

0 Prologue

In these lecture notes, we will see how mounting experimental evidence at the beginning of the 20th century forced scientists to re-think their understanding of the world in terms of distinct wave and particle concepts, and to look for a unified description that encompasses both. This lead to the formulation of wave-particle duality and quantum mechanics.

We shall discuss, at a phenomenological level, some of the key concepts that one must introduce in order to formulate a new description of reality that encompasses the wave-particle duality required by quantum mechanics. We will focus in the first instance on the notion of wave function and its meaning in regards to physical properties of a system and the outcome of measurements.

We will introduce a wave equation that describes the time evolution of a wave function, and we will discuss its understanding in terms of energy and momentum operators. Moreover, we will introduce the concept of probability current and the corresponding continuity equation.

We shall begin our investigation of quantum mechanical systems by solving the Schrödinger equation for the case of 1D unbound particles in (piece-wise) constant potentials. We will discuss the physical significance of the results and remark on the difference between classical and quantum behaviour – encountering the exquisitely quantum mechanical concept of tunnelling.

We will continue the study by turning our attention to the behaviour of bound particles. In this context we

will discuss the 1D simple harmonic oscillator, and the notions of zero point energy and of parity of the wavefunction. A correspondence principle between classical and quantum mechanics will also be formulated.

In this first part of the notes, we will develop an initial understanding of quantum mechanics, starting from the fundamental physics questions that it aimed to address, and continuing to propose a formalism based on wave functions and the Schrödinger equation. We will then venture into solving a range of problems concerning the motion of a particle in 1D, encompassing piece-wise constant potentials and the simple harmonic oscillator. In the process, we will encounter and discuss several new and important concepts brought about by quantum mechanics.

Following this first part of the course, we will then move beyond the wave function formulation and introduce a new, more powerful description based on vector spaces and operators – a description that has come to take centre stage and is widely used in modern physics. In this language, we shall revisit the notion of position, momentum and energy operators, and discuss the concepts of eigenvalues and eigenvectors, of expectation values and the uncertainty principle. The full force of quantum mechanics only emerges when the tools of linear operators are brought to bear. Operator algebra is central to quantum mechanics. Its advanced forms pervade many concepts that are now regarded as ‘physical reality’.

We will demonstrate the power and importance of the operator formulation of quantum mechanics by considering again the simple harmonic oscillator, to be contrasted with the Schrödinger description in earlier chapters, leading to the discussion of ladder operators. The quantisation of the harmonic oscillator is central to quantum mechanics because it highlights the relationship between particle-like and wave-like behaviour. Ultimately,

it defines what one means by elementary excitations such as photons, phonons, magnons, etc., and it provides an elegant scheme for describing their behaviour and interactions. We will close this part of the notes by introducing the notion of density operator and by discussing how to form functions of operators.

Up to this point, we will have introduced two formulations of quantum mechanics, in terms of Schrödinger's equation and wave functions, and in terms of vector spaces and operators. In both cases, time is a parameter. For example, the operator algebra we will have discussed applies at an instant in time, and at every time. We can only calculate the probability of observing some outcome at a chosen time t .

Next, we will turn our attention to time evolution in quantum mechanics. We shall start from the so-called *Schrödinger picture*, where it is quantum states and not observables (operators) that change with time. We will introduce the concept of time translation operator and Ehrenfest's theorem describing the evolution of the expectation value of operators. We will also discuss the connection to classical mechanics and the time-energy uncertainty principle, and we will give an example involving a particle in an infinite well. We will also introduce the so-called *Heisenberg picture*, where quantum states do not change with time, but operators do; in many ways, this is closer to the classical viewpoint, where position and momentum are dynamical variables that vary with time in familiar ways. We will finally give an example involving the density operator.

In the lecture notes, we will have hitherto considered 1D systems, for simplicity. While this allows us to introduce some of the main concepts and formalisms used in quantum mechanics, it also has substantial limitations – for example, there is no notion of angular momentum in 1D. In order to progress in our understanding of quantum mechanics, we will have to expand our discussion to higher dimensions, and specifically to 3D systems.

Through the investigation of quantum mechanical systems in 3D, we will introduce the notion of angular momentum and we shall see how it plays a fundamental part in quantum mechanics. We will define the angular momentum operator and derive its properties, commutation and uncertainty relations, eigenstates and eigenvalues. We will encounter again ladder operators. And we will further discuss the connection between angular momentum and magnetic moment, and give detailed examples such as diatomic molecules and the Hydrogen-like atom.

This will conclude the part of the lecture notes dedicated to single particle mechanics. The rest of the course will be dedicated to multi-particle / multi-state systems. This will set the stage to discuss the notion of spin and total angular momentum, as well as the role of interactions and identical particles.

We will start going beyond single particle mechanics by considering a system composed of two particles, and how one can describe the states and their physical properties, and how to solve the quantum two-body problem. This will lead to the important concept of entangled states and spooky action at a distance.

We will introduce the notion of intrinsic angular momentum, or spin, of a particle. We shall see how one can combine the spin and orbital angular momentum into the total angular momentum, and how this can be

formulated in terms of eigenstates and eigenvalues, and composite wave functions. We will also study how to combine spins of different particles.

We will then encounter another iconic example of a quantum mechanical system that requires the combination of different states: the two level system. We will describe it using the vector and operator-matrix formalism and present a complete perturbative study. This will lead us to introduce the notions of overlap integrals and level splitting / repulsion, typical of covalent / ionic bonds between atoms. The notes will also include two asides that illustrate simple applications of the matrix representation to orbital and spin angular momentum, and two further non-examinable asides on spin matrices and density matrices.

Finally, we shall close the lecture notes by returning to the description of systems comprising more than one particle, paying particular attention to the situation where the particles are *identical*. We will encounter the concept of exchange symmetry and the spin statistics theorem, whereby particles at the quantum mechanical level can either be symmetric upon exchange (bosons) or antisymmetric (fermions). We will finish with a brief discussion of how this affects multi-particle states.

I The Quantum Revolution

We shall start these lecture notes by reviewing how mounting experimental evidence at the beginning of the 20th century forced scientists to re-think their understanding of the world in terms of distinct wave and particle concepts, and to look for a unified description that encompasses both. This lead to the formulation of wave-particle duality and quantum mechanics.

1 Background

At the beginning of the 20th century physics underwent a revolution. A revolution that changed the direction of science.

It was driven by the inability to account for the results of a succession of different experiments.

Physicists such as Bohr, Heisenberg, Einstein, Schrödinger, Born and others debated, and indeed argued about, the nature of reality, and by 1927 had created a mathematical framework that could account for the phenomena seen.

The principles of quantum mechanics were discussed at the 5th Solvay Conference, Fig. 1, which is regarded as one of the most influential conferences in physics. Of the 29 attendees, 17 were or became Nobel Prize winners.

The new formalism could have not been created from

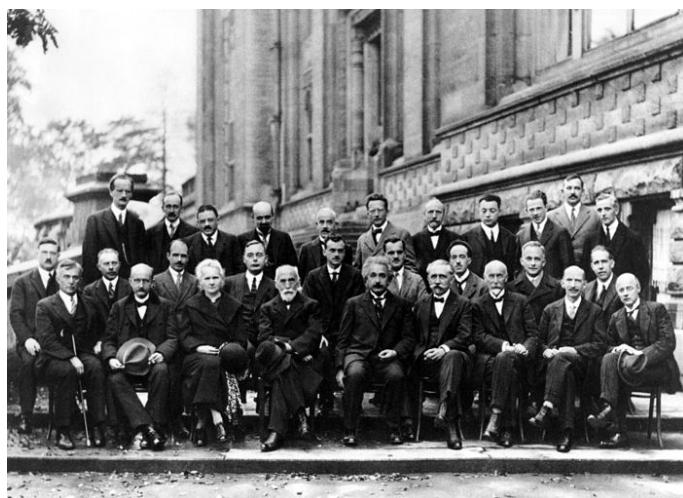


Figure 1: The 1927 Solvay Conference on Electrons and Photons.

prior theories by extending the rules, but needed fundamental new concepts.

The mathematical framework goes under the name of *quantum mechanics*. It has been remarkably resilient against numerous attempts to break it, both experimentally and theoretically.

In 1928, Dirac created a relativistic formulation, which ultimately led to the discovery of many more unexpected and counterintuitive phenomena, such as antimatter.

Measurement lies at the heart of quantum mechanics:

- The primary value of any theory lies in its ability to predict, either in a deterministic or probabilistic way, the outcome of a pertinent experiment. It is the predictive power of a physical theory that makes it valuable.
- Even at the theoretical and philosophical levels, the notion of ‘measurement’ is central to quantum mechanics.

When one talks about position, velocity, momentum, etc., one is referring to a quantity, actually a quality, that is determined by making an observation of some kind.

The position of a ball has no meaning unless one has some, possibly abstract, way of measuring position, and indeed defining what is meant by ‘measuring position’.

In quantum mechanics an observable, such as position, is not simply a scalar quantity, it is a mathematical operator that indicates that a measurement is being made.

The observer is not merely a passive bystander, but an active part of the process.

Measurement changes the future behaviour of a system; it is not merely that the value of the quantity is now known.

The order in which measurements are carried out can affect the results obtained.

A disconcerting aspect of quantum mechanics, which troubled the inventors themselves, is that the quantum state of a system, which is a mathematical object that encodes the likelihood of the possible outcomes of an experiment occurring, evolves in time deterministically, but when a measurement is made it changes discontinuously, undeterministically, and in an irreversible way, to take up the state associated with the observed outcome.

This apparent process is referred to as the *collapse of the wavefunction*, and it leads to a number of conceptual challenges and famous paradoxes.

The collapse of a state is compelling and needed because it ensures that the result of a second, identical measurement immediately after the first gives the same result with certainty.

The probabilistic nature of quantum mechanics does not correspond to the fact that one simply has a large collection of slightly different classical systems and incomplete knowledge about which system is being tested.

Probabilistic behaviour is intrinsic to measurements carried out on a single system.

As the description of a quantum state is enhanced to encompass an increasing range of possible physical observables, the mathematical description necessarily becomes more complicated.

Profoundly, a quantum state describes a complete system, not just the behaviour of individual particles.

For example, if a system comprises two electrons, then a measurement of the properties of one electron might well affect the outcome of a measurement performed on the other.

2 Outline of the lectures

During the next 24 lectures we shall learn the language of quantum mechanics, we shall build the mathematical tools necessary to use it, and we shall apply the principles to a range of physical phenomena and examples.

The basic approach will be as follows:

- (1) Understand why quantum mechanics is needed at all.

- (2) Establish a mathematical description of (the state of) a system. The state encodes the likelihood of the possible results of a measurement being recorded.
- (3) Establish mathematical operators that correspond to measurements of interest, and that encode the actual values that might be recorded when a measurement is made.
- (4) Understand how operators and states can be used to determine the likelihoods of outcomes, the distribution of possible outcomes, and the effects of measuring quantities in different orders.

These elements form the *Copenhagen interpretation* of quantum mechanics.

We shall begin with a brief overview of the experimental challenges that (classical) physics faced at the start of the 20th century. With building evidence of waves behaving as particles, and particles as waves, it became clear that a description of the physical world where these two concepts are clearly distinct was no longer tenable. A new description was needed, which lead to the birth of quantum mechanics.

We shall introduce quantum mechanics in the first instance by discussing, in a phenomenological way, how a system that behaves both as a wave and as a particle can be described. This will require the introduction of the concept of a wavefunction. We shall review some of its basic properties – partly familiar from other courses – and how it relates to measurements in quantum mechanics.

We will then propose a differential equation for the time evolution of the wavefunction – akin to Newton's equation governing the evolution of classical particles. We

will introduce Schrödinger's equation and study its solutions in various 1-dimensional cases of interest; this will allow us a first glimpse into the quantum world, illustrating concepts such as zero point energy and tunnelling that do not have a classical mechanical counterpart.

While Schrödinger's description of quantum mechanics has its values, an equivalent but mathematically more powerful description can be formulated in terms of vector spaces and operators. The next part of the course will turn to this, introducing its basic concepts and definitions, and re-building the quantum mechanics that we have seen so far in the so-called operator formalism. In this new language, we shall encounter again well known quantities such as position, momentum and energy – and their ‘measured’ expectation values – and we shall formulate more crisply the notion of uncertainty principle. We will finally apply the new description to the case of the simple harmonic oscillator – to contrast it with the way in which we understood it earlier in Schrödinger's formalism. We shall close this first step into quantum mechanics by revisiting the concept of time evolution in the operator formalism.

For the sake of simplicity, we hitherto restricted our discussion to the motion of a particle in 1-dimensional systems. While many of the concepts and formalism do carry over to higher dimensions, such simple choice suffers from important limitations. For example, there is no angular momentum in 1-dimensional motion. Some of the key concepts in modern quantum mechanics require us indeed to explore higher dimensions, and the lectures will turn to it next.

As we move into 3-dimensions, we will introduce the angular momentum operator and observe that its components do not commute and thence cannot be measured simultaneously to arbitrary accuracy! We will discuss its representation in terms of ladder operators, and its

significance in regards to magnetic properties of a quantum system. We will then be in the position to carry out a complete study a basic model for the Hydrogen atom – one of the most significant early achievements that asserted the validity and importance of quantum mechanics in describing our world.

The last part of the course will turn to a handful of advanced topics, which will provide further knowledge about quantum mechanics concepts and techniques, as well as give a prospective into the intricacies that one needs to deal with in order to develop a quantum mechanical description of systems beyond the simple cases studied in these introductory lectures:

- Quantum mechanics of multi-particle systems: two particle states, separability and entanglement
- Intrinsic angular momentum (spin) and combination of angular momenta
- Two level systems, double-well potentials and tunnelling
- Quantum mechanics of multi-particle systems: identical particles and exchange symmetry, spin statistics theorem, bosons and fermions, Pauli exclusion principle

3 Physics challenges of the early 20th century that lead to the birth of quantum mechanics

What was so perplexing about the experiments of the early 20th century?

By the end of the 19th century, two *distinct* kinds of behaviour accounted for most of physics: particles and waves.

The behaviour of macroscopic objects, from peas to planets, were known to be describable by the principles of classical mechanics, or for large collections, Boltzmann statistics.

Conversely, the behaviour of sound, and Maxwell's electromagnetic radiation, were known to be describable by waves.

This deceptively clear distinction become blurred when it was found that atoms and electrons display wavelike behaviour, such as diffraction and interference, and light displays particulate behaviour, such as the absorption and emission of radiation in the form of discrete packets of energy.

The particulate behaviour of light led to the notion of photons having energy $E = h\nu = \hbar\omega$ and momenta $p = h/\lambda = \hbar\omega/c$.

What was needed was a way of encompassing wavelike and particle-like behaviour in a single formalism, but this formalism took many years to emerge.

In the following subsection we shall review some of the key experiments and their interpretation.

We shall make extensive use of Planck's constant,

$$h = 6.626\,069\,57(29) \times 10^{-34} \text{ Js}, \quad (1)$$

which for notational convenience is often written

$$\hbar = \frac{h}{2\pi}, \quad (2)$$

or

$$\hbar = 1.054\,571\,726(47) \times 10^{-34} \text{ Js}, \quad (3)$$

This constant appears throughout quantum mechanics,

and essentially scales what would otherwise be a scale-free theory.

It can only be found experimentally.

3.1 Photoelectric effect

Consider the situation where a beam of monochromatic light is incident on a photo-electric surface, called the *cathode*.

The surface emits electrons, which are collected by a second electrode called the *anode*.

An external voltage source allows the potential between the electrodes to be held constant, and an ammeter records the flow of charge.

Arrangements of this kind can be understood using the energy diagram shown in Fig. 2.

Prior to illumination, the electrons in the metal have energies up to (essentially) a well-defined maximum, called the *Fermi energy*.

An electron exiting the metal leaves a positive charge behind, and thence some energy W , called the *work function*, is needed before the electron can escape.

As electrons traverse the system, an external current is generated, which indicates that electrons have been released and collected.

The external voltage source establishes an electric field between the plates, and the potential is a linear function of position, as shown in Fig. 2.

By sweeping the voltage from a positive to negative

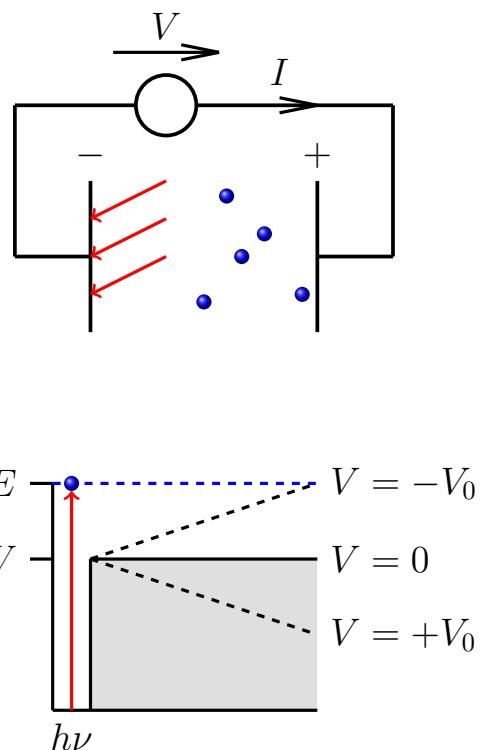


Figure 2: Photoelectric effect

value, the current can be turned off, because the released electrons do not have enough energy to make it to the anode.

The negative voltage that reduces the current to zero is a measure of the maximum kinetic energy of the electrons released.

How do the energy and rate of the emitted electrons vary with the intensity and frequency of the light?

Classically, the electric field in an optical beam increases with intensity (power flow), but does an increase in intensity create more photo-electrons having the same energy, or higher-energy photo-electrons at the same rate?

As the electric field strength is increased, the photo-electrons should be accelerated to higher energies.

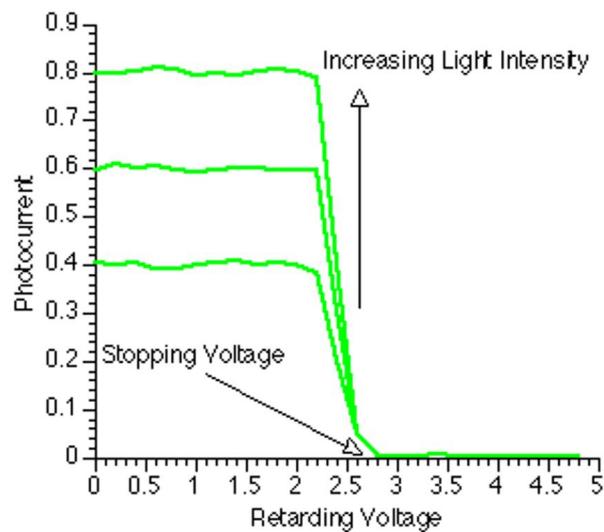


Figure 3: Electron energy is independent of intensity.

It is observed that the energy of the emitted photo-electrons does not depend on the intensity of the light, but the photo-current does: Fig. 3

It seems that each electron can only absorb a certain quantum of energy, whatever the intensity of the impinging radiation is. Greater intensity just gives more electrons having the same kinetic energy.

Conversely, increasing the frequency does increase the energy absorbed by each electron, but it does not increase the photo-current: Fig. 4.

No matter how faint the optical beam, some photo-electrons are always generated, and they always have the same energy.

The discrete packets of energy are called *photons*, and each one carries energy $h\nu$.

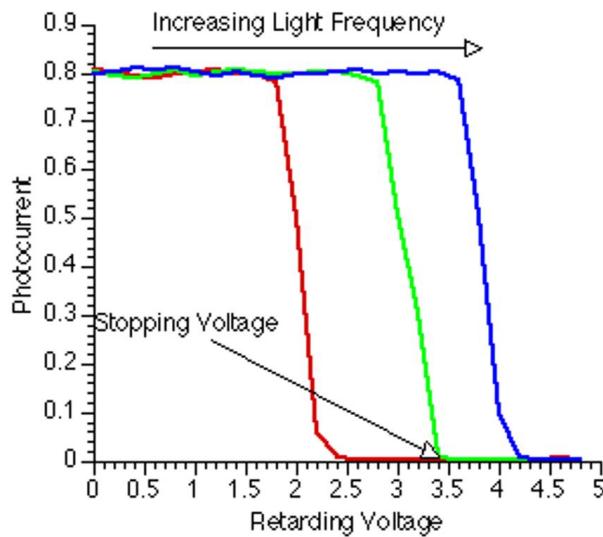


Figure 4: Electron energy depends on frequency.

