theta)$. They can be obtained by reducing the equation to a standard form, using the substitution $w = \cos\theta$. Since θ is in the range $0 \leq \theta \leq \pi$, we have $-1 \leq w \leq 1$. We obtain the equation

$$-\hbar^2 \frac{d}{dw} \left((1-w^2) \frac{dY}{dw} \right) - \left(\lambda - \frac{m^2}{1-w^2} \right) Y = 0. \tag{7.59}$$

For $m = 0$ and $\lambda = \hbar^2 l(l+1)$ this is Legendre's differential equation for functions of degree l , which has solution $P_l(w)$. For general m it's an associated Legendre differential equation.²⁵

The associated Legendre functions can be obtained from the Legendre polynomials P_l by

$$P_{l,m}(\theta) = (\sin\theta)^{|m|} \frac{d^{|m|}}{d(\cos\theta)^{|m|}} P_l(\cos\theta). \tag{7.60}$$

We thus have the overall solution given by the *spherical harmonic* with *total angular momentum quantum number* l and \hat{L}_3 *quantum number* m :

$$Y_{l,m}(\theta, \phi) = P_{l,m}(\theta) \exp(im\phi),$$

²⁵Details can be found in Whittaker and Watson, "A course of modern analysis", 4th edition, (C.U.P., 1996).

an eigenfunction of \hat{L}^2 and \hat{L}_3 with eigenvalues $\hbar^2 l(l+1)$ and $\hbar m$ respectively.
 (For plots of some spherical harmonics see e.g. mathworld.wolfram.com/SphericalHarmonic.html.)

7.7 Solving the 3D Schrödinger equation for a spherically symmetric potential

The time-independent SE is

$$-\frac{\hbar^2}{2M}\nabla^2\psi + V(r)\psi = E\psi. \quad (7.61)$$

Recall that in spherical polar coordinates

$$\begin{aligned} \nabla^2 &= \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \\ \hat{L}^2 &= -\hbar^2 \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right). \end{aligned} \quad (7.62)$$

So we have

$$-\hbar^2 \nabla^2 = -\hbar^2 \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) + \frac{1}{r^2} \hat{L}^2. \quad (7.63)$$

We can thus rewrite the SE as

$$-\frac{\hbar^2}{2M} \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) + \frac{1}{2Mr^2} \hat{L}^2 \psi(r, \theta, \phi) + V(r)\psi(r, \theta, \phi) = E\psi(r, \theta, \phi). \quad (7.64)$$

If we separate variables, writing $\psi(r, \theta, \phi) = \psi(r)Y_{l,m}(\theta, \phi)$, this gives

$$-\frac{\hbar^2}{2M} \left(\frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} \right) \psi(r) + \left(\frac{\hbar^2}{2Mr^2} l(l+1) + V(r) \right) \psi(r) = E\psi(r). \quad (7.65)$$

So, we have a standard 1D radial Schrödinger equation for $\psi(r)$, with the modified potential

$$V(r) + \frac{\hbar^2 l(l+1)}{2Mr^2}.$$

Comments:

- We have seen that we can find a basis of simultaneous eigenfunctions of \hat{H} , \hat{L}^2 and \hat{L}_3 , with eigenvalues E , l and m respectively. Since the modified potential depends on l but not m , if an energy eigenspace with eigenvalue E contains any state with \hat{L}^2 eigenvalue l , it must contain states with *all* the associated \hat{L}_3 eigenvalues: $m = -l, -l+1, \dots, l$. This greatly simplifies the analysis of orbital angular momentum eigenstates associated with a given energy level.
- If the angular momentum $l = 0$ then also $m = 0$, and the function $Y_{00}(\theta, \phi)$ is constant. Thus all zero angular momentum states are spherically symmetric. Conversely, since the Y_{lm} for $l \neq 0$ are orthogonal to Y_{00} , all spherically symmetric states have zero angular momentum. This makes sense physically, since a state ψ with $\langle \underline{L} \rangle_\psi \neq \underline{0}$ by definition has a nonzero vector associated with it, which breaks spherical symmetry.