Consider the energy diagram shown in Fig. 2:

$$\frac{1}{2}mv^2 = h\nu - W \quad (4)$$

$$eV_0 = h\nu - W,$$

$$V_0 = \frac{h\nu}{e} - \frac{W}{e}.$$

where V_0 is the *stopping potential*.

By plotting the stopping potential as a function of frequency, it is possible to measure the work function, W , and the ratio h/e : Fig. 5.

The slope of the line is independent of the material used.

Energy can only be extracted from an electromagnetic field in discrete amounts, or quanta, called photons:

$$E = h\nu = \hbar\omega$$

3.2 Blackbody radiation

A *black body* is an opaque object that emits thermal electromagnetic radiation. A perfect black body is one that

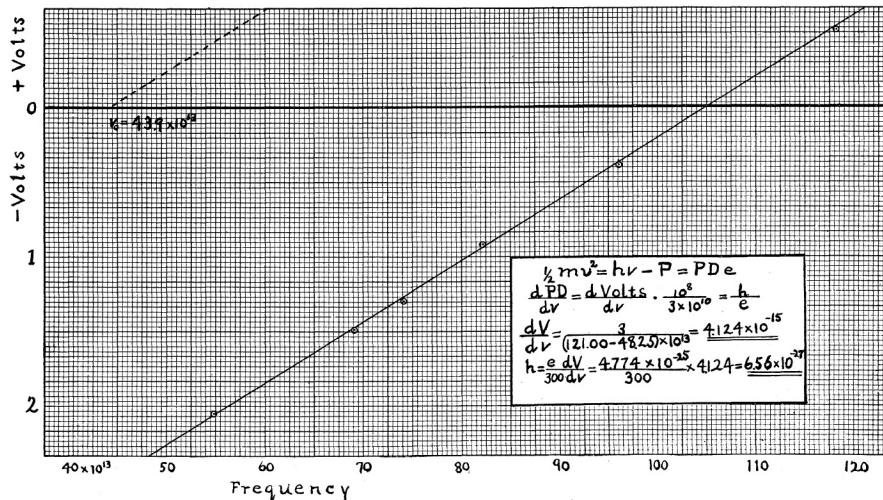


Figure 5: Millikan won the Nobel Prize in 1923 for experimentally proving Einstein's quantisation hypothesis. He was desperately trying to disprove it! The cut-off wavelength is 683 nm, corresponding to the work function of the sodium target used (R.A. Millikan, Phys. Rev. vol. VII, p. 355, 1916).

absorbs all incoming radiation and does not *reflect* any.

But why does the spectrum have the form shown in Fig. 6?

The spectrum is the same as the electromagnetic field in a cavity having temperature T , which supports a finite number of electromagnetic modes per unit volume, with frequency between ν and $\nu + d\nu$:

$$dn = N(\nu) d\nu. \quad (5)$$

It can be shown that

$$N(\nu) d\nu = \frac{8\pi\nu^2}{c^3} d\nu. \quad (6)$$

Derive (6) and explain how the derivation leads to the notion of energy per unit volume. How can the 'modes of a cavity' be used to describe free-space black-body radiation without a cavity?

Classically, the average energy per mode is $k_B T$, and so

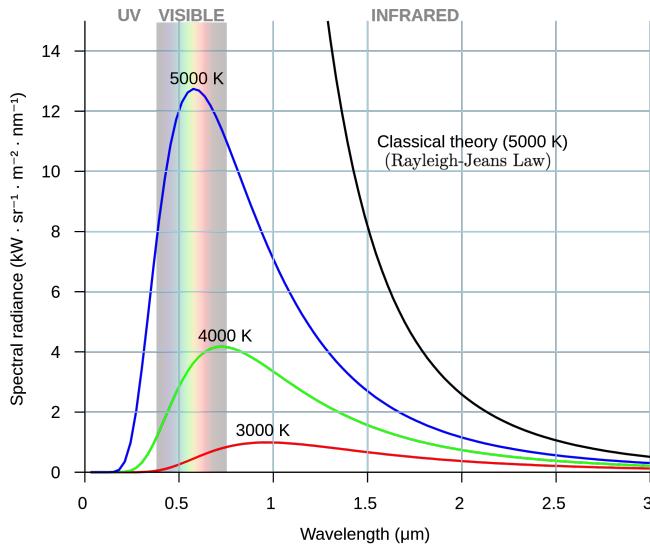


Figure 6: Spectral radiance of a black body (Wikipedia: black-body radiation).

the energy density is given by

$$\rho(\nu, T) d\nu = \frac{8\pi\nu^2}{c^3} k_B T d\nu, \quad (7)$$

which is the Rayleigh-Jeans Law.

In terms of wavelength

$$\rho(\lambda, T) d\lambda = \frac{8\pi}{\lambda^4} k_B T d\lambda. \quad (8)$$

Accordingly, $\rho(\lambda, T)$ increases without bound as the wavelength is reduced (or the frequency is increased).

This erroneous prediction is called the *ultra-violet catastrophe*.

Planck hypothesized that for each electromagnetic mode, at frequency ν , the energy is quantised into discrete states having energies $h\nu$: ie $h\nu, 2h\nu, 3h\nu, \dots$

The average energy $\bar{\epsilon}$ of each mode at frequency ν is then

$$\bar{\epsilon} = \frac{\sum_{n=0}^{\infty} nh\nu e^{-nh\nu/k_B T}}{\sum_{n=0}^{\infty} e^{-nh\nu/k_B T}}, \quad (9)$$

which evaluates to

$$\bar{\epsilon} = \frac{h\nu}{e^{h\nu/k_B T} - 1}. \quad (10)$$

Because of the Boltzmann factor, high-frequency states are unlikely to be occupied.

Then

$$\rho(\nu, T) = \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{h\nu/k_B T} - 1}, \quad (11)$$

which is called the *Planck radiation formula*.

The following is observed:

- The Planck radiation formula fits experimental data, avoiding the ultra-violet catastrophe. The energy density goes to zero as $\nu \rightarrow \infty$.
- The Rayleigh-Jeans formula is recovered as $\nu \rightarrow 0$.
- The classical result is obtained as $h \rightarrow 0$, which is a manifestation of the *correspondence principle*.

The Cosmic Microwave Background Radiation provides a profound example of black-body radiation: Fig. 7.

Once again, we find evidence that the energy in an electromagnetic mode must be an integer number of $h\nu$.

3.3 Parenthesis: the de Broglie hypothesis

In 1924, de Broglie hypothesized that *individual* particles have wave-like properties. The hypothesis is instrumental in the interpretation of the two experiments that we are about to discuss, and this is the reason why we review it beforehand in a separate subsection.

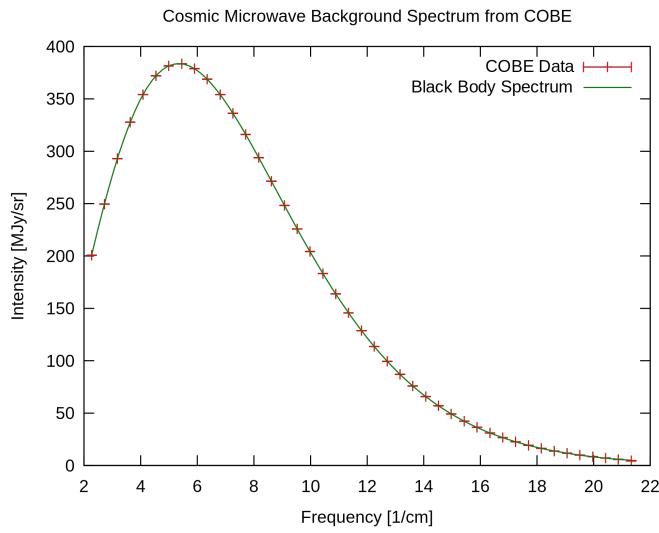


Figure 7: Cosmic microwave background radiation (Wikipedia: Cosmic Background Explorer).

For particles having momentum p , or as a vector \mathbf{p} ,

$$p = h/\lambda \quad (12)$$

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

where λ is called the *de Broglie wavelength*.

\mathbf{p} is written as a vector, in accordance with \mathbf{k} being the wavevector.

The association of momentum with wavelength has profound implications.

3.4 Bohr's atom

In 1913, Neils Bohr constructed a model of atomic hydrogen, which was based largely on classical theory, but which introduced the idea that angular momentum is quantised.

The nucleus of an atom is typically confined to a radius of about 10^{-14} m, as shown by α -particle scattering experiments, whereas the electrons typically extend out to

a distance of about 10^{-10} m.

If the electrons ‘orbit’ the nucleus, why don’t they radiate electromagnetic energy?

For a classical system the orbiting electrons would spiral into the nucleus emitting a near-continuous spectrum of electromagnetic radiation as they decay. It can be shown that this should occur on a timescale of 100 ps.

Why are atoms stable?

A possible explanation is that each electron can only exist in certain stable, discrete, orbits having definite energy. An electron can only exist in one of these states, and cannot move continuously between them.

When the electron changes from one state to a neighbouring state, a single photon having energy $h\nu$ is emitted, or absorbed.

This behaviour is explained by attributing wave-like behaviour to each electron, and requiring that in a stable orbit, the circumference (assuming a circular orbit) must be an integer number of de Broglie wavelengths:

$$2\pi r_n = n\lambda. \quad (14)$$

The angular momentum L in a stationary state is given by

$$L = m_e v_n r_n = p \frac{n\lambda}{2\pi} = n\hbar. \quad (15)$$

It seems that angular momentum is quantised, and this accounts for the discreteness of the orbits and for the stability of atoms.

Using Bohr’s model, it is straightforward to show that

the total energy of an orbit, kinetic and potential, is

$$\begin{aligned} E_n &= -\frac{e^2}{8\pi\epsilon_0 r_n} \\ &= -\frac{1}{n^2} \frac{m_e e^4}{8\epsilon_0^2 h^2} \\ &= -\frac{hcR}{n^2}, \end{aligned} \tag{16}$$

where R is called the *Rydberg constant*.

The behaviour of other molecules can be found by scaling by the atomic number Z .

Note that the energy levels are negative compared with the energy of a free particle.

This expression was derived empirically by Rydberg in 1889, well before the founding of quantum mechanics.

The theoretical value found this way is in good, but not perfect, agreement with the experimental value:

$R = 1.09737 \cdot 10^7 \text{ m}^{-1}$ for hydrogen.

Atomic emission-line spectra are explained by transitions between orbitals, $h\nu_{nm} = E_n - E_m$, giving

$$\begin{aligned} h\nu_{nm} &= \frac{hc}{\lambda_{nm}} \\ &= hcR \left(\frac{1}{n^2} - \frac{1}{m^2} \right), \end{aligned} \tag{17}$$

where n and m are integers: Fig. 8.

Through this same model, the Bohr radius can be derived for $n = 1$ giving a value of $5.29 \cdot 10^{-11} \text{ m}$.

Likewise, the velocity of the electron in the first Bohr orbit can be written $v = \alpha c$, where α is the fine structure constant, $1/137$.

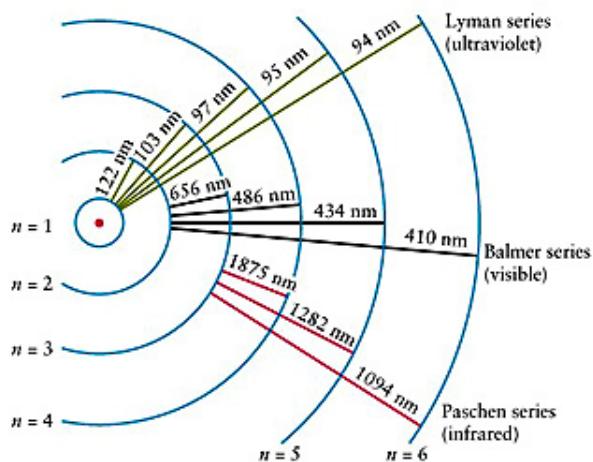


Figure 8: Energy levels that give rise to the discrete emission and absorption line spectra of hydrogen.

The stability of atoms can be understood if the angular momentum of an orbiting electron is quantised:

$$L = n\hbar.$$

n is called the *principal quantum number*.

3.5 Electron diffraction

The de Broglie hypothesis was confirmed by Davisson (who won the Nobel Prize 1937) and Germer, and independently by G.P. Thomson in 1927 by observing diffraction patterns associated with electrons.

Because the wavelengths are small, and the scattering object must have periodic structure on the scale size of the wavelength, crystalline arrays of atoms were used.

Davisson and Germer directed a beam of electrons at the [111] plane of crystalline nickel, for which $d = 2.15 \text{ \AA}$, and a detector was swept in angle to measure the direction in which most of the electrons were scattered: Fig. 9.

In accordance with Bragg reflection, maximum reflection occurs at the angle θ for which $d \sin \theta = n\lambda$.

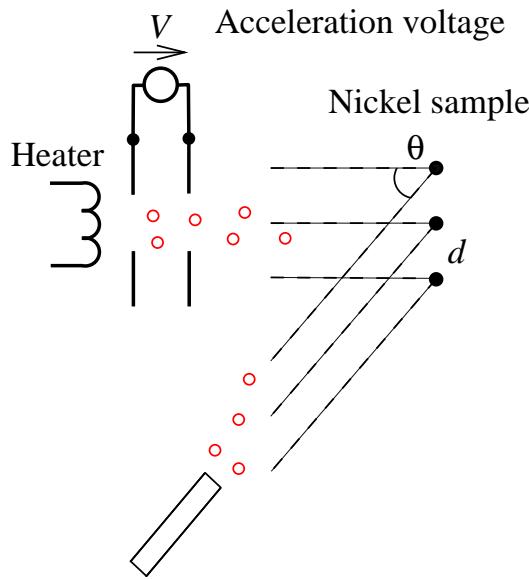


Figure 9: Davisson and Germer experiment.

For $n = 1$, it follows that

$$\frac{1}{\lambda} = \frac{p}{h} \quad \text{de Broglie} \quad (18)$$

$$= \frac{\sqrt{2eVm}}{h} \quad \text{classical momentum} \quad (19)$$

$$= \frac{1}{d \sin \theta} \quad \text{Bragg.} \quad (20)$$

This establishes a relationship between the accelerating voltage V and the angle for which the intensity is at a maximum.

$$\sin \theta = \frac{h}{d\sqrt{2eVm}}. \quad (21)$$

Davisson and Germer found that for a voltage of 54 V, corresponding to a de Broglie wavelength of 1.67 Å, the intensity was at a maximum for 50°, which corresponds to a lattice spacing of $d = 2.18$ Å.

In good agreement with the value known from x-ray measurements.

4 Planck's constant

In all the examples reviewed above, we have seen that the constant of proportionality

$$h = 6.626\,069\,57(29) \times 10^{-34} \text{ Js}, \quad (22)$$

plays a central role.

The dimensions of $\hbar = h/2\pi$,

$$[M][L]^2[T]^{-1}, \quad (23)$$

correspond to momentum \times length, or energy \times time or angular momentum.

Planck's constant is divided by wavelength to give linear momentum, multiplied by frequency to give energy, or simply used as a natural unit for angular momentum.

5 The quantum challenge

We have seen that a variety of different experiments reveal serious problems when trying to use classical physics to describe the microscopic world:

- The photoelectric effect shows that energy can only be extracted from an electromagnetic field in discrete amounts: $E = \hbar\omega$.
- Bohr's model of the atom suggests that in order to explain why electron orbits do not decay, angular momentum must be quantised in integer multiples of \hbar .
- The above two phenomena are not unrelated, because when an atom changes state, the electromagnetic radiation takes the form of a spectral line having a specific wavelength.

- The Davisson-Germer experiment shows that electrons display wavelike phenomena, and in particular that the characteristic wavelength is related to linear momentum through $p = \hbar k$ (or more generally $\mathbf{p} = \hbar \mathbf{k}$).
- Planck avoided the ultra-violet catastrophe in the spectrum of a black body source by requiring that electromagnetic energy is quantised in units of $\hbar\omega$.

These observations, and others, provided mounting evidence that physical processes cannot be divided strictly into waves, based on spatially continuous fields, and particles, based on discrete point-like objects, but that every physical process displays both wave-like and particle-like behaviour.

How can wave-like and particle-like behaviour be combined into a single mathematical framework?

This was the quantum challenge.

II Wavefunctions

In the previous chapter we have seen how mounting experimental evidence at the beginning of the 20th century forced scientists to re-think their understanding of the world in terms of distinct wave and particle concepts, and to look for a unified description that encompasses both. This lead to the formulation of wave-particle duality and quantum mechanics. In this chapter we shall begin to lay its foundations by discussing – phenomenologically – how such quantum mechanical states can be formulated in terms of wavefunctions, and how the latter are to be understood to describe a physical system.

6 Quantum interference

Young's double-slit experiment brings to the fore certain key features of quantum mechanics: Fig.10 and Fig.11.

We know from optics that when a coherent source of radiation from a slit at S1 illuminates a pair of parallel slits at S2, an interference fringe is formed on a screen at F.

When the experiment is carried out using low-intensity light, such that individual photons can be detected, a number of remarkable observations are made:

- Initially, individual photons are detected at random positions, without a discernible pattern. After a period of time, when a large number of photons have been collected, a diffraction pattern emerges.

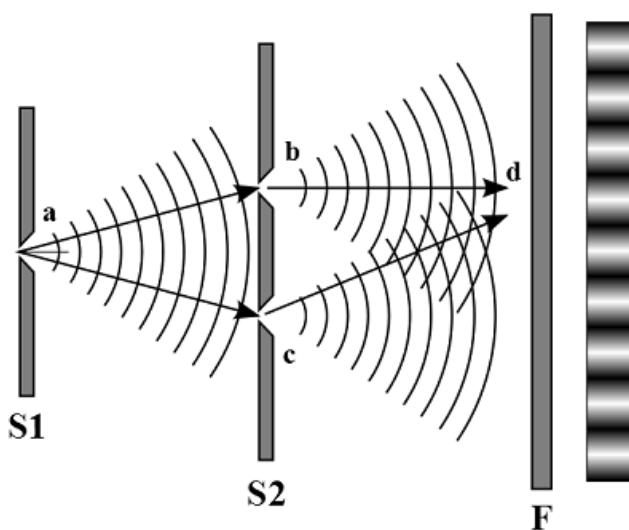


Figure 10: Young's double slit experiment.

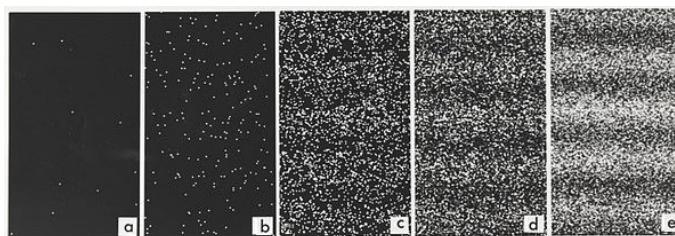


Figure 11: Interference fringes.

The diffraction pattern has the form that would be predicted on the basis of classical electromagnetism.

- If the experiment is repeated, individual photons are detected at different random positions, but nevertheless they build up the same overall diffraction pattern.
- If the distance between the slits is changed, a similar behaviour occurs, but with a different distribution.
- Even for extremely low light levels, with very long time intervals between events, the same behaviour is seen.
- It seems as if there is an underlying probability distribution, which interferes in the manner of waves, and which determines the likelihood of a photon being detected at a particular location.

Does it make sense to speculate about which slit a photon has passed through prior to a detection being made? Does a photon exist as a localised entity until it has been detected?

The notion that probability functions can interfere in a wavelike manner is very peculiar, but experiments compel us to a mathematical model of this kind.

It is necessary to introduce *wavefunctions* that encode information about the probability of achieving certain outcomes when an observation is made.

This idea pervades the whole of quantum mechanics.

By the *state of a system* we mean that we know the form, both in space and time, of the underlying distribution that describes the probabilities of achieving possible outcomes when an observation is made.

Because of the double-slit experiment, and indeed every classical electromagnetic scattering problem, one might expect that a plane wave is one possible form of wavefunction.

In one dimension

$$\Psi(x, t) = A e^{i(kx - \omega t)}. \quad (24)$$

The wavefunction has amplitude and phase, giving rise to interference, but a probability density must be real valued.

In the case of a single particle, and for all wavefunctions,

$$P(x, t) dx = |\Psi(x, t)|^2 dx \quad (25)$$

is the probability of detecting the particle in the region of width dx at x , at time t .

This interpretation requires that the wavefunction has been normalised:

$$\int_{-\infty}^{+\infty} P(x, t) dx = \int_{-\infty}^{+\infty} |\Psi(x, t)|^2 dx = 1, \quad (26)$$

at all times.

Presumably, the particle must exist for all time, and then one will certainly find it somewhere.

The wavefunction evolves in a causal manner if the system is left alone, but the outcome of any particular experiment, such as detecting the presence of a particle at a particular location, is intrinsically uncertain.

One might invoke the existence of ‘matter waves’ having the same functional form, but now taking into account the observed relationships between momentum and wavelength, and energy and frequency.

For example,

$$\Psi(x, t) = A e^{i(px - Et)/\hbar}. \quad (27)$$

In this case,

$$P(x, t) = |\Psi(x, t)|^2 = |A|^2 \quad (28)$$

and a detection is equally likely everywhere at all times.

In the case of ‘matter waves’, the quantum mechanical interpretation is the same: the wavefunction encodes the likelihood of finding the object at some location when a measurement is made.

In classical mechanics, the state of a system is defined once the position and momentum of the particle are known; in quantum mechanics, the state of a system involves position and momentum in an indirect way.

The wavefunction is an intrinsically complex quantity: unlike in classical physics where one always takes the real part of a complex exponential in order to represent a real quantity, such as voltage, displacement, pressure etc.

To be describable by a travelling wave, the wavevector and the frequency must be known, single-valued, and fixed. Thus, the linear momentum and kinetic energy must be known.

A travelling wave represents a free particle whose momentum and energy are known precisely, but whose position is completely unknown.

For a non-relativistic particle,

$$E = \frac{1}{2}mv^2 = \frac{p^2}{2m} = \frac{\hbar^2 k^2}{2m}, \quad (29)$$

but because $E = \hbar\omega$,

$$\omega = \frac{\hbar k^2}{2m}, \quad (30)$$

which shows that, in the case of a free particle, the wavefunction has a quadratic dispersion relationship.

More properly, in three dimensions

$$\Psi(\mathbf{r}, t) = A e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}, \quad (31)$$

where \mathbf{k} is the wave vector, \mathbf{r} a general position vector with respect to the phase reference point, ω the angular frequency, and t time.

In terms of classical mechanical properties

$$\Psi(\mathbf{r}, t) = A e^{\frac{i}{\hbar}(\mathbf{p} \cdot \mathbf{r} - Et)}. \quad (32)$$

Diffraction and interference have been observed using many different kinds of particles, including macromolecules: Fig. 12.

Figure 13 shows interference patterns generated us-

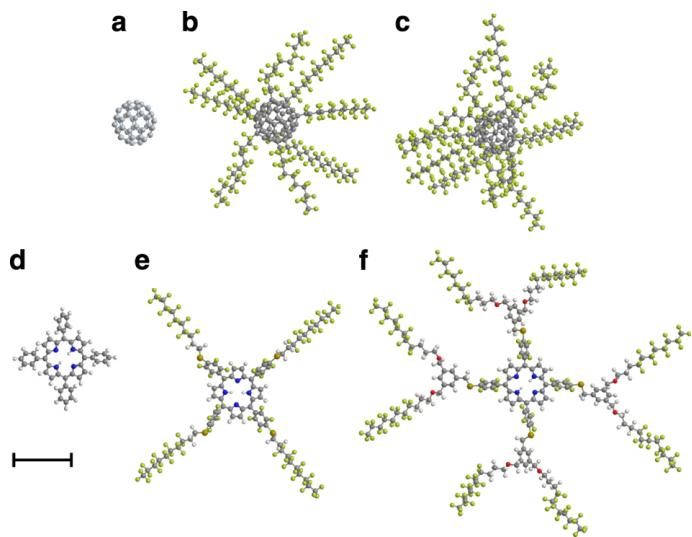


Figure 12: Interference phenomena have been seen with complex molecules (Nature Communications, vol. 2, p. 263, 2011).

ing molecules comprising up to 430 atoms, with sizes of 60 Å, masses of up to 6,910 AMU giving de Broglie wavelengths, $\lambda = h/mv$, as small as 1 pm.

The wavelike behaviour of matter is not just a feature of fundamental particles, it is an intrinsic feature of all material bodies.

7 Local wavefunctions

The travelling plane-wave form is not square integrable, and therefore cannot be normalised to give the probability of finding the particle somewhere as being unity.

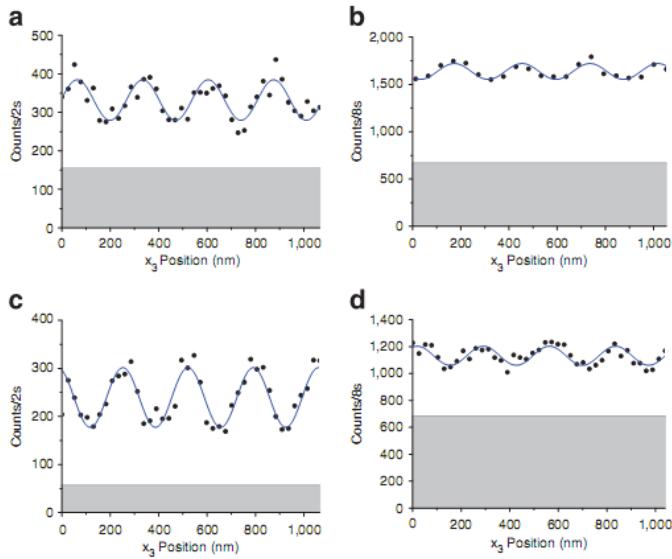


Figure 13: Interference fringes observed with the complex molecules shown in Fig. 12 .

Strictly, it is not a ‘meaningful’ wavefunction, but it is the limiting case of something that is meaningful.

If something is known about the position of a particle, $\Psi(\mathbf{r}, t)$ must to some extent be localised: $|\Psi(\mathbf{r}, t)| \rightarrow 0$ as $|\mathbf{r}| \rightarrow \pm\infty$, for every t .

The region of localisation may, however, change with time.

We have not yet discussed what differential equation the wavefunction must obey, but for a linear differential equation, if a traveling wave is a solution, then a weighted linear combination of traveling waves is also a solution.

A localised wavefunction can be constructed through

$$\Psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(k) e^{i(kx - \omega t)} dk, \quad (33)$$

where one should remember that ω is a function of k .

At $t = 0$, this corresponds to the Fourier integral:

$$\Psi(x, 0) = \Psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(k) e^{ikx} dk. \quad (34)$$

From Fourier's theorem it follows that

$$g(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \Psi(x) e^{-ikx} dx. \quad (35)$$

For the Fourier integral to exist in the ordinary sense, $\Psi(x)$ must be square integrable, which it is, and therefore $g(k)$ can be found.

If, at $t = 0$, $\Psi(x)$ fully describes the state of a system, then so does $g(k)$ because it can be calculated without additional information.

Eq. (33) may seem odd, because the wavefunction is no longer associated with a single, well-defined momentum.

$|g(k)|^2 dk$ is the probability that a measurement of the particle's wavenumber (momentum) will yield a value in the range dk at k .

$g(k)$ is called the *momentum wavefunction*.

The Fourier transform of a Gaussian function is a Gaussian, and therefore it can be used to indicate the precision with which position and momentum can be measured 'simultaneously'.

Consider the momentum wavefunction

$$g(k) = \left(\frac{a^2}{\pi} \right)^{1/4} e^{-a^2(k-k_0)^2/2}, \quad (36)$$

which, when used in Eq. (34), leads to a Gaussian-weighted linear superposition of travelling waves centred

at k_0 .

The prefactor normalises the wavefunction in k -space:

$$\int_{-\infty}^{+\infty} |g(k)|^2 dk = \left(\frac{a^2}{\pi}\right)^{1/2} \int_{-\infty}^{+\infty} e^{-a^2(k-k_0)^2} dk. \quad (37)$$

Using the substitution $u = a(k - k_0)$,

$$\int_{-\infty}^{+\infty} |g(k)|^2 dk = \left(\frac{1}{\sqrt{\pi}}\right) \int_{-\infty}^{+\infty} e^{-u^2} du = 1. \quad (38)$$

Now calculate the resulting form of the position wavefunction:

$$\Psi(x) = \frac{1}{\sqrt{2\pi}} \left(\frac{a^2}{\pi}\right)^{1/4} \int_{-\infty}^{+\infty} e^{-a^2(k-k_0)^2/2} e^{ikx} dk. \quad (39)$$

Write the exponent as

$$-\frac{a^2}{2} (k - k_0)^2 + ikx = -\frac{a^2}{2} \left(k - k_0 - \frac{ix}{a^2}\right)^2 + ik_0x - \frac{x^2}{2a^2}, \quad (40)$$

such that

$$\Psi(x) = \frac{1}{\sqrt{2\pi}} \left(\frac{a^2}{\pi}\right)^{1/4} e^{ik_0x} e^{-\frac{x^2}{2a^2}} \int_{-\infty}^{+\infty} e^{-\frac{a^2}{2}(k-k_0-\frac{ix}{a^2})^2} dk. \quad (41)$$

Using the substitution

$$u = \frac{a}{\sqrt{2}} \left(k - k_0 - \frac{ix}{a^2}\right); \quad du = \frac{a}{\sqrt{2}} dk \quad (42)$$

gives

$$\Psi(x) = \frac{1}{a\sqrt{\pi}} \left(\frac{a^2}{\pi}\right)^{1/4} e^{ik_0x} e^{-\frac{x^2}{2a^2}} \int_{-\infty}^{+\infty} e^{-u^2} du \quad (43)$$

or

$\Psi(x) = \left(\frac{1}{\pi a^2}\right)^{1/4} e^{ik_0x} e^{-\frac{x^2}{2a^2}}$

(44)

which is already normalised.

The wavefunction comprises a plane-wave ‘carrier’ having wave number k_0 modulated by a Gaussian envelope centred at $x = 0$.

The momentum and position wavefunctions at $t = 0$ are shown in Fig. 14.

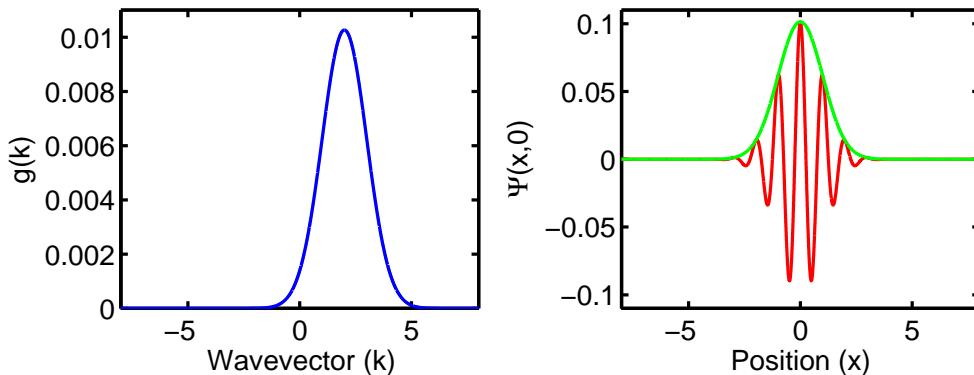


Figure 14: (Left) Gaussian momentum wavefunction $g(k)$ of a free particle. (Right) Real part of the associated position wavefunction $\Psi(x)$ at $t = 0$. The fringes travel at the phase velocity associated with the central wavevector k_0 . The green line shows $|\Psi(x)|$, but remember that the probability density is given by the square magnitude.

At $t = 0$, the square magnitude of the momentum wavefunction takes the value e^{-1} when $|k - k_0| = 1/a$, whereas the square magnitude of the position wavefunction takes the value e^{-1} when $|x| = a$.

As the position wavefunction is constrained more tightly, the momentum wavefunction widens.

Given that an infinite range of momenta is not physically possible, it will never be possible to measure the position of a particle with infinite precision.

8 Measurements

Consider what is meant by ‘simultaneous measurement’.

The absolute value squared of a wavefunction gives the probability of a measurement. Suppose that the position of a particle is measured, locating it within a region Δx . If the particle is measured again, immediately afterwards, the wavefunction must have changed, because it is now known that the particle lies within the region Δx with certainty!

Is this the same as the particle being ‘disturbed’ mechanically by the measurement apparatus? If this is simply about the apparatus disturbing the system, would we need quantum mechanics at all, or would classical mechanics be sufficient?

If the position is known precisely, the wavefunction must have collapsed to a delta function merely as a consequence of the measurement.

The process by which the wavefunction changes as a consequence of a measurement is called the *collapse of the wavefunction*.

It is controversial because the wavefunction changes abruptly when a measurement is made, and is not described by a differential equation that predicts a continuous evolution of the state.

The most common, contemporary interpretation is that collapse is a manifestation of a thermodynamically irreversible interaction with the classical environment that constitutes the experiment.

Observations are often used as a way of *preparing* a wavefunction for a later experiment. This idea is needed to define what is meant by a *simultaneous measurement*.

It is seen that the order in which measurements are made

can lead to very different outcomes as a consequence of the ‘disturbing effects’ of the measurements.

If the position of a particle is measured and then the momentum, one will get different answers to the case where the momentum is measured and then the position.

It is not possible to repeat a measurement in the same way, unless the state is again preprepared in its original form.

State preparation is a major consideration in experimental quantum physics.

Ordering is exceedingly important in quantum mechanics, and in due course we will need a non-commutative algebra to describe the act of measurement.

If a sequence of measurements is carried out, to determine the probability distribution associated with some state, then the system must be prepared in that state before each measurement is carried out.

If two quantities, say A and B, can be measured ‘simultaneously’, then the act of measuring A does not have any influence on the likely outcome of a subsequent measurement of B; likewise, the act of measuring B does not have any influence on the likely outcome of a subsequent measurement of A.

In this case, the order in which the measurements are made does not matter.

9 Expectation values and uncertainty

Once the position wavefunction is known, it is possible to calculate the expectation value of position, or indeed the expectation value of any function of position.

Physically, this corresponds to measuring the position multiple times, *each time ensuring that the system is preprepared in the same state*, and averaging the results.

Thus

$$\langle x \rangle = \int_{-\infty}^{+\infty} x P(x) dx = \int_{-\infty}^{+\infty} x |\Psi(x)|^2 dx. \quad (45)$$

Likewise,

$$\langle x^2 \rangle = \int_{-\infty}^{+\infty} x^2 P(x) dx = \int_{-\infty}^{+\infty} x^2 |\Psi(x)|^2 dx. \quad (46)$$

Define the *width* of the distribution according to the standard deviation,

$$\Delta x^2 = \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2, \quad (47)$$

or

$$\boxed{\Delta x = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}.} \quad (48)$$

Eq. (48) is better than using e^{-1} points because it applies to any reasonable, localised wavefunction.

It quantifies the quantum uncertainty in a measurement: the spread of the likely outcomes.

Calculate this quantity for the Gaussian wave-packet, using Eq. (44),

$$\langle x \rangle = \left(\frac{1}{\pi a^2} \right)^{1/2} \int_{-\infty}^{+\infty} x e^{-(x^2/a^2)} dx = 0 \quad (49)$$

$$\langle x^2 \rangle = \left(\frac{1}{\pi a^2} \right)^{1/2} \int_{-\infty}^{+\infty} x^2 e^{-(x^2/a^2)} dx = \frac{a^2}{2},$$

and therefore

$$\Delta x = \frac{a}{\sqrt{2}}. \quad (50)$$

Likewise using Eq. (36)

$$\langle k \rangle = \left(\frac{a^2}{\pi} \right)^{1/2} \int_{-\infty}^{+\infty} k e^{-a^2(k-k_0)^2} dk = k_0 \quad (51)$$

$$\langle k^2 \rangle = \left(\frac{a^2}{\pi} \right)^{1/2} \int_{-\infty}^{+\infty} k^2 e^{-a^2(k-k_0)^2} dk = \frac{1}{2a^2} + k_0^2.$$

and therefore

$$\Delta k = \frac{1}{a\sqrt{2}} = \frac{1}{2\Delta x}. \quad (52)$$

Thus

$$\Delta x = \frac{a}{\sqrt{2}} \text{ and } \Delta p = \frac{\hbar}{a\sqrt{2}}. \quad (53)$$

Combining these results,

$\Delta x \Delta p = \frac{\hbar}{2},$

(54)

which is independent of a .

In summary:

- To localise a particle to within a range Δx appears to introduce an uncertainty in its momentum of Δp .
- We have shown that $\Delta x \Delta p = \hbar/2$ for a Gaussian wavepacket, but for wavefunctions in general we will see that $\Delta x \Delta p \geq \hbar/2$.
- The Gaussian wavepacket is a state of *minimum uncertainty*, meaning that the product of the individual uncertainties takes on its smallest possible value: $\hbar/2$.

In three dimensions, Heisenberg's Uncertainty Principle becomes

$$\boxed{\begin{aligned}\Delta x \Delta p_x &\geq \hbar/2 \\ \Delta y \Delta p_y &\geq \hbar/2 \\ \Delta z \Delta p_z &\geq \hbar/2\end{aligned}} \quad (55)$$

Note that $\Delta x \Delta p_y = 0$, etc., and so spatially orthogonal quantities can, in principle, be measured simultaneously to any precision.

As we shall see, Heisenberg's Uncertainty Principle applies to other pairs of observables.



Werner Heisenberg
(1901-1976)

10 Time evolution of wave packets

According to Eq. (33), a general wavefunction can be constructed from a weighted linear combination of 'plane waves':

$$\Psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(k) e^{i(kx - \omega t)} dk. \quad (56)$$

Each of the elemental travelling waves is associated with some dispersion relation $\omega(k)$.

Consider the situation where the weighting function $g(k)$ is highly localised in k -space, and use Eq. (56) to calculate the time evolution of $\Psi(x, t)$.

This situation gives rise to what is known as a *wave packet*.

But what is meant by 'highly localised in k -space'?

Figure 15 shows the quadratic dispersion relationship of a free particle; $g(k)$ is also marked.

The effect of localising $g(k)$ depends on how much the dispersion relation changes over the region for which $g(k)$ is appreciably non-zero.

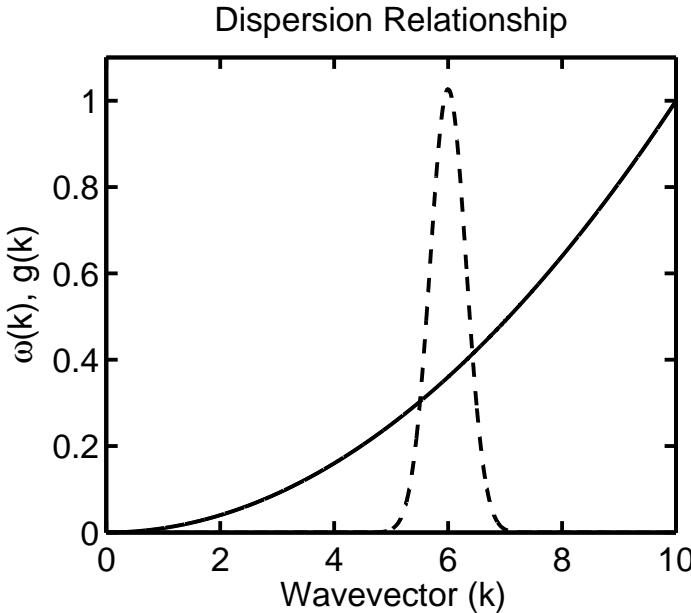


Figure 15: (Solid) The dispersion relationship is a plot of angular frequency against wavevector, $\omega(k)$. (Dashed) The momentum wavefunction is the Fourier-component weighting function in the k domain. The degree of dispersion depends on the curvature of the dispersion relationship over the region of the weighting function.

The analysis can be formalised as follows.

Let ω_0 and k_0 be the central values of ω and k in the wave packet. For a given elemental wave (ω, k) ,

$$\begin{aligned} k &= k_0 + \delta k \\ \omega(k) &= \omega(k_0) + \delta\omega(k), \end{aligned} \tag{57}$$

and then

$$\begin{aligned} \Psi(x, t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(k) e^{i[(k_0 + \delta k)x - (\omega_0 + \delta\omega(k))t]} d(\delta k) \\ &= \underbrace{e^{i(k_0 x - \omega_0 t)}}_{\text{Carrier}} \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(k) e^{i[\delta k x - \delta\omega(k)t]} d(\delta k)}_{\text{Envelope}}, \end{aligned} \tag{58}$$

which comprises a travelling wave multiplied by an envelope.