- *The fact that, in quantum mechanics, we can express $-\hbar^2\nabla^2$ in terms of a differential operator involving only r together with a term proportional to the operator L^2 is a consequence of the fact that, in classical mechanics, kinetic energy can be expressed as a sum of terms proportional to the radial momentum squared and the angular momentum squared. However, to derive the first result from the second requires addressing one or two subtleties beyond our scope. (A discussion can be found in Dirac’s book “The Principles of Quantum Mechanics”, 4th edition, Section 38.)*
- *We can understand qualitatively why we should expect the radial SE to depend on the angular momentum l , by noting that the “extra potential energy” term $\frac{l(l+1)\hbar^2}{2Mr^2}$ corresponds to the potential needed to produce the centripetal force $\frac{L^2}{mr^3}$ which would keep a classical particle of angular momentum L in a circular orbit, if we set $L = \hbar\sqrt{l(l+1)}$.*

7.7.1 Degeneracies

As noted above, the values of E for which this equation is solvable clearly may depend on l but not on m . As there are $(2l+1)$ possible values of $m = -l, -l+1, \dots, l$, each energy level would have degeneracy $(2l+1)$, assuming there are no further degeneracies.

The ground state We can now prove a result we stated earlier.

Theorem 7.2. *The ground state (i.e. lowest energy bound state) solution of the 3D Schrödinger equation for a spherically symmetric potential must have $l = m = 0$ and is thus spherically symmetric.*

Proof. The proof is by contradiction. Suppose that $\psi(r, \theta, \phi) = \psi(r)Y_{lm}(\theta, \phi)$, for some $l > 0$, is the lowest energy solution and has energy E . We have that

$$-\frac{\hbar^2}{2M}\left(\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr}\right)\psi(r) + \left(\frac{\hbar^2}{2Mr^2}l(l+1) + V(r)\right)\psi(r) = E\psi(r). \quad (7.66)$$

Now as \hat{H} , \hat{L}^2 and \hat{L}_3 are commuting hermitian operators, the space of wavefunctions is spanned by their simultaneous eigenstates. In particular, the space of zero angular momentum wavefunctions is spanned by orthonormal eigenstates $\psi_i(r, \theta, \phi)$ of \hat{H} with $E = E_i$ and $l = m = 0$, which have the form

$$\psi_i(r, \theta, \phi) = \psi_i(r)Y_{00}(\theta, \phi) = \psi_i(r). \quad (7.67)$$

(In other words, the eigenstates ψ_i are all spherically symmetric solutions.) We can thus write $\psi(r) = \sum_i c_i \psi_i(r)$ for some constants c_i such that $\sum_i |c_i|^2 = 1$.

$$\begin{aligned}
E &= \int_{r=0}^{\infty} \psi^*(r) \left(-\frac{\hbar^2}{2M} \left(\frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} \right) \psi(r) \right. \\
&\quad \left. + \left(\frac{\hbar^2}{2Mr^2} l(l+1) + V(r) \right) \psi(r) \right) \quad (7.68)
\end{aligned}$$

$$\begin{aligned}
&= \int_{r=0}^{\infty} \psi^*(r) \left(-\frac{\hbar^2}{2M} \left(\frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} + V(r) \right) \psi(r) \right. \\
&\quad \left. + \int_{r=0}^{\infty} \psi^*(r) \left(\frac{\hbar^2}{2Mr^2} l(l+1) \right) \psi(r) \right) \quad (7.69)
\end{aligned}$$

$$\begin{aligned}
&> \int_{r=0}^{\infty} \psi^*(r) \left(-\frac{\hbar^2}{2M} \left(\frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} + V(r) \right) \psi(r) \right) \\
&= \int_{r=0}^{\infty} \sum_i c_i \psi_i^*(r) \left(-\frac{\hbar^2}{2M} \left(\frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} + V(r) \right) \right) \sum_j c_j \psi_j(r). \quad (7.70)
\end{aligned}$$

Now this last term is

$$\sum_i c_i^* c_i E_i = \sum_i |c_i|^2 E_i. \quad (7.71)$$

Since we have that $E > \sum_i |c_i|^2 E_i$ and that $\sum_i |c_i|^2 = 1$, we must have that $E > E_i$ for at least one value of i . Hence E is not the lowest energy eigenvalue, in contradiction to our original assumption. ■

7.8 The Hydrogen atom

We can now obtain the general bound state solution for particles in the potential $V(r) = -\frac{e^2}{4\pi\epsilon_0 r}$.