Consider the expansion

$$\begin{aligned}\omega(k) &= \omega(k_0) + \frac{d\omega}{dk} \Big|_{k_0} \delta k + \frac{1}{2} \frac{d^2\omega}{dk^2} \Big|_{k_0} (\delta k)^2 + \dots \quad (59) \\ \delta\omega &= \frac{d\omega}{dk} \Big|_{k_0} \delta k + \frac{1}{2} \frac{d^2\omega}{dk^2} \Big|_{k_0} (\delta k)^2 + \dots.\end{aligned}$$

If δk is sufficiently small that the second-order and higher terms can be ignored,

$$\begin{aligned}\delta\omega &= \frac{d\omega}{dk} \Big|_{k_0} \delta k \quad (60) \\ &= v_g \delta k,\end{aligned}$$

giving

$$\begin{aligned}\Psi(x, t) &= e^{ik_0(x-v_p t)} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(k) e^{i\delta k(x-v_g t)} d(\delta k), \\ &= e^{ik_0(x-v_p t)} f(x - v_g t),\end{aligned}\tag{61}$$

and the envelope simply shifts along the x -axis with a constant group velocity v_g .

The phase velocity, $v_p = \omega_0/k_0$, characterises the rate of progression of the carrier wave's phase front.

The group velocity, $v_g = (d\omega/dk)|_0$, characterises the rate of progression of the envelope.

If ω/k is a constant, every Fourier component travels at the same velocity, and every wavefunction propagates at the same velocity.

If $\omega(k)$ is linear over the spectral range of the wave

packet, the wavefunction does not spread as it propagates. Its envelope propagates at the group velocity $(d\omega/dk)|_0$: Fig. 16.

In anomalous dispersion, the phase and group velocities have different signs, corresponding to different directions of travel.

If the quadratic term in Eq. (59) is non-zero, the wave packet changes shape as it propagates: it *disperses*.

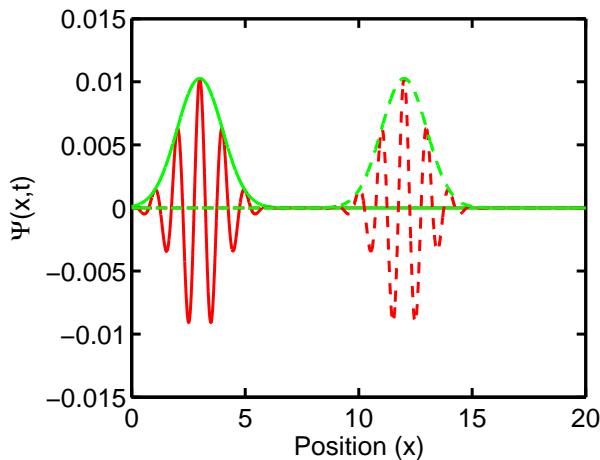


Figure 16: If the dispersion relationship is linear, the wavepacket does not spread as it propagates. The envelope travels at the group velocity v_g , and the underlying wave travels at the phase velocity v_p . These may or may not be equal, and can even have different signs.

In crystals, an electron's dispersion relationship can have regions where its slope tends to zero. How does the electron's wavefunction change with time for wave packets having spectral content that spans these regions. What does this imply for a measurement of the electron's position?

11 Time evolution of a Gaussian wavepacket

Consider the specific case of a Gaussian wave packet:

$$\Psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(k) e^{i[kx - \omega(k)t]} dk, \quad (62)$$

where

$$g(k) = \left(\frac{a^2}{\pi} \right)^{1/4} e^{-a^2(k-k_0)^2/2}, \quad (63)$$

giving, as shown in Eq. (44),

$$\Psi(x, 0) = \left(\frac{1}{\pi a^2} \right)^{1/4} e^{-\frac{x^2}{2a^2}} e^{ik_0 x}. \quad (64)$$

But how does this wave packet evolve with time?

For a free particle,

$$E = \frac{p^2}{2m} = \frac{\hbar^2 k^2}{2m} = \hbar\omega, \quad (65)$$

and

$$\omega = \frac{\hbar k^2}{2m}, \quad (66)$$

which is the dispersion relationship.

Now

$$\begin{aligned} \omega(k) &= \omega(k_0) + \frac{d\omega}{dk} \Big|_{k_0} \delta k + \frac{1}{2} \frac{d^2\omega}{dk^2} \Big|_{k_0} (\delta k)^2 + \dots \quad (67) \\ &\approx \omega_0 + \alpha \delta k + \beta (\delta k)^2, \end{aligned}$$

(which is exact for quadratic – free particle – dispersion).

Hence

$$\alpha = \frac{d\omega}{dk} \Big|_{k_0} = \frac{\hbar k_0}{m} = \frac{p_0}{m}, \quad (68)$$

and

$$\beta = \frac{1}{2} \frac{d^2\omega}{dk^2} \Big|_{k_0} = \frac{\hbar}{2m}. \quad (69)$$

Now evaluate the integral:

$$\begin{aligned}\Psi(x, t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(k) e^{i[kx - \omega(k)t]} dk, \\ &= \frac{1}{\sqrt{2\pi}} \left(\frac{a^2}{\pi}\right)^{1/4} \int_{-\infty}^{+\infty} e^{-a^2\delta k^2/2} e^{i[(k_0 + \delta k)x - (\omega_0 + \alpha\delta k + \beta(\delta k)^2)t]} d(\delta k), \\ &= e^{i(k_0 x - \omega_0 t)} \frac{1}{\sqrt{2\pi}} \left(\frac{a^2}{\pi}\right)^{1/4} \int_{-\infty}^{+\infty} e^{-(a^2/2 + i\beta t)\delta k^2} e^{i(x - \alpha t)\delta k} d(\delta k).\end{aligned}\quad (70)$$

Use the substitutions $X = x - \alpha t$ and $Y = a^2/2 + i\beta t$ to get

$$\begin{aligned}\Psi(x, t) &= e^{i(k_0 x - \omega_0 t)} \frac{1}{\sqrt{2\pi}} \left(\frac{a^2}{\pi}\right)^{1/4} \int_{-\infty}^{+\infty} e^{-Y\delta k^2} e^{iX\delta k} d(\delta k), \\ &= e^{i(k_0 x - \omega_0 t)} \frac{1}{\sqrt{2\pi}} \left(\frac{a^2}{\pi}\right)^{1/4} e^{-X^2/4Y} \int_{-\infty}^{+\infty} e^{-Y(\delta k - iX/2Y)^2} d(\delta k),\end{aligned}\quad (71)$$

and using $u = \sqrt{Y}(\delta k - iX/2Y)$ and $du = \sqrt{Y}d(\delta k)$ gives

$$\begin{aligned}\Psi(x, t) &= e^{i(k_0 x - \omega_0 t)} \frac{1}{\sqrt{2\pi}} \left(\frac{a^2}{\pi}\right)^{1/4} e^{-X^2/4Y} \frac{1}{\sqrt{Y}} \int_{-\infty}^{+\infty} e^{-u^2} du, \\ &= e^{i(k_0 x - \omega_0 t)} \frac{1}{\sqrt{2\pi}} \left(\frac{a^2}{\pi}\right)^{1/4} e^{-X^2/4Y} \sqrt{\frac{\pi}{Y}}, \\ &= \left(\frac{a^2}{4\pi}\right)^{1/4} \sqrt{\frac{1}{Y}} e^{-X^2/4Y} e^{i(k_0 x - \omega_0 t)}.\end{aligned}\quad (72)$$

Finally, substituting back in for X and Y , and using $\alpha = \hbar k_0/m$ and $\beta = \hbar/2m$, gives

$$\begin{aligned}\Psi(x, t) &= \left(\frac{a^2}{4\pi}\right)^{1/4} \frac{1}{(a^2/2 + i\hbar t/2m)^{1/2}} \exp\left\{\frac{-(x - \hbar k_0 t/m)^2}{4(a^2/2 + i\hbar t/2m)}\right\} \\ &\quad \exp\{i(k_0 x - \omega_0 t)\}\end{aligned}\quad (73)$$

and

$$|\Psi(x, t)|^2 = \left(\frac{a^2}{4\pi}\right)^{1/2} \frac{1}{(a^4/4 + \hbar^2 t^2/4m^2)^{1/2}} \exp\left\{\frac{-(x - \hbar k_0 t/m)^2}{(a^2 + \hbar^2 t^2/m^2 a^2)}\right\}. \quad (74)$$

Notice that

- $|\Psi(x, t)|^2$ is centred on $x = \hbar k_0 t/m = v_g t$. The centre of the wave packet travels at the group velocity.
 - $|\Psi(x, t)|^2$ has width $(\Delta x)^2 = a^2/2 (1 + \hbar^2 t^2/a^4 m^2)$, and so spreads as time evolves.
 - It can also be shown that $|\Psi(x, t)|^2$ remains normalised as it propagates.
-

According to Heisenberg's Uncertainty Principle, if the initial ($t = 0$) uncertainty in the particle's position, x_0 , is Δx_0 , then there is a spread in the particles initial momentum and hence velocity of $\Delta v_0 = \hbar/2m\Delta x_0$.

At later times $x(t) = x_0 + vt$, but this is subject to uncertainty.

At time t , classically one would write

$$[\Delta x(t)]^2 = [\Delta x_0]^2 + [\Delta v_0 t]^2 = (\Delta x_0)^2 + (\hbar t/2m\Delta x_0)^2, \quad (75)$$

where the errors have been combined in quadrature on the assumed basis that measurement errors are uncorrelated.

Given that $\Delta x_0^2 = a^2/2$, we find

$$[\Delta x(t)]^2 = a^2/2[1 + \hbar^2 t^2/a^4 m^2], \quad (76)$$

as before.

The situation is shown in Fig. 17.

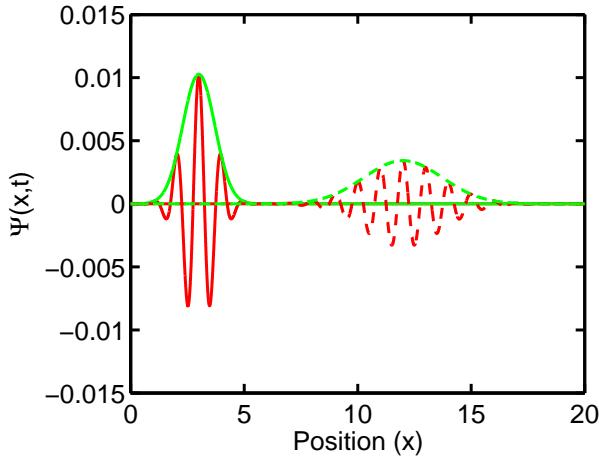


Figure 17: If the dispersion relationship is quadratic, the wavepacket spreads as it propagates. Its increasing width can be self-consistently interpreted as being due to the fact that both the original position and momentum (velocity) were uncertain. In this case, the wavefunction remains Gaussian as it propagates. Normalisation requires it to decrease in height.

12 Momentum representation

To this point we have used the wavevector k as the argument of the momentum wavefunction, which emphasises the relationship with the Fourier transform.

Physically, however, it is more elegant to write $g(k)$ directly in terms of momentum:

$$\begin{aligned}\Psi(x, t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(k) e^{i(kx - \omega t)} dk, \quad (77) \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(k, t) e^{ikx} dk \\ &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \Phi(p, t) e^{ipx/\hbar} dp,\end{aligned}$$

where in the last line we have used

$$\Phi(p, t) = \frac{1}{\sqrt{\hbar}} g\left(\frac{p}{\hbar}, t\right). \quad (78)$$

Also

$$g(p/\hbar, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \Psi(x, t) e^{-ipx/\hbar} dx, \quad (79)$$

$$\Phi(p, t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \Psi(x, t) e^{-ipx/\hbar} dx.$$

Thus we are left with the symmetrical expressions

$$\boxed{\Psi(x, t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \Phi(p, t) e^{ipx/\hbar} dp} \quad (80)$$

and

$$\boxed{\Phi(p, t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \Psi(x, t) e^{-ipx/\hbar} dx.} \quad (81)$$

Using this definition, the Gaussian wave packet becomes

$$\Phi(p, 0) = \left(\frac{a^2}{\pi\hbar^2} \right)^{1/4} \exp^{-a^2(p-p_0)^2/2\hbar^2}. \quad (82)$$

$\Psi(x, t)$ is the wavefunction in the *position representation*, and $\Phi(p, t)$ is the wavefunction in the *momentum representation*.

Each of $\Psi(x, t)$ and $\Phi(p, t)$ contains complete information about the state, because one can be derived from the other.

$|\Psi(x, t)|^2 dx$ is the probability that the particle is found in the region dx centred on x at time t .

$|\Phi(p, t)|^2 dp$ is the probability that the particle is measured to have momentum in the range dp centred on p at time t .

Show that if $\Psi(x, t)$ is normalised for all time then so is $\Phi(p, t)$. This is nothing but Parseval's theorem.

III Schrödinger's Equation

In the previous chapter, we discussed, at a phenomenological level, some of the key concepts that one must introduce in order to formulate a new description of reality that encompasses the wave-particle duality required by quantum mechanics. In particular, we focused on the notion of wave function and its meaning in regards to physical properties of a system and the outcome of measurements. In this chapter, we will introduce a wave equation that describes the time evolution of a wave function – in a sense the quantum mechanical equivalent of Newton's equation of motion, which describes the time evolution of a classical object.

13 One dimensional wave equation

Can we find a differential equation that the wavefunction $\Psi(x, t)$ must obey?

For a general wavefunction

$$\Psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(k) e^{i(kx - \omega t)} dk, \quad (83)$$

it is clear that

$$\frac{\partial \Psi(x, t)}{\partial t} = \frac{-i}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \omega g(k) e^{i(kx - \omega t)} dk, \quad (84)$$

and

$$\begin{aligned} \frac{\partial^2 \Psi(x, t)}{\partial x^2} &= \frac{-1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} k^2 g(k) e^{i(kx - \omega t)} dk \\ &= \frac{-2m}{\hbar} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \omega g(k) e^{i(kx - \omega t)} dk, \end{aligned} \quad (85)$$

where we have used the quadratic dispersion relation for a free particle

$$\omega = \frac{\hbar k^2}{2m}. \quad (86)$$

Comparing Eqs. (84) and (85) gives

$$\boxed{i\hbar \frac{\partial \Psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi(x, t)}{\partial x^2}}. \quad (87)$$

Eq. (87) is Schrödinger's wave equation for a free particle in one dimension.

The potential is everywhere zero, or constant, and the particle is free, which justifies the use of the quadratic dispersion relationship.

Every solution constitutes a possible wavefunction, and therefore a possible probability distribution.

Consider the specific form

$$\Psi(x, t) = A e^{i(kx - \omega t)}. \quad (88)$$

We observe that

$$i\hbar \frac{\partial \Psi(x, t)}{\partial t} = E\Psi(x, t) \quad (89)$$

$$-i\hbar \frac{\partial \Psi(x, t)}{\partial x} = p\Psi(x, t) \quad (90)$$

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \Psi(x, t)}{\partial x^2} = \frac{p^2}{2m} \Psi(x, t). \quad (91)$$

In quantum mechanics, *operators* are denoted by a ‘hat’:

$$\boxed{\hat{E} \equiv i\hbar \frac{\partial}{\partial t} \quad \text{Total Energy Operator}} \quad (92)$$

$$\boxed{\hat{p} \equiv -i\hbar \frac{\partial}{\partial x} \quad \text{Momentum Operator}} \quad (93)$$

$$\boxed{\frac{\hat{p}^2}{2m} \equiv -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \quad \text{Kinetic Energy Operator.}} \quad (94)$$

These operators ‘represent’ the total energy, momentum and kinetic energy respectively when they act on the wavefunction.

The wave equation is derived from the usual classical equation for the conservation of energy, but replacing the classical variables with the appropriate operators:

$$E = \frac{p^2}{2m} \quad \rightarrow \quad \hat{E} = \frac{\hat{p}^2}{2m}. \quad (95)$$

This expression gives Scrödinger’s equation.

When \hat{E} , \hat{p} and $\hat{p}^2/2m$ operate on the travelling wave, which represents a particle having well-defined values of energy and momentum, they return the values of each of the variables they represent, multiplied by the wavefunction.

In quantum mechanics, dynamical variables are replaced by operators, which act on wavefunctions in accordance with the usual classical equations of motion.

An operator is a mathematical indicator that a measurement, or observation, is taking place.

The eigenvalues of an operator are the possible values that might be returned by an experiment.

Operators are not commutative, and so if two measurements are carried out in different orders, different results may be obtained.

In the presence of an external, possibly time-dependent, potential, $V(x, t)$,

$$\begin{aligned} E &= \frac{p^2}{2m} + V(x, t) \\ \hat{E} &= \frac{\hat{p}^2}{2m} + \hat{V}(x, t). \end{aligned} \quad (96)$$

Remember that multiplication by a constant is also an operator.



Erwin Schrödinger
(1887-1961)

Schrödinger's one-dimensional wave equation then generalises to

$$i\hbar \frac{\partial \Psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi(x, t)}{\partial x^2} + V(x, t)\Psi(x, t) \quad (97)$$

A relativistic version of Schrödinger's equation can be derived in a similar way, taking into account the rest energy $E = mc^2$, and this leads to the notion of antiparticles. This topic is beyond the scope of this course.

14 Wave equation in three dimensions

The three-dimensional momentum and kinetic energy operators are

$$\hat{\mathbf{p}} \equiv -i\hbar\nabla \quad \text{3D momentum operator} \quad (98)$$

and

$$\frac{\hat{p}^2}{2m} \equiv -\frac{\hbar^2}{2m}\nabla^2 \quad \text{3D kinetic energy operator} \quad (99)$$

respectively. From these one can straightforwardly write the 3D Schrödinger equation,

$$i\hbar \frac{\partial \Psi(\mathbf{r}, t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi(\mathbf{r}, t) + V(\mathbf{r}, t)\Psi(\mathbf{r}, t). \quad (100)$$

Although we have ‘reverse-engineered’ this equation, it is true for all physical systems, and it is the starting point for all quantum wave mechanics.

It can be derived more formally through Lagrangian mechanics, but we shall not do so here.

In classical mechanics the sum of the kinetic and potential energies is the total energy, otherwise known as the *Hamiltonian*.

Create the Hamiltonian operator

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + \hat{V}(\mathbf{r}, t), \quad (101)$$

which itself is an observable.

Schrödinger’s equation then takes on the notationally compact form

$$i\hbar\frac{\partial\Psi(\mathbf{r}, t)}{\partial t} = \hat{H}\Psi(\mathbf{r}, t). \quad (102)$$

It then follows that

$\hat{H} \equiv i\hbar\frac{\partial}{\partial t}$

(103)

is the total energy operator, as we anticipated earlier.

15 Time-independent wave equation

Suppose that the potential is a function of position only, $V(\mathbf{r})$.

According to the method of the separation of variables assume that

$$\Psi(\mathbf{r}, t) = \psi(\mathbf{r})T(t). \quad (104)$$

Then

$$i\hbar \left[\frac{1}{T(t)} \right] \frac{\partial T(t)}{\partial t} = \frac{1}{\psi(\mathbf{r})} \left[-\frac{\hbar^2}{2m} \nabla^2 \psi(\mathbf{r}) + V(\mathbf{r})\psi(\mathbf{r}) \right]. \quad (105)$$

The LHS is independent of \mathbf{r} and the RHS is independent of t , and so both must equal the same constant, say E :

$$E = \frac{1}{\psi(\mathbf{r})} \left[-\frac{\hbar^2}{2m} \nabla^2 \psi(\mathbf{r}) + V(\mathbf{r})\psi(\mathbf{r}) \right]. \quad (106)$$

For a one-dimensional particle

$$\begin{aligned} (E - V) \psi(x) &= -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x)}{\partial x^2} \\ &= \frac{p^2}{2m} \psi(x). \end{aligned} \quad (107)$$

where we have used Eq. (91).

E is the total energy:

$$E = i\hbar \left[\frac{1}{T(t)} \right] \frac{\partial T(t)}{\partial t}, \quad (108)$$

giving

$$T(t) = A e^{-iEt/\hbar}. \quad (109)$$

We conclude that if the potential does not change with time

$$\boxed{\Psi(\mathbf{r}, t) = \Psi(\mathbf{r}) e^{-iEt/\hbar}}, \quad (110)$$

where A is absorbed into $\Psi(\mathbf{r})$.

$\Psi(\mathbf{r})$ is the solution to the Schrödinger equation

$$\boxed{-\frac{\hbar^2}{2m} \nabla^2 \psi(\mathbf{r}) + V(\mathbf{r})\psi(\mathbf{r}) = E\psi(\mathbf{r})}. \quad (111)$$

Notice that the probability distribution does not change with time

$$|\Psi(\mathbf{r}, t)|^2 = |\Psi(\mathbf{r})|^2, \quad (112)$$

and so the solutions are referred to as *stationary states*.

Finally, using the Hamiltonian operator Eq. (101), Schrödinger's time-independent wave equation becomes

$$\boxed{\hat{H}\Psi(\mathbf{r}) = E\Psi(\mathbf{r})}. \quad (113)$$

16 Probability current

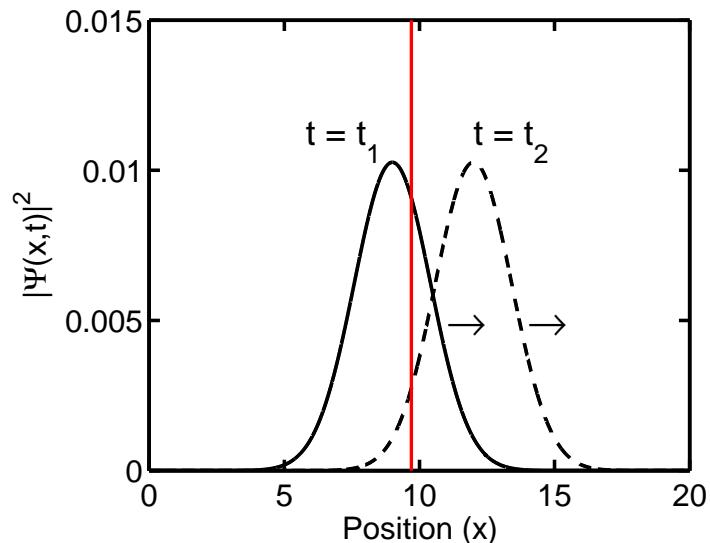


Figure 18: Probability distributions at two snapshots in time for a propagating particle.

When the square modulus of the wavefunction is integrated over all space, the result must be unity, because the particle will be found somewhere with certainty.

If the probability distribution changes with time, then it must be possible to define a *probability current*.

If the probability in one region decreases, the probability elsewhere must increase to ensure that probability is conserved overall. Correspondingly, the probability current must satisfy a continuity equation.

As the probability distribution passes through a surface of interest, there is a finite chance of finding the particle on one side or the other. In due course the particle transitions from being definitely on one side to being definitely on the other.

How can this concept be formulated mathematically?

Consider the expression

$$i\hbar \frac{\partial}{\partial t} \int_{\mathcal{V}} P(\mathbf{r}, t) dV = i\hbar \frac{\partial}{\partial t} \int_{\mathcal{V}} \Psi^*(\mathbf{r}, t) \Psi(\mathbf{r}, t) dV; \quad (114)$$

apart from the factor $i\hbar$, this is the time rate of change of the probability of finding the particle in a closed region \mathcal{V} .

Now

$$\begin{aligned} & i\hbar \frac{\partial}{\partial t} \int_{\mathcal{V}} P(\mathbf{r}, t) dV \\ &= i\hbar \int_{\mathcal{V}} \left[\Psi^*(\mathbf{r}, t) \frac{\partial \Psi(\mathbf{r}, t)}{\partial t} + \frac{\partial \Psi^*(\mathbf{r}, t)}{\partial t} \Psi(\mathbf{r}, t) \right] dV, \\ &= \int_{\mathcal{V}} \left[\Psi^*(\mathbf{r}, t) \hat{H} \Psi(\mathbf{r}, t) - (\hat{H} \Psi^*(\mathbf{r}, t)) \Psi(\mathbf{r}, t) \right] dV, \\ &= -\frac{\hbar^2}{2m} \int_{\mathcal{V}} \left[\Psi^*(\mathbf{r}, t) \nabla^2 \Psi(\mathbf{r}, t) - (\nabla^2 \Psi^*(\mathbf{r}, t)) \Psi(\mathbf{r}, t) \right] dV, \\ &= -\frac{\hbar^2}{2m} \int_{\mathcal{V}} \nabla \cdot \left[\Psi^*(\mathbf{r}, t) \nabla \Psi(\mathbf{r}, t) - \Psi(\mathbf{r}, t) \nabla \Psi^*(\mathbf{r}, t) \right] dV, \\ &= -\frac{\hbar^2}{2m} \oint_{\mathcal{S}} \left[\Psi^*(\mathbf{r}, t) \nabla \Psi(\mathbf{r}, t) - \Psi(\mathbf{r}, t) \nabla \Psi^*(\mathbf{r}, t) \right] \cdot d\mathbf{s} \end{aligned} \quad (115)$$

where the third line follows from Eq. (102), and the fourth line follows from applying the Hamiltonian, Eq. (101), and noting that the terms involving the potential cancel.

The fifth line follows by applying the vector differential identity

$$\nabla \cdot (f\mathbf{F}) = \mathbf{F} \cdot \nabla f + f \nabla \cdot \mathbf{F}, \quad (116)$$

which is valid for any scalar function f and vector field \mathbf{F} . The sixth line comes from applying the divergence theorem.

Thus

$$\frac{\partial}{\partial t} \int_{\mathcal{V}} P(\mathbf{r}, t) dV = - \oint_{\mathcal{S}} \mathbf{J}(\mathbf{r}, t) \cdot d\mathbf{s} \quad (117)$$

where

$$\boxed{\mathbf{J}(\mathbf{r}, t) = \frac{\hbar}{i2m} [\Psi^*(\mathbf{r}, t) \nabla \Psi(\mathbf{r}, t) - \Psi(\mathbf{r}, t) \nabla \Psi^*(\mathbf{r}, t)]} \quad (118)$$

which is a vector field.

Eq. (117) shows that the time rate of increase of the probability within some closed region is equal to the negative surface integral of the outward normal component of some flux over the surface bounding the region.

We can interpret $\mathbf{J}(\mathbf{r}, t)$ as a *probability current*.

Eq. (117) must be true for all volumes, and so

$$\boxed{\frac{\partial P}{\partial t} + \nabla \cdot \mathbf{J} = 0,} \quad (119)$$

which is the *continuity equation* for the probability.

Notice that Eq. (118) can be written as

$$\boxed{\mathbf{J}(\mathbf{r}, t) = \Re[\Psi^*(\mathbf{r}, t) \frac{\hbar}{im} \nabla \Psi(\mathbf{r}, t)],} \quad (120)$$

and so $\mathbf{J}(\mathbf{r}, t)$ is a real vector field, which can be plotted as a function of space and time.

It can be evaluated easily once the wavefunction is known.

We shall make use of the probability current later on in the course.

Also, using $\hat{\mathbf{p}} = -i\hbar\nabla$, Eq. (120) becomes

$$\boxed{\mathbf{J}(\mathbf{r}, t) = \Re[\Psi^*(\mathbf{r}, t) \frac{\hat{\mathbf{p}}}{m} \Psi(\mathbf{r}, t)],} \quad (121)$$

which is a commonly used representation.

Consider the form of the operator $(\hbar/im)\nabla$ when acting on a wavefunction in one dimension, and therefore gain an appreciation of the physical meaning of Eq. (121).

IV Unbound Particles

In the previous chapter we introduced a wave equation that describes the time evolution of a wave function, and we discussed its understanding in terms of energy and momentum operators. Moreover, we introduced the concept of probability current and the corresponding continuity equation. In this chapter we begin our investigation of how one solves the Schrödinger equation, focusing in the first instance on the case of 1D systems with (piece-wise) constant potentials. We shall discuss the physical significance of the results and remark on the difference between classical and quantum behaviour. This chapter addresses the case of unbound particles, whereas bound ones will be the subject of the next chapter.

17 Solving Schrödinger's equation

We can use Schrödinger's time-independent wave equation to investigate the behavior of particles under the influence of time-independent potentials, or forces.

What happens when a particle moves in a region where the potential changes with position?

Consider the one-dimensional case.

If the particle's total energy E is constant, its kinetic energy $T(x)$ must vary with x :

$$T(x) = E - V(x). \quad (122)$$

Classically, if $V(x) > E$, the particle cannot enter the region.

In the quantum mechanical case, we must solve Schrödinger's time-independent wave equation,

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x)}{\partial x^2} + V(x)\psi(x) = E\psi(x). \quad (123)$$

Consider the situation where the potential is constant, V_0 , but there is no limit on its value:

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x)}{\partial x^2} = [E - V_0] \psi(x). \quad (124)$$

The solution is

$$\Psi(x, t) = A e^{i(\pm kx - \omega t)}, \quad (125)$$

with

$$k = \left[\frac{2m(E - V_0)}{\hbar^2} \right]^{1/2}. \quad (126)$$

Both signs, \pm , constitute allowable solutions.

For $E > V_0$ the solution has a real wave number, the momentum is well defined, and the solution takes the form of a travelling wave having a position-independent amplitude.

If $E < V_0$, which corresponds to a classically forbidden region, k becomes complex, say $k = i\kappa$, and the solution is

$$\Psi(x, t) = A e^{\mp \kappa x} e^{-i\omega t}, \quad (127)$$

where

$$\kappa = \left[\frac{2m(V_0 - E)}{\hbar^2} \right]^{1/2}. \quad (128)$$

The solution is *not* that of a travelling wave. It has the form of a decaying, or growing, exponential having real and imaginary parts that oscillate in time: Fig. 19.

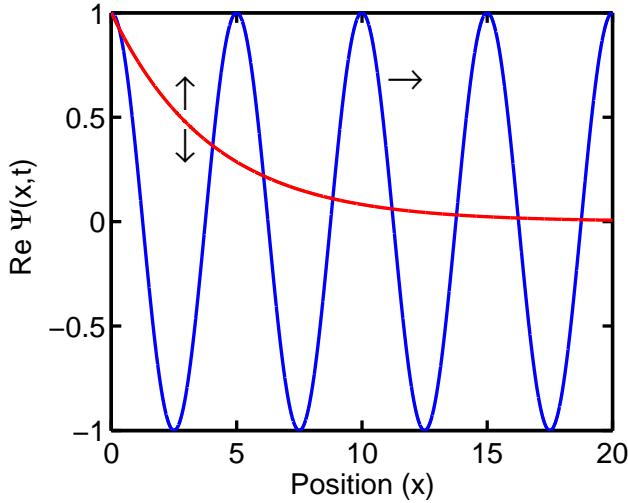


Figure 19: Blue: Real part of a travelling wavefunction at a snapshot in time. Red: Real part of an evanescent wavefunction at a snapshot in time.

There are no *wavelike* variations in x .

Solutions that have purely imaginary wavenumbers, and therefore decay exponentially without any wave-like behaviour, are called *evanescent waves*.

They are found in electromagnetics and classical elasticity, as well as quantum mechanics.

The probability of finding the particle within some region dx ,

$$|\Psi(x, t)|^2 dx = |A|^2 e^{\mp 2\kappa x} dx, \quad (129)$$

does not vanish, and one wonders whether this can ever occur in reality - time for an experiment!

Notice that the higher the potential, the shorter the decay length.

The wavefunction is forced to zero everywhere as the potential is increased to infinity.

Calculate the probability current at some reference position for the wavefunction given by Eq. (127). Does this contradict the notion that there is some finite chance of finding the particle on either side of the reference position?

18 Beams

Before solving problems involving spatially varying potentials, it is beneficial to revisit the wavefunction

$$\Psi(x, y, z, t) = A e^{i(kx - \omega t)}, \quad (130)$$

interpreted as a wavefunction in three-dimensions, but constant in the y and z directions.

$|A|^2 dx$ is the probability of finding the particle in a box of length dx and unit cross section in the $y - z$ plane.

The problem with this wavefunction is that it cannot be normalised.

It is not square integrable and so it is not meaningful, not even as a probability density:

$$\int_{-\infty}^{+\infty} |\Psi(x, t)|^2 dx = \infty. \quad (131)$$

In other words, because there is a finite chance of finding the particle anywhere, $|A|^2$ must tend to zero in order to give an integrated probability of unity.

Likewise, the square amplitude of the momentum wavefunction, a delta function squared, is not normalisable, because the integral of the square of a delta function is not defined.

It is possible, however, to calculate the probability current density

$$\mathbf{J}(\mathbf{r}, t) = \Re \left[\Psi^*(\mathbf{r}, t) \frac{\hbar}{im} \nabla \Psi(\mathbf{r}, t) \right], \quad (132)$$

which for an x -directed wave becomes

$$\mathbf{J} = \frac{\hbar k}{m} |A|^2 \hat{\mathbf{i}} = v |A|^2 \hat{\mathbf{i}} \quad (133)$$

as expected, because the box of unit cross section and length dx moves across the reference plane in time dt .

Equivalently,

$$J = \frac{dP_{\text{box}}}{dt} = \frac{|A|^2 dx}{dt} = v |A|^2 \quad (134)$$

is the probability current per unit cross-sectional area travelling in the x direction.

Because, for a plane wave, $|A|^2$ is vanishingly small, the probability current becomes vanishingly small.

It is often argued that if one has a beam of particles, all having the same momentum, then the plane-wave form Eq. (125) can be used to represent their behaviour as long as the probability current density, $J = |\mathbf{J}|$, is associated with the average particle flux, F , in particles per second per unit area.

Suppose that there are N particles present, all of which behave the same and independently of each other; then the average particle flux F , in particles per second, is

$$F = N |A|^2 \frac{dx}{dt}. \quad (135)$$

The vanishingly small probability amplitude associated with having an infinitely extended and uniform wavefunction – an infinitely long beam – is overcome by having an infinite number of particles!

Also, there is an implication that a single complex wavefunction is now being used to describe a collection of particles, rather than a single particle.

This sudden jump in reasoning hides many issues associated with multiple-particle systems, and it should be avoided if at all possible. It is superficially appealing, but problematic.

For example, it assumes that the particles are not *entangled*; in other words that the probability of finding any particle at some location is the sum of the probabilities of finding each of the particles there separately - we shall talk about this later in the course.

19 Boundary conditions

To apply Schrödinger's equation to problems where the potential varies in space (and in particular to the case we will study shortly of piece-wise constant, discontinuous potentials), it is necessary to know the boundary conditions that the wavefunction must obey.

In the one-dimensional case, with time-independent potentials, the wave function satisfies the equation

$$\frac{\partial^2 \psi(x)}{\partial x^2} = -\frac{2m}{\hbar^2} [E - V(x)] \psi(x). \quad (136)$$

We also know that $\psi(x)$ must be finite everywhere otherwise the probability density would be unphysical.

Thus, the RHS of Eq. (136) must be finite and $\partial^2 \psi(x)/\partial x^2$ must be finite. It follows that $\partial \psi(x)/\partial x$ must be finite.

As a result, $\partial \psi(x)/\partial x$ and $\psi(x)$ must be continuous.

In conclusion:

- $\psi(x)$ must be finite and continuous.
- $\partial \psi(x)/\partial x$ must be finite and continuous.

These conditions impose a certain level of ‘smoothness’ on the solutions, which means that the probability amplitude cannot change abruptly across a boundary.

The same reasoning applies in all three Cartesian directions.

In the case where V changes abruptly to infinity, $\psi(x)$ must tend to zero and therefore $\partial\psi(x)/\partial x$ varies discontinuously.

20 Potential step

Recall that, classically, a particle is unbound if there is some infinite region over which the particle’s energy is *everywhere* greater than the potential energy.

Consider the situation when an unbound particle exists in a region where there is a step discontinuity in the potential.

One example is a photo-electron at the surface of a metal. The electron sees a step in potential, V_0 – the workfunction encountered in the Chapter I – as it exits the metal: Fig. 20.

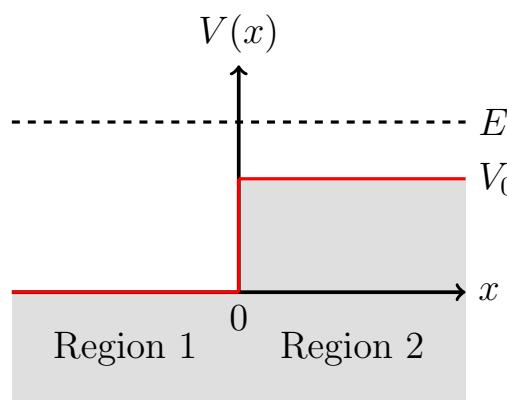


Figure 20: Potential step.

Can we find a solution to the time-independent Schrödinger equation?

Suppose that the particle has total energy E , and the potential has a step of V_0 at $x = 0$. Then the solutions in Region 1 (to the left of the step) and Region 2 (to the right of the step) are known to be

$$\begin{aligned}\Psi_1^+(x, t) &= Ae^{+ik_1x}e^{-i\omega t} \\ \Psi_1^-(x, t) &= rAe^{-ik_1x}e^{-i\omega t} \\ \Psi_2^+(x, t) &= tAe^{+ik_2x}e^{-i\omega t},\end{aligned}\tag{137}$$

where

$$k_1 = \left[\frac{2mE}{\hbar^2} \right]^{1/2} \quad k_2 = \left[\frac{2m(E - V_0)}{\hbar^2} \right]^{1/2}, \tag{138}$$

and r and t simply scale the coefficients of the ‘transmitted’ (Ψ_2^+) and ‘reflected’ (Ψ_1^-) wavefunctions to the complex amplitude of the ‘incident’ (Ψ_1^+) wavefunction.

Notice the following:

- The zero reference potential is defined with respect to the left region.
- The total energy is the same on both sides of the discontinuity, and so the wavefunctions have the same frequency $\omega = E/\hbar$.
- There is a tacit assumption that wavefunctions behave as waves, having reflected and transmitted parts.
- We allow incoming and outgoing waves in region 1, but only allow an outgoing wave in region 2, in correspondence with the notion of a probability flux incoming from the left.

The total wavefunctions in the two regions are then

$$\begin{aligned}\Psi_1(x, t) &= A[e^{+ik_1x} + re^{-ik_1x}]e^{-i\omega t} \\ \Psi_2(x, t) &= tAe^{+ik_2x}e^{-i\omega t}.\end{aligned}\tag{139}$$

We finally apply the boundary conditions at $x = 0$:

$$\Psi_1(0, t) = \Psi_2(0, t), \quad \left. \frac{\partial \Psi_1(x, t)}{\partial x} \right|_{x=0} = \left. \frac{\partial \Psi_2(x, t)}{\partial x} \right|_{x=0} \quad (140)$$

for all times.

From these we find

$$1 + r = t, \quad k_1(1 - r) = k_2 t, \quad (141)$$

and therefore the reflection and transmission amplitudes are

$$\boxed{r = \frac{k_1 - k_2}{k_1 + k_2}, \quad t = \frac{2k_1}{k_1 + k_2}.} \quad (142)$$

For a travelling wave, the magnitudes of the incident, reflected and transmitted probability fluxes are

$$\begin{aligned} J_1^+ &= |A|^2 \frac{\hbar k_1}{m} \\ J_1^- &= |Ar|^2 \frac{\hbar k_1}{m} = \left| A \frac{k_1 - k_2}{k_1 + k_2} \right|^2 \frac{\hbar k_1}{m} \\ J_2^+ &= |At|^2 \frac{\hbar k_2}{m} = \left| A \frac{2k_1}{k_1 + k_2} \right|^2 \frac{\hbar k_2}{m}. \end{aligned} \quad (143)$$

The probability flux reflection coefficient R is defined as

$$R = \frac{J_1^-}{J_1^+} = |r|^2 = \left| \frac{k_1 - k_2}{k_1 + k_2} \right|^2, \quad (144)$$

and the probability flux transmission coefficient T as

$$T = \frac{J_2^+}{J_1^+} = |t|^2 \frac{k_2}{k_1} = \frac{4k_1 k_2}{(k_1 + k_2)^2}. \quad (145)$$

Probability flux is conserved

$$J_1^- + J_2^+ = J_1^+, \quad (146)$$

as required; or, equivalently, $R + T = 1$.

20.1 Special case $V_0 < 0$

In this case $k_2 > k_1$, but the particle is still free everywhere.