As we did in obtaining spherically symmetric solutions, we define the quantities $a = \frac{e^2 M}{2\pi\epsilon_0 \hbar^2}$, $b = \frac{\sqrt{-2ME}}{\hbar}$. We obtain from Eqn. (7.65) the equation

$$\left(\frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} \right) \psi(r) + \left(-\frac{1}{r^2} l(l+1) + \frac{a}{r} \right) \psi(r) = b^2 \psi(r). \quad (7.72)$$

As we saw in discussing Eqn. (7.26), we see that the ansatz $\psi(r) \approx \exp(-br)$ means that the two terms which are largest asymptotically (the first term on the LHS and the term on the RHS) cancel. This again suggests trying an ansatz of the form $\psi(r) = f(r) \exp(-br)$, for a power series $f(r)$.

However, the new singular term $(\frac{1}{r^2} l(l+1))$ means that the previously obtained solutions are not generally valid.

It turns out to be convenient to write the power series in the form $f(r) = \sum_{n=0}^{\infty} a_n r^{n+\sigma}$, where σ , a constant to be determined, is chosen so that $a_0 \neq 0$: i.e. the power series begins with a term proportional to r^σ .

Considering the coefficient of $r^{\sigma-2}$, we have $-(\sigma(\sigma-1) + 2\sigma) + l(l+1) = 0$ or $\sigma(\sigma+1) = l(l+1)$, a quadratic equation with roots $\sigma = l$ and $\sigma = -(l+1)$. As $l \geq 0$, we choose $\sigma = l$ to avoid a divergence at $r = 0$.

We now have

$$a_n = \frac{(n+l)2b - a}{n(n+2l+1)} a_{n-1} \text{ for } n \geq 1. \quad (7.73)$$

As before, if the power series does not terminate this reduces to $a_n \approx \frac{2b}{n} a_{n-1}$ for large n , which would give us $f(r) \approx \exp(2br)$ and

$$\psi(r) \approx \exp(2br) \exp(-br) \approx \exp(br),$$

a divergent and unnormalisable wavefunction, which is physically unacceptable.

The power series must thus terminate, so we have $a = 2b(n+l)$, for some $n \geq 1$. Rewriting, we have $b = \frac{a}{2N}$ for some $N \geq l+1$, giving the same overall set of solutions for b , and thus the same energy levels (i.e. the Bohr energy levels), as the spherically symmetric case with $l=0$ we considered earlier:

$$E = -\frac{Me^4}{32\pi^2\epsilon_0^2\hbar^2} \frac{1}{N^2}.$$

7.8.1 Energy level degeneracies

Each value of N is consistent with

$$l = 0, 1, \dots, (N-1); \quad (7.74)$$

each value of l is consistent with

$$m = -l, -(l-1), \dots, l. \quad (7.75)$$

(The first of these degeneracies occurs only for a Coulomb force law; the second, as we have seen, holds for any central potential.) The total number of values of (m, l) consistent with N is thus

$$\sum_{l=0}^{N-1} \sum_{m=-l}^l 1 = \sum_{l=0}^{N-1} (2l+1) = 2\left(\frac{1}{2}N(N-1)\right) + N = N^2. \quad (7.76)$$

In fact, the true degeneracy of the N th energy level of the hydrogen atom in a full non-relativistic quantum mechanical treatment is $2N^2$: the extra factor of 2 arises from an intrinsically quantum mechanical degree of freedom, the electron *spin*, which has no direct classical analogue. (This is covered in the Part II Principles of Quantum Mechanics course.)

7.9 Towards the periodic table

We could try to generalize 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

$$\psi(\mathbf{x}_1, \dots, \mathbf{x}_Z) = \psi_1(\mathbf{x}_1) \dots \psi_Z(\mathbf{x}_Z), \quad (7.77)$$

where the ψ_j are rescaled solutions for the hydrogen atom (the rescaling is because the nucleus has charge $+Ze$ instead of $+e$). The energy is just the sum

$$E = \sum_{i=1}^Z E_i. \quad (7.78)$$

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.²⁶ 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, as above, we have $2N^2$ states in the N th energy level. This gives us an atom with a full energy level with $Z = 2, 10 = 8+2, \dots$ for $N = 1, 2, \dots$; these are the elements helium, neon, \dots . 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 (unreactive) 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.

²⁶See Part II Principles of Quantum Mechanics for details.