The real and imaginary parts of the incident and reflected wavefunctions are shown in Fig. 21

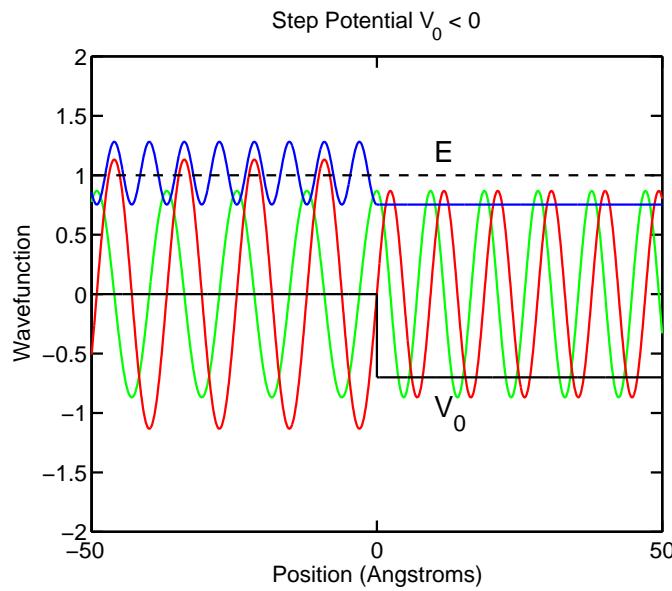


Figure 21: Black solid: Potential as a function of position. Dashed black: Energy as a function of position. Green: Real part of the wave function. Red: Imaginary part of the wave function. Blue: Squared-amplitude of the wave function.

Because the energy is constant, ω is constant.

Notice that the wavelength is long in the region where the potential energy is large, which is associated with the fact that the momentum, and the kinetic energy, is small.

The probability of detecting the particle has ripples to the left of the potential step as a consequence of the reflected probability wave interfering with the incident wave.

The real and imaginary parts of the wavefunction, and their spatial derivatives, are continuous at the boundary.

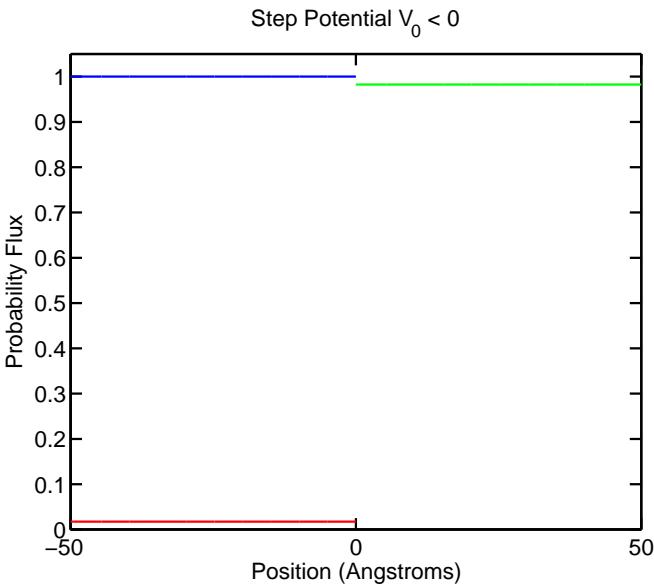


Figure 22: Blue: Probability flux of the forward travelling wave. Green: Probability flux of the transmitted wave. Red: Probability flux of the reflected wave.

Notice that:

$r < 0$ and real, and so there is reflection with a phase change of π ;

$t > 0$ and real, and so the transmitted wave is in phase with the incident wave.

The probability fluxes are shown in Fig. 22.

The reflections are relatively small, because there is nothing confining the particle.

There is a discontinuity in the fluxes at $x = 0$ because the wavevectors, and therefore speeds, are different on the two sides of the step.

Even though the probability is continuous, the fluxes are discontinuous due to the different ‘flow’ rates.

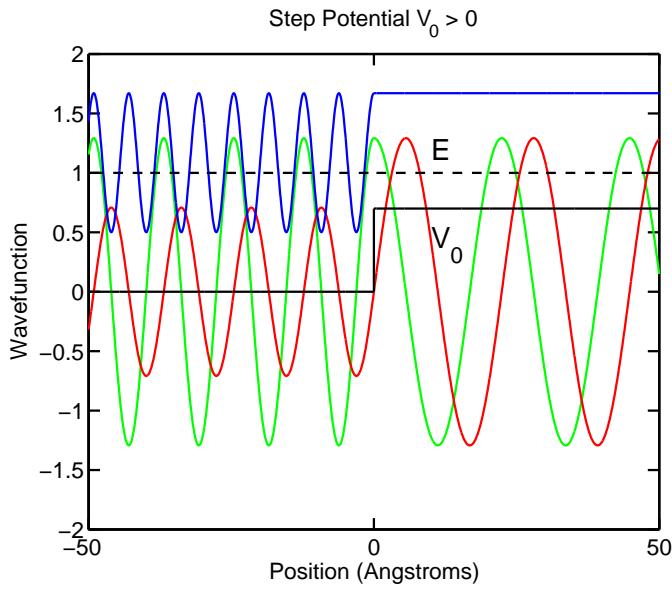


Figure 23: Black solid: Potential as a function of position. Dashed black: Energy as a function of position. Green: Real part of the wave function. Red: Imaginary part of the wave function. Blue: Squared amplitude of the wave function.

20.2 Special case $V_0 > 0$

In this case $k_2 < k_1$, but the particle is still free everywhere.

The real and imaginary parts of the incident and reflected wavefunctions are shown in Fig. 23.

The wavelength is large in the region where the kinetic energy, and therefore momentum, is small.

The associated probability fluxes are shown in Fig. 24.

Notice that:

$r > 0$ and real, and so there is reflection in phase with the incident wave;

$t > 0$ and real, and so the transmitted wave is in phase with the incident wave.

The reflections are still relatively small, because there is nothing constraining the particle.

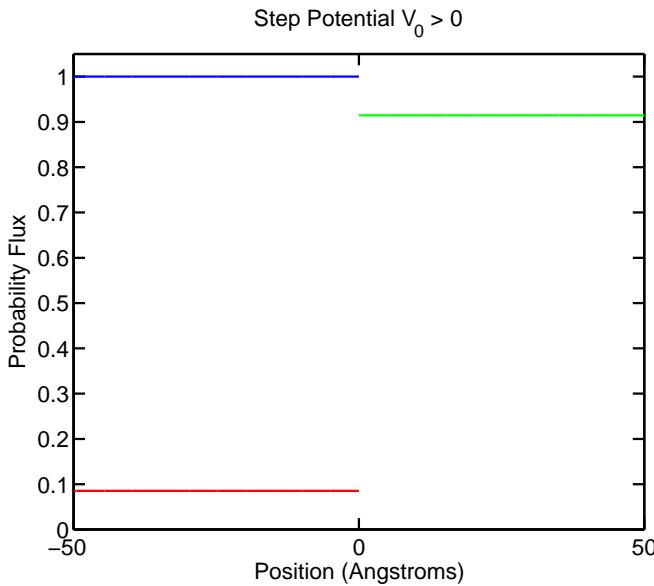


Figure 24: Blue: Probability flux of the forward travelling wave. Green: Probability flux of the transmitted wave. Red: Probability flux of the reflected wave.

There is a discontinuity in the fluxes at $x = 0$ because the wavevectors, and therefore speeds, are different on the two sides of the step.

Even though the probability is continuous, the fluxes are discontinuous due to the different ‘flow’ rates.

Classically, no reflection would occur.

20.3 Special case $0 < E < V_0$

In this case k_2 is imaginary.

The real and imaginary parts of the incident and reflected wavefunctions are shown in Fig. 25.

Then

$$\Psi_2(x) = te^{-\kappa_2 x} = \frac{2k_1}{k_1 + i\kappa_2} e^{-\kappa_2 x}, \quad (147)$$

and $\Psi_2(x)$ is non-zero, but decays in the x direction.

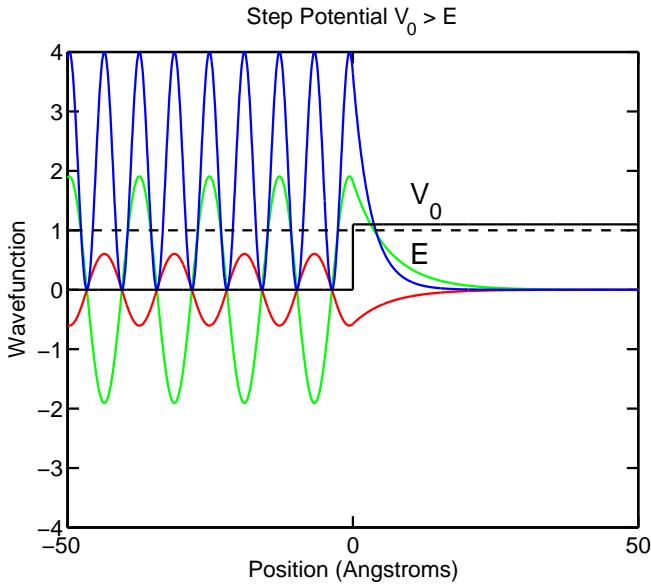


Figure 25: Black solid: Potential as a function of position. Dashed black: Energy as a function of position. Green: Real part of the wave function. Red: Imaginary part of the wave function. Blue: Squared amplitude of the wave function.

The exponentially increasing solution is eliminated on ‘physical grounds’.

There is a finite probability of finding the particle to the right of the barrier, which is not allowed classically.

The associated probability fluxes are shown in Fig. 26.

Also,

$$R = \left| \frac{k_1 - i\kappa_2}{k_1 + i\kappa_2} \right|^2 = 1, \quad (148)$$

and

$$T = 0, \quad (149)$$

because k_2 is imaginary.

The probability flux is completely reflected, even though there is the possibility of finding the particle in the classically forbidden region.

In the limit $V_0 \rightarrow \infty$, the penetration into the barrier becomes exceedingly small, and the classical case is recovered, as shown in Fig. 27, but now the probability of

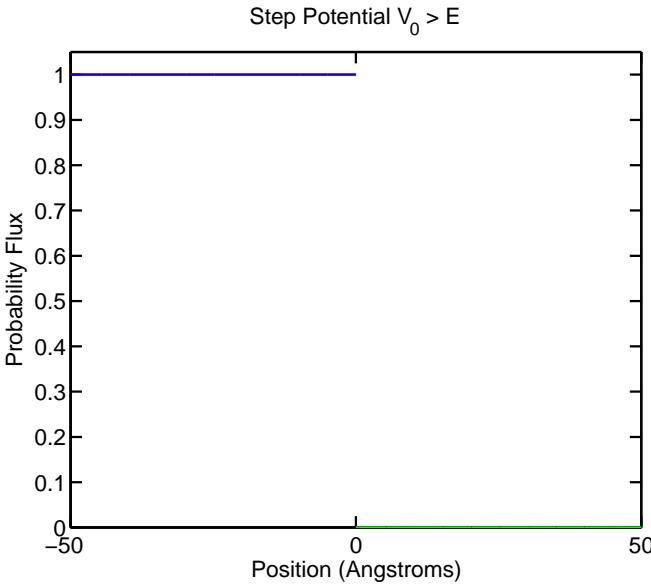


Figure 26: Blue: Probability flux of the forward travelling wave. Green: Probability flux of the transmitted wave. Red: Probability flux of the reflected wave.

finding the particle on the left is fully spatially modulated.

This occurs because $\kappa_2 \rightarrow \infty$,

$$r = \frac{k_1 - i\kappa_2}{k_1 + i\kappa_2} \rightarrow -1, \quad (150)$$

and

$$\Psi_1(x) = e^{+ik_1x} - e^{-ik_1x} = 2i \sin(k_1x), \quad (151)$$

as shown in Fig. 9.

In this case the inference is extreme.

Why can the particle be found to the right of the barrier?

Can the behaviour be understood in terms of Heisenberg's Uncertainty Principle?

If the energy of the particle were uncertain by an amount

$$\Delta E \geq V_0 - E \quad (152)$$

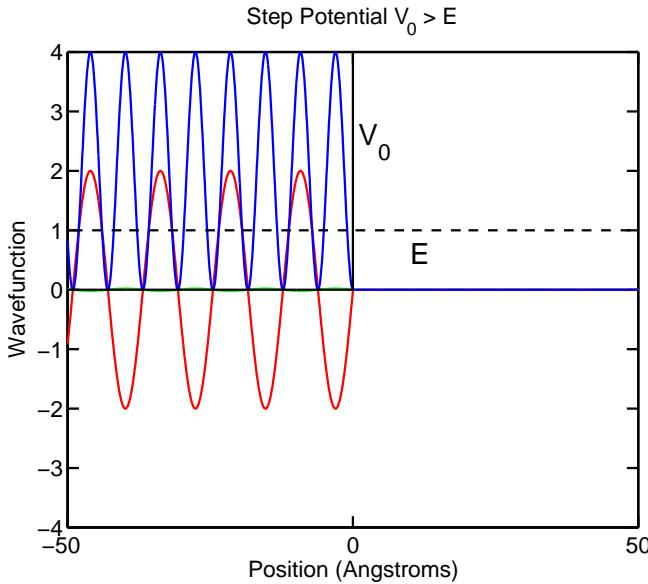


Figure 27: Black solid: Potential as a function of position. Dashed black: Energy as a function of position. Green: Real part of the wave function. Red: Imaginary part of the wave function. Blue: Squared amplitude of the wave function.

then the particle would have a probabilistic chance of surmounting the barrier.

In this case, the momentum would also be uncertain

$$(\Delta p)^2 = 2m\Delta E, \quad (153)$$

and the position would be uncertain

$$\Delta x \leq \frac{\hbar}{2\Delta p} = \frac{1}{2} \left[\frac{\hbar^2}{2m(V_0 - E)} \right]^{1/2}. \quad (154)$$

Eq. (154) can be written as

$$\Delta x \leq \frac{1}{2\kappa_2}, \quad (155)$$

which indeed is the *characteristic length* to the right of the barrier where the particle may be found (recall that the probability, Eq. (129), is the absolute value squared of the wave function, hence the factor of 2 on the RHS).

In other words, the probability distribution cannot go to zero abruptly at a finite-height barrier.

The uncertainty in position suggests that the particle could have an energy greater than $\Delta E \geq V_0 - E$, and therefore surmount the barrier.

Although this argument is very ‘hand-waving’, it illustrates the subtlety of the probabilistic interpretation of quantum mechanics.

21 Potential barriers

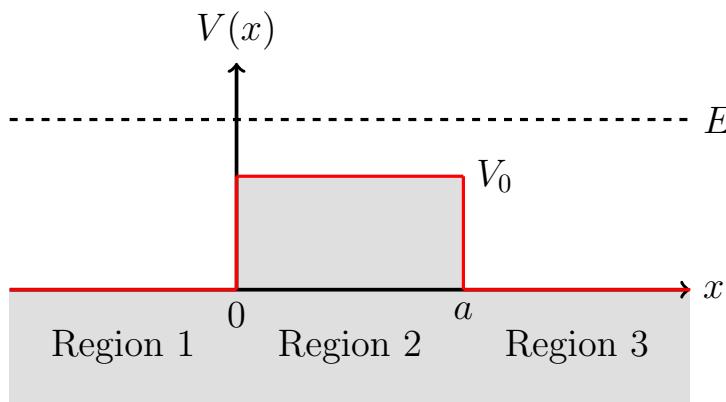


Figure 28: Potential barrier.

If wavefunctions can penetrate into potential steps, can particles tunnel through barriers?

Consider a barrier having height $V_0 > 0$ located between $x = 0$ and $x = a$: Fig. 28.

In Regions 1 and 3, the wavevector is k_1 , whereas in Region 2 it is k_2 . Make no assumptions about the height of the barrier and therefore the nature of k_2 .

Define a complex-valued reflection coefficient r for the backward-going wave in Region 1, and a complex-valued transmission coefficient t for the forward-going wave in Region 3.

The forward and backward travelling/decaying waves in Region 2 have complex coefficients α and β respectively.

As before,

$$k_1 = \left[\frac{2mE}{\hbar^2} \right]^{1/2} \quad k_2 = \left[\frac{2m(E - V_0)}{\hbar^2} \right]^{1/2}. \quad (156)$$

The wavefunctions in the three regions become

$$\begin{aligned} \Psi_1(x, t) &= A[e^{+ik_1x} + re^{-ik_1x}]e^{-i\omega t} \\ \Psi_2(x, t) &= A[\alpha e^{+ik_2x} + \beta e^{-ik_2x}]e^{-i\omega t} \\ \Psi_3(x, t) &= A[te^{+ik_1x}]e^{-i\omega t}. \end{aligned} \quad (157)$$

Applying the boundary conditions, $\psi(x)$ and $\partial\psi(x)/\partial x$ are continuous at $x = 0$ and $x = a$, gives

$$1 + r = \alpha + \beta \quad (158)$$

$$k_1(1 - r) = k_2(\alpha - \beta) \quad (159)$$

$$\alpha e^{ik_2a} + \beta e^{-ik_2a} = te^{ik_1a} \quad (160)$$

$$k_2(\alpha e^{ik_2a} - \beta e^{-ik_2a}) = k_1 te^{ik_1a}. \quad (161)$$

Combining Eqs. (158) and (159) gives

$$2 = \alpha \left(1 + \frac{k_2}{k_1} \right) + \beta \left(1 - \frac{k_2}{k_1} \right). \quad (162)$$

Likewise combining Eqs. (160) and (161) gives

$$\alpha = \frac{t}{2} \left(1 + \frac{k_1}{k_2} \right) e^{-ik_2a} e^{ik_1a}, \quad (163)$$

and

$$\beta = \frac{t}{2} \left(1 - \frac{k_1}{k_2} \right) e^{ik_2a} e^{ik_1a}. \quad (164)$$

Substituting (163) and (164) into (162) results in

$$2 = \frac{t}{2} \left[\left(1 + \frac{k_1}{k_2} \right) \left(1 + \frac{k_2}{k_1} \right) e^{-ik_2a} + \left(1 - \frac{k_1}{k_2} \right) \left(1 - \frac{k_2}{k_1} \right) e^{ik_2a} \right] e^{ik_1a}, \quad (165)$$

and finally

$$t = \left[\frac{2k_1 k_2}{2k_1 k_2 \cos(k_2 a) - i(k_1^2 + k_2^2) \sin(k_2 a)} \right] e^{-ik_1 a}. \quad (166)$$

It can also be shown that

$$r = \frac{(k_1^2 - k_2^2) \sin(ak_2)}{2ik_1 k_2 \cos(ak_2) + (k_1^2 + k_2^2) \sin(ak_2)}. \quad (167)$$

From which follows that $R + T = 1$, as expected.

21.1 Special case $V_0 < 0$

Consider a potential well of depth V_0 . In this case the wavevectors are all real, and $k_2 > k_1$: Fig. 29.

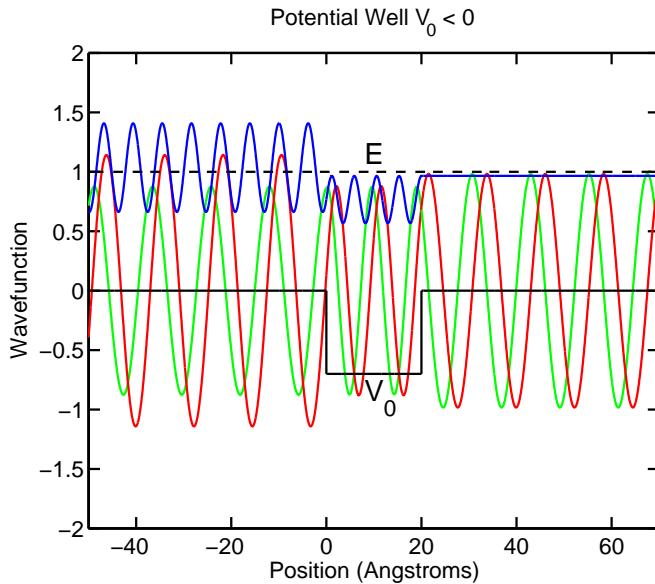


Figure 29: Square potential well of depth V_0 and width a .

Even though the particle's energy is everywhere greater than the potential energy, the probability density has spatial structure.

The probability of finding the particle in the region of the well is lower than the surroundings. A ‘classical interpretation’ is that because the particle is travelling faster in the well, it spends less time in the well.

21.2 Special case $E > V_0 > 0$

The case of a potential barrier is shown in Fig. 30 and the interpretation is straightforward.

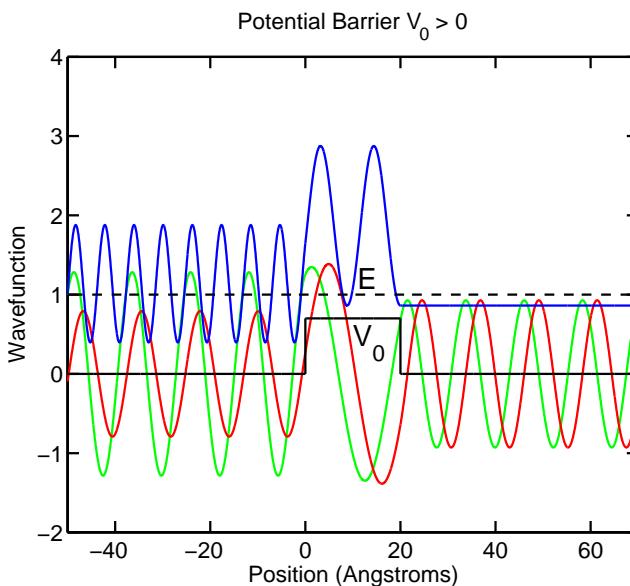


Figure 30: Square potential barrier of height V_0 and width a .

21.3 Special case $V_0 > E > 0$

The case of a potential barrier having a height greater than the particle's free energy is shown in Fig. 31.

The probability density decays exponentially within the barrier, giving rise to a small, but non-zero, probability flux on the far side of the barrier.

This effect is non-classical and is called *tunnelling*.

Fig. 32 shows the flux reflection and transmission coefficients, R and T , as a function of particle energy.

For small and large energies, the barrier tends towards classical behaviour, but for particle energies comparable with the barrier height, quantum resonances are seen.

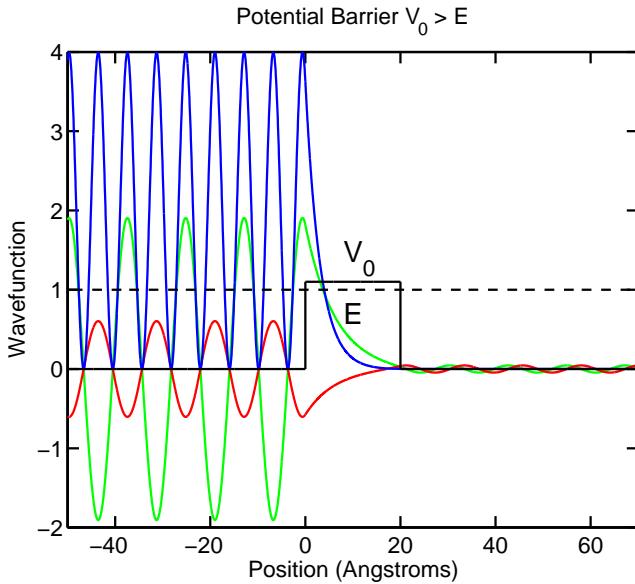


Figure 31: Square potential barrier of height V_0 and width a .

The resonant behaviour occurs because of the relationship between the de Broglie wavelength and barrier width.

According to Eq. (166), $T = 1$ when $k_2a = n\pi$, where n is an integer.

Explore the behaviour when $V_0 > E > 0$ as follows.

We know that $k_2 = i\kappa_2$, and therefore from (166)

$$t = \left[\frac{i2k_1\kappa_2}{i2k_1\kappa_2 \cos(i\kappa_2a) - i(k_1^2 - \kappa_2^2) \sin(i\kappa_2a)} \right] e^{-ik_1a}. \quad (168)$$

Using $\sin(iz) = i \sinh(z)$ and $\cos(iz) = \cosh(z)$

$$t = \left[\frac{2ik_1\kappa_2}{2ik_1\kappa_2 \cosh(\kappa_2a) + (k_1^2 - \kappa_2^2) \sinh(\kappa_2a)} \right] e^{-ik_1a}, \quad (169)$$

and

$$T = |t|^2 = \frac{4k_1^2\kappa_2^2}{4k_1^2\kappa_2^2 \cosh^2(\kappa_2a) + (k_1^2 - \kappa_2^2)^2 \sinh^2(\kappa_2a)}. \quad (170)$$

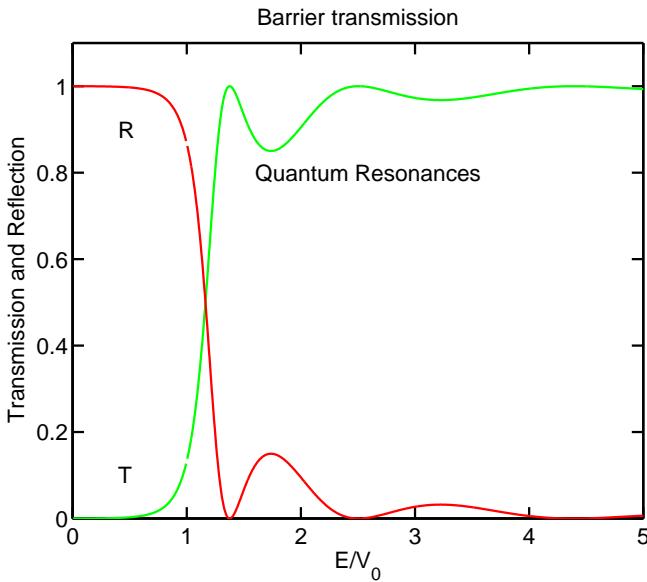


Figure 32: Transmission (green) and reflection (red) as a function of energy for a tunnel barrier.

When the barrier is thick compared with the decay length, $\kappa_2 a$ becomes large, and $\sinh(\kappa_2 a) \approx \cosh(\kappa_2 a) \approx \exp(\kappa_2 a)/2$ and

$$T = \frac{16k_1^2\kappa_2^2 e^{-2\kappa_2 a}}{(k_1^2 + \kappa_2^2)^2}, \quad (171)$$

where

$$\kappa_2 = \left[\frac{2m(V_0 - E)}{\hbar^2} \right]^{1/2}. \quad (172)$$

T falls exponentially as the thickness of the barrier a is increased.

Essentially the same effect is seen in Fig. 32: T falls to zero exponentially as V_0 is increased.

It may seem odd that energy can be transferred through such a region, but consider two horizontal elastic strings connected together by a rigid rod that can only move up and down. Energy can obviously travel along such a structure, even though there is no travelling wave-like behavior in the rod.

This analogy should not be pushed too far, but the principle is clear.

22 Examples of Tunnelling

There are numerous areas of physics where tunnelling is required in order to account for the observed behaviour.

22.1 Field Emission

Under normal circumstances, a low-energy electron cannot leave the surface of a metal because of the Coulombic interaction between the electron and the ions that make up the lattice.

The presence of an external electric field ($V = -Ex$) creates a potential having the form of a narrow barrier: Fig. 33. Electrons can tunnel across this barrier giving rise to a process known as *field emission*.

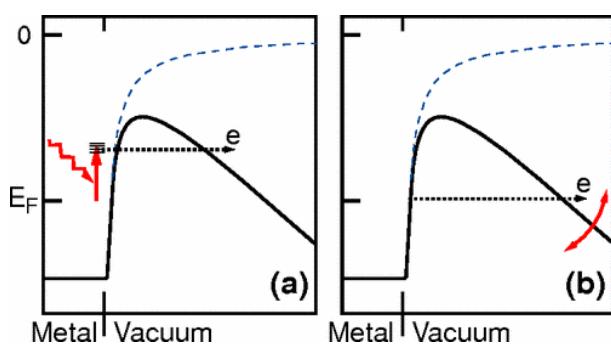


Figure 33: Potential at the surface of a metal when there is no externally applied electric field (dashed blue), and when there is a strong external field (solid black). The likelihood of tunnelling depends on the energy of the electron. The tunnelling rate can be controlled by changing the strength of the external field.

This process has been developed into a form of microscopy: Fig. 34.

A sharp tip is scanned over the surface of the sample, and a voltage is applied such that field emission occurs.

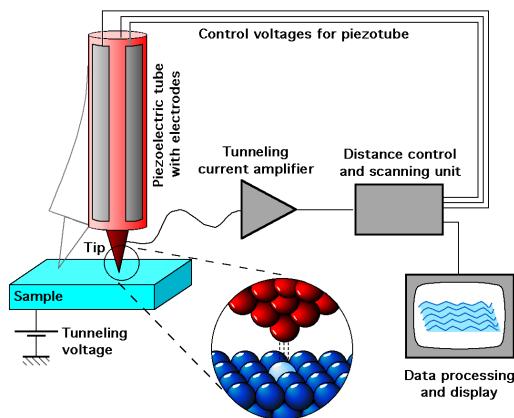


Figure 34: Scanning Tunnelling Microscope.

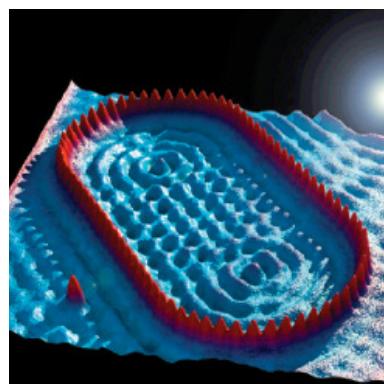


Figure 35: Quantum coral made of iron atoms on copper. M.F. Crommie, C.P. Lutz, D.M. Eigler, E.J. Heller, "Waves on a metal surface and quantum corrals", Surface Review and Letters 2 (1), 127-137 (1995).

The tunneling current, which is exponentially dependent on spacing, is used to hold the tip at a fixed height above the surface.

The feedback control to the piezo displacer is then used as a measure of the height of the surface.

Remarkably, height can be measured to better than 0.1 nm, giving extraordinary images

See Eric J. Heller (2007), "Electrons in Two Dimensions: Quantum Corrals and Semiconductor Microstructures," <https://nanohub.org/resources/3253>: Fig. 35.

22.2 Radioactive decay

Quantum mechanical tunnelling is central to explaining radioactive decay.

Large unstable nuclei can decay via the emission of an α -particle, which is the nucleus of a He atom comprising two protons and two neutrons.

The energy of the emitted α -particle is small (4-9 MeV) compared with the nuclear barrier (20-25 MeV), and so how is the particle able to escape?

Russian-born physicist George Gamow explained the process in 1928 in terms of quantum mechanical tunneling.



George Gamow
(1904-1968)

A striking observation is that the lifetime of a nucleus is strongly correlated with the energy released: Fig. 36.

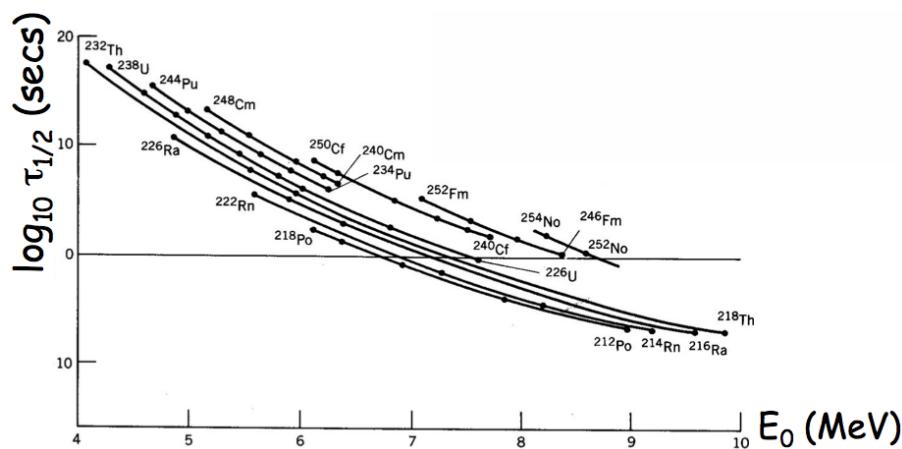


Figure 36: Half life of various radioactive nuclei as a function of the energy released.

For example,

$$\begin{aligned} ^{238}\text{U} &\rightarrow ^{234}\text{Th} + \alpha \quad E_o = 4.2\text{MeV} \quad \tau_{1/2} = 5 \times 10^9 \text{years} \\ ^{212}\text{Po} &\rightarrow ^{208}\text{Pb} + \alpha \quad E_o = 8.8\text{MeV} \quad \tau_{1/2} = 3 \times 10^{-7} \text{secs} \end{aligned}$$

A factor of 2.1 in kinetic energy corresponds to a factor of 10^{24} in half life.

The transmission probability, and therefore the rate of escape, depends near-exponentially on the kinetic energy of the released particle.

Small changes in the energy of the released particle correspond to very large changes in the half life of the material.

A high-energy particle comes from a nucleus with small half life.

V Bound Particles

In the previous chapter we began to study solutions to the Schrödinger equation, focusing on unbound particles in 1D with piece-wise constant potentials, and we encountered the exquisitely quantum mechanical concept of tunnelling. In this chapter, we shall continue the study by turning our attention to the behaviour of bound particles. We will then discuss the 1D simple harmonic oscillator, and the notions of zero point energy and of parity of the wavefunction. A correspondence principle between classical and quantum mechanics will also be formulated.

23 Bound vs unbound

Classically, a particle is *bound* if its energy is only greater than the potential energy over some finite region of space.

If a particle is *unbound*, the solutions of Schrödinger's equation form a continuum of states: they exist for all values E within some range.

If a particle is *bound*, the solutions of Schrödinger's equation form a set of discrete states: they only exist for certain specific values of E within some range.

In general, the potential-energy landscape can be sufficiently complicated that the solutions of Schrödinger's equation comprise a continuum of states, several bands of states, and discrete sets of states.

Consider a potential well having width a and depth $-V_0$.

If the particle has enough energy to escape the box, it has a continuum of states available; whereas if it is confined to the box, it can only occupy one of a number of discrete states.

Certain physical properties of bound particles, for example energy and momentum, can only take discrete values, and packets of energy must be released or absorbed when a particle moves between these states.

The *spectrum* of available states is a key concept when describing the behaviour of physical systems.

It figures highly in atomic physics, molecular physics and solid-state physics.

24 Infinite square well

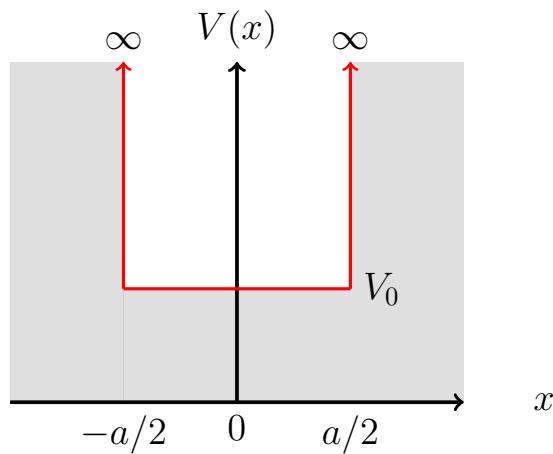


Figure 37: Particle in an infinite potential well.

Consider the potential

$$V(x) = \begin{cases} V_0 & \text{for } |x| \leq a/2 \\ \infty & \text{for } |x| > a/2. \end{cases} \quad (173)$$

The floor is at V_0 and the well has infinitely high sides: Fig. 37.

Under these circumstances, the particle cannot escape: there are no states available on the outside of the well.

For $|x| > a/2$, $\psi(x) = 0$, since $V(x) = \infty$. Because $\psi(x)$ must be continuous,

$$\psi(\pm a/2) = 0. \quad (174)$$

For $|x| \leq a/2$, $\psi(x)$ satisfies Schrödinger's equation with $V(x) = V_0$,

$$\frac{\partial^2 \psi(x)}{\partial x^2} + \frac{2m(E - V_0)}{\hbar^2} \psi(x) = 0. \quad (175)$$

Consider solutions for different values of E .

24.1 Special case $E < V_0$

Suppose that $E < V_0$, then k is imaginary, $i\kappa$, and

$$\psi(x) = Ae^{\kappa x} + Be^{-\kappa x}, \quad (176)$$

using the boundary condition $\psi(\pm a/2) = 0$ gives

$$\psi(x) = Ae^{\kappa a/2} + Be^{-\kappa a/2} = 0, \quad (177)$$

from which it follows that $A = B = 0$.

A particle having energy $E < V_0$ can never be found inside the well.

24.2 Special case $E = V_0$

In this case

$$\frac{\partial^2 \psi(x)}{\partial x^2} = 0, \quad (178)$$

and

$$\psi(x) = Ax + B. \quad (179)$$

The boundary condition $\psi(\pm a/2) = 0$ requires $A = B = 0$.

This situation is more difficult to understand because it says that a stationary particle, one for which $E - V_0 = 0$, cannot exist within the well.

Classically, this would be a perfectly reasonable solution!

24.3 Special case $E > V_0$

In this case, k is real and

$$\begin{aligned}\psi(x) &= Ae^{ikx} + Be^{-ikx} \\ &\text{or} \\ &= A \sin(kx) + B \cos(kx).\end{aligned}\tag{180}$$

The boundary conditions $\psi(\pm a/2) = 0$ give

$$\begin{aligned}A \sin(ka/2) + B \cos(ka/2) &= 0 \\ -A \sin(ka/2) + B \cos(ka/2) &= 0.\end{aligned}\tag{181}$$

Non-zero solutions only exist for

$$\begin{aligned}A &\neq 0 & B &= 0 \\ ka &= n\pi, \quad n = 2, 4, 6, \dots & \text{and} & \psi(x) = A \sin(kx) \\ A &= 0 & B &\neq 0 \\ ka &= n\pi, \quad n = 1, 3, 5, \dots & \text{and} & \psi(x) = B \cos(kx).\end{aligned}\tag{182}$$

Note that $ka = n\pi$ is equivalent to $a = n\lambda/2$.

There can be no solution for which $n = 0$, such that the amplitude of the wavefunction is constant across the well.

Bound states can be normalised according to the requirement

$$\int_{-\infty}^{+\infty} |\psi(x)|^2 dx = 1, \quad (183)$$

which becomes

$$\int_{-a/2}^{+a/2} |A \sin(n\pi x/a)|^2 dx = 1 \text{ for } n = 2, 4, 6, \dots, \quad (184)$$

and

$$\int_{-a/2}^{+a/2} |B \cos(n\pi x/a)|^2 dx = 1 \text{ for } n = 1, 3, 5, \dots. \quad (185)$$

Clearly $A = B = \sqrt{2/a}$.

Thus for an infinite square well

$$k_n = \frac{n\pi}{a} \quad (186)$$

$$E_n - V_0 = \boxed{\frac{\hbar^2 n^2 \pi^2}{2ma^2}}. \quad (187)$$

$E_n - V_0$ is the kinetic energy of the particle in the n 'th state, where $n \geq 1$. It is the energy with respect to the bottom of the well.

n is called a *quantum number*. Later we shall see that sets of different quantum numbers are needed to label more complicated states.

Eq. (187) gives the discrete energy levels of an infinitely high square potential well.

They are associated with the wavefunctions

$$\boxed{\psi_n(x) = \begin{cases} \sqrt{\frac{2}{a}} \cos(k_n x) & \text{for } n \text{ odd} \\ \sqrt{\frac{2}{a}} \sin(k_n x) & \text{for } n \text{ even.} \end{cases}} \quad (188)$$

It is possible to make devices – e.g., electrons in quantum dots – sufficiently small that the energy difference between neighbouring states is large enough for the discrete states to be observed.

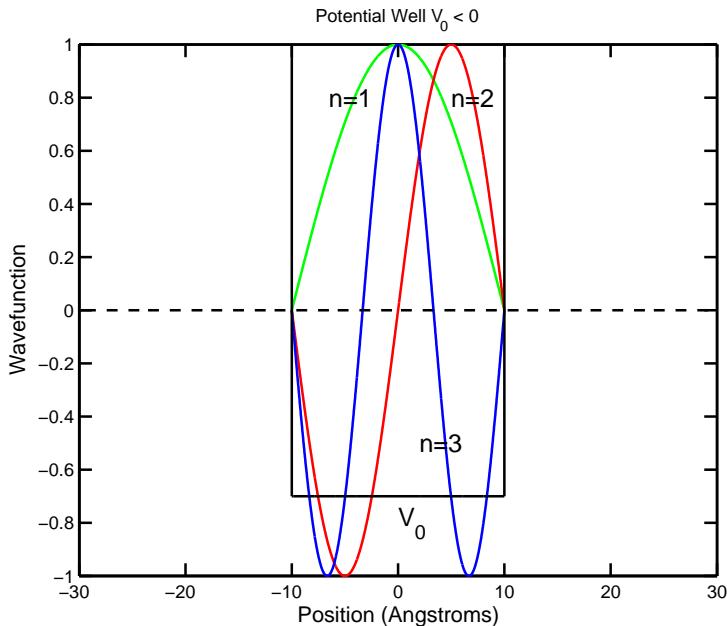


Figure 38: Wave functions of the infinite square potential well.

The situation is shown in Fig. 38:

- The allowed energy levels are discrete.
- Quantisation is a direct consequence of confinement.
- There is an infinite number of bound states for the infinite well.
- The lowest allowed energy level ($n = 1$) is non-zero, and is called the *zero point energy*. The particle cannot be stationary.
- The wave functions are the eigenfunctions of the Hamiltonian operator.
- The energies are the associated eigenvalues.
- The wave functions associated with the different states are orthogonal according to the inner product

$$\int_{-\infty}^{+\infty} \psi_m^*(x) \psi_n(x) dx = \delta_{mn} = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{otherwise.} \end{cases} \quad (189)$$

25 Finite square well

Consider the finite square potential well shown in Fig. 39:

$$V(x) = \begin{cases} -V_0 & \text{for } |x| \leq a/2 \\ 0 & \text{for } |x| > a/2 \end{cases}. \quad (190)$$

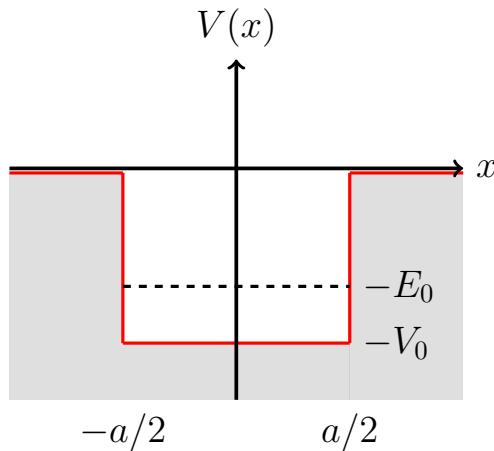


Figure 39: Particle in a square potential well.

For a particle having $E > 0$ there is a continuous range of unbound solutions.

For $E < -V_0$ no solutions are possible.

For $-V_0 < E < 0$, a number of bound states exist having discrete energy levels.

A mixed spectrum results.

Real wavevectors are needed within the well and imaginary wavenumbers outside of the well.

Use the solutions to Schrödinger's equation found earlier, and set

$$\begin{aligned} V &= -V_0 \\ E &= -E_0. \end{aligned} \quad (191)$$

For $|x| \leq a/2$ (region 2):

$$\psi(x) = A \sin(kx) + B \cos(kx) \text{ where } k = \sqrt{\frac{2m(V_0 - E_0)}{\hbar^2}}. \quad (192)$$

For $|x| > a/2$ (regions 1 and 3):

$$\psi(x) = Ce^{\kappa x} + De^{-\kappa x} \text{ where } \kappa = \sqrt{\frac{2mE_0}{\hbar^2}}. \quad (193)$$

Now

$$\psi(x) \rightarrow 0 \text{ as } x \rightarrow \pm\infty, \quad (194)$$

and therefore

$$\begin{aligned} \psi_1(x) &= Ce^{+\kappa x} \\ \psi_3(x) &= De^{-\kappa x}. \end{aligned} \quad (195)$$

Satisfying the boundary conditions at $x = a/2$ and $x = -a/2$ gives

$$A \sin(ka/2) + B \cos(ka/2) = De^{-\kappa a/2} \quad (196)$$

$$kA \cos(ka/2) - kB \sin(ka/2) = -\kappa D e^{-\kappa a/2} \quad (197)$$

$$-A \sin(ka/2) + B \cos(ka/2) = Ce^{-\kappa a/2} \quad (198)$$

$$kA \cos(ka/2) + kB \sin(ka/2) = \kappa C e^{-\kappa a/2}. \quad (199)$$

Substracting and adding Eqs. (196) and (198):

$$2A \sin(ka/2) = (D - C)e^{-\kappa a/2} \quad (200)$$

$$2B \cos(ka/2) = (D + C)e^{-\kappa a/2}. \quad (201)$$

Substracting and adding Eqs. (197) and (199):

$$2kB \sin(ka/2) = \kappa(D + C)e^{-\kappa a/2} \quad (202)$$

$$2kA \cos(ka/2) = -\kappa(D - C)e^{-\kappa a/2}. \quad (203)$$

Dividing (203) by (200):

$$k \cot(ka/2) = -\kappa \quad (204)$$

as long as $(C - D) \neq 0$ and $A \neq 0$.

Dividing (202) by (201):

$$k \tan(ka/2) = \kappa \quad (205)$$

as long as $(C + D) \neq 0$ and $B \neq 0$.

Eqs. (204) and (205) give rise to two classes of solution:

$$\begin{aligned} \text{Type 1 } & A = 0 \quad C = D \quad k \tan(ka/2) = \kappa \quad \text{even} \\ \text{Type 2 } & B = 0 \quad C = -D \quad k \cot(ka/2) = -\kappa \quad \text{odd.} \end{aligned} \quad (206)$$

These must be solved numerically. Using $X = ka/2$ and $Y = \kappa a/2$ gives

$$\begin{aligned} \text{Type 1: } & X \tan(X) = Y \\ \text{Type 2: } & -X \cot(X) = Y. \end{aligned} \quad (207)$$

Also

$$X^2 + Y^2 = \frac{a^2}{4\hbar^2} [2m(V_0 - E_0) + 2mE_0] = \frac{mV_0a^2}{2\hbar^2}. \quad (208)$$

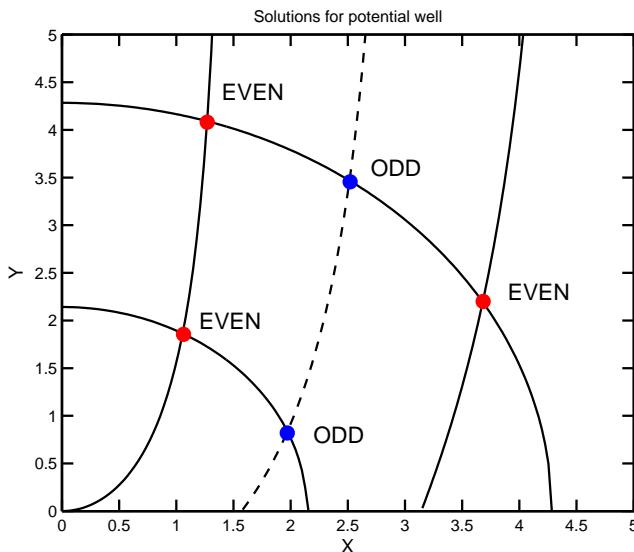


Figure 40: Graphical solutions for a square potential well. The top three intersections correspond to a wide well, and the bottom two to a well having half of the width.

The intersections of Eq. (207) and Eq. (208) in the $X - Y$ plane give the solutions: Fig. 40.

The different branches have even and odd *parity*, and different energies.

The energies can be calculated once X and Y are known:

$$E_n = -\frac{\hbar^2 \kappa^2}{2m} = -\frac{2\hbar^2 Y_n^2}{ma^2} = \frac{Y_n^2 V_0}{X_n^2 + Y_n^2}. \quad (209)$$

Collecting the results, the even parity solutions are

$$\psi_n(x) = \begin{cases} B \cos(k_n x) & |x| \leq a/2 \\ B \cos(k_n a/2) e^{\kappa a/2} e^{-\kappa x} & x > a/2 \\ B \cos(k_n a/2) e^{\kappa a/2} e^{\kappa x} & x < -a/2 \end{cases}, \quad (210)$$

where we have used Eqs. (196) and (198) to express the coefficients C and D in terms of B .

Likewise, the odd parity solutions are

$$\psi_n(x) = \begin{cases} A \sin(k_n x) & |x| \leq a/2 \\ A \sin(k_n a/2) e^{\kappa a/2} e^{-\kappa x} & x > a/2 \\ -A \sin(k_n a/2) e^{\kappa a/2} e^{\kappa x} & x < -a/2 \end{cases}. \quad (211)$$

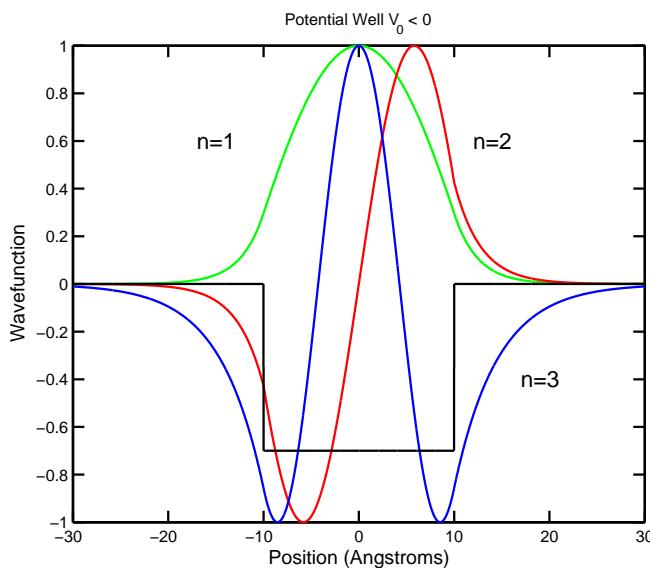


Figure 41: Wavefunctions of a finite symmetric square potential well.

The three wavefunctions corresponding to the top three solutions in Fig. 40, are illustrated in Fig. 41.

In this case, the energies are -0.63 eV, -0.49 eV and -0.18 eV with respect to the zero potential, corresponding to 0.07 eV, 0.21 eV and 0.52 eV above the bottom of the well.

The associated probability densities are shown in Fig. 42.

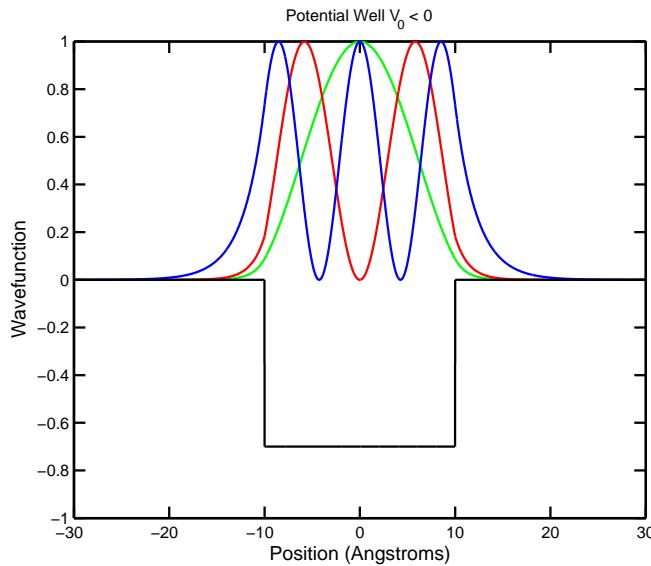


Figure 42: Probability densities of a symmetric square potential well.

Observe the following:

- The ground state is even, and high-order states alternate between even and odd.
- The associated energies are not distributed evenly.
- The number of bound states reduces as the well is made narrower.
- According to Fig. 40, and however small V_0 , there is always at least one bound even state. There is always some finite chance of a particle being captured.
- The overall spectrum of available states has a continuous part (free particle) and a discrete part (bound particle).

- There is always some probability of finding the particle inside the classically forbidden region $|x| > a/2$.
- The probability of finding the particle internally is not uniformly distributed. For odd states it will never be found in the middle!
- As the potential well is made deeper, the solutions tend to the sine and cosine solutions discussed earlier.

Many of these general observations hold in the case of more complicated potentials.

A particularly important potential landscape comprises a periodic array of potential wells: for example, the potential experienced by an electron in metal.

In that case, the solutions to Schrödinger's equation are called *Bloch wave functions*, and they will be discussed in the solid-state physics course.

They give rise to the band structure of crystals, which is central to describing the behaviour of semiconductors.

26 One-dimensional harmonic oscillator

Real potential wells are rarely square and can take on a variety of forms.

Because many systems are close to their ground states, the potential can be modelled locally in terms of a Taylor expansion:

$$V(x_0 + x) \approx V(x_0) + \frac{1}{2}\alpha x^2 + O(x^3) \dots \quad (212)$$

At low enough energies, a quadratic approximation may be sufficiently accurate to describe the behaviour of the system.

This is a situation of considerable importance because it corresponds to a *harmonic oscillator*.

Consider a particle of mass m moving in the quadratic potential,

$$V(x) \approx \frac{1}{2}\alpha x^2, \quad (213)$$

where α is the ‘spring constant’.

Classically, the particle oscillates with a natural frequency of $\omega = \sqrt{\alpha/m}$, from which we can write

$$V(x) = \frac{1}{2}m\omega^2x^2, \quad (214)$$

which expresses the potential in terms of the classical frequency of oscillation.

It is worth finding the quantum mechanical solutions of this problem because quadratic potentials occur extensively in physics.

To find the allowable solutions, we must solve the one-dimensional time-independent Schrödinger wave equation:

$$\frac{\partial^2\psi(x)}{\partial x^2} + \left[\frac{2mE}{\hbar^2} - \frac{m^2\omega^2x^2}{\hbar^2} \right] \psi(x) = 0. \quad (215)$$

I. Write the wave equation in dimensionless form:

Defining

$$q = x\sqrt{\frac{m\omega}{\hbar}} \text{ and } \epsilon = \frac{2E}{\hbar\omega}, \quad (216)$$

gives

$$\frac{\partial^2\chi(q)}{\partial q^2} + [\epsilon - q^2]\chi(q) = 0, \quad (217)$$

where we have used the substitution

$$\chi(q) = \psi \left[q\sqrt{\frac{\hbar}{m\omega}} \right]. \quad (218)$$

II. Look at the limiting form as $|q| \rightarrow \infty$:

In this case, $|q| \gg \epsilon$, and one can draw inspiration from considering the equation

$$\frac{\partial^2 \chi(q)}{\partial q^2} - q^2 \chi(q) = 0, \quad (219)$$

whereby

$$\chi(q) \propto e^{-q^2/2}, \quad (220)$$

at large $|q|$.

The wavefunction decays rapidly away from the centre of the well.

III. Assume a solution:

Assume a solution of the form

$$\chi(q) = H(q)e^{-q^2/2}, \quad (221)$$

where $H(q)$ is a well-behaved function yet to be found.

Substituting Eq. (221) into Eq. (217) gives

$$\boxed{\frac{\partial^2 H(q)}{\partial q^2} - 2q \frac{\partial H(q)}{\partial q} + (\epsilon - 1)H(q) = 0,} \quad (222)$$

which is *Hermite's equation*.

Hermite's equation is one of the classic equations that can be solved using an assumed power series solution.

IV. Final form:

Eq. (222) gives rise to a set of functions called *Hermite polynomials*, which are given by *Rodrigue's formula*,

$$\boxed{H_n(q) = (-1)^n e^{q^2} \frac{d^n}{dq^n} \left(e^{-q^2} \right),} \quad (223)$$

where $n = (\epsilon - 1)/2$ is an integer, and therefore

$$\boxed{E_n = \left(n + \frac{1}{2} \right) \hbar\omega.} \quad (224)$$

Eq. (223) can be evaluated to reveal the polynomials

$$\begin{aligned} H_0(q) &= 1 && \text{even} \\ H_1(q) &= 2q && \text{odd} \\ H_2(q) &= 4q^2 - 2 && \text{even} \\ H_3(q) &= 8q^3 - 12q && \text{odd.} \end{aligned} \quad (225)$$

A convenient set of recursion relations allow these polynomials to be calculated easily.

Collecting the results, we arrive at the expression

$$\psi_n(x) = A_n H_n \left[x \sqrt{\frac{m\omega}{\hbar}} \right] e^{-\frac{m\omega}{2\hbar}x^2}, \quad (226)$$

where A_n is a normalising factor.

A key feature of these functions is that they are orthogonal, and can therefore be normalised to form an orthonormal set.

We can compute the inner product between two wavefunctions

$$\int_{-\infty}^{\infty} \chi_n(q) \chi_m^*(q) dq = \int_{-\infty}^{\infty} A_n H_n(q) A_m^* H_m(q) e^{-q^2} dq, \quad (227)$$

and show that it is equal to zero unless $m = n$:

$$\begin{aligned} \int_{-\infty}^{\infty} \chi_n(q) \chi_m^*(q) dq &= \delta_{mn} |A_n|^2 \int_{-\infty}^{\infty} H_n^2(q) e^{-q^2} dq \\ &= \delta_{mn} |A_n|^2 \sqrt{\pi} 2^n n!. \end{aligned} \quad (228)$$

Thus in order to normalise the wavefunctions, we need to set

$$A_n = [\sqrt{\pi} 2^n n!]^{-1/2}. \quad (229)$$

The normalised solutions become

$$\psi_n(x) = \frac{1}{[\sqrt{\pi} 2^n n!]^{1/2}} H_n \left[x \sqrt{\frac{m\omega}{\hbar}} \right] e^{-\frac{m\omega}{2\hbar}x^2}. \quad (230)$$

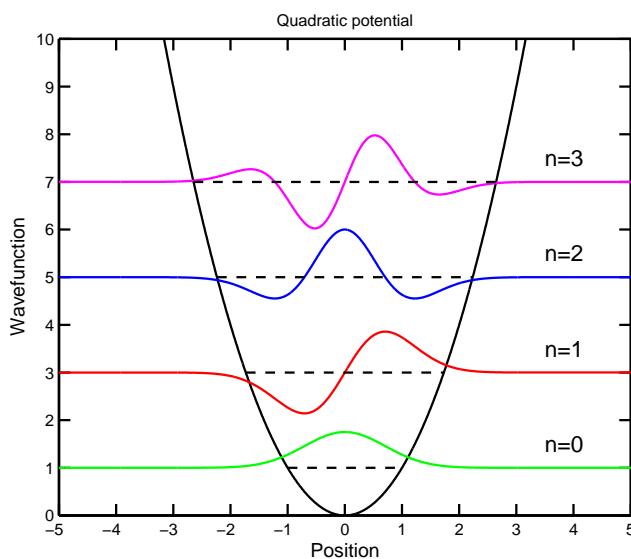


Figure 43: Wavefunctions of a particle in a quadratic potential. These have not been normalised so that the high-order functions can be easily seen. The dashed lines indicate the classical confinement of the particle.

Several low-order solutions are shown in Fig. 43.

Note the following:

- The wavefunctions form an orthonormal set.
- There is a finite chance of finding the particle outside of the classically allowed region.
- The energy levels are uniformly separated by $\hbar\omega$.
- The system can only change state by emitting or absorbing energy in integer units of $\hbar\omega$. Thus, the dynamical behaviour of the ‘mechanical’ system is quantised.
- The lowest energy allowed, $n = 0$, is $\hbar\omega/2$, which is the ground state. Otherwise, the wavefunction would be zero and the particle would not exist at all.
- A stationary particle is not allowed to exist in the well!
- As the order increases, the confinement agrees more closely to the classical limit (*correspondence principle*).

27 Parity

As we shall see, operators play an extensive role in quantum mechanics, and here we consider the notion of parity.

The parity operator \hat{P} has the effect of ‘reflecting’ a wavefunction through the process $x \rightarrow -x$.

\hat{P} is not associated with a dynamical variable, but it is an operator, and has the mathematical properties of an operator.

We have seen for example in the case of the simple harmonic oscillator that each bound eigenstate is either even or odd under this operation:

$$\hat{P}\psi_n = (-1)^{(n-1)}\psi_n \quad \text{where} \quad n \in \{1, 2, 3, \dots, \infty\}. \quad (231)$$

In this case, the lowest order state is indexed by $n = 1$.

Be careful, because often the lowest order state is indexed by 0, and Eq. (231) then needs changing accordingly. This is simply a matter of convention.

This is true in general when the parity operator commutes with the Hamiltonian. Namely, in the cases considered above, when the potential is symmetric under parity transformation, $V(x) = V(-x)$.

Each allowable wave function is an eigenstate of the parity operator, with an eigenvalue of either -1 or $+1$.

Wave functions for which the eigenvalue is $+1$ are called *even*, and wave functions for which the eigenvalue is -1 are called *odd*.

This reasoning follows from the spatial symmetry of the wave function.

A particle bound in a symmetric potential $V(x) = V(-x)$, must be equally likely of being found at $\pm x$.

$$|\psi(-x)|^2 = |\psi(x)|^2. \quad (232)$$

The wavefunction must be even or odd.

In three dimensions, the parity operator \hat{P} has the effect of reflecting a wavefunction through the origin $\mathbf{r} \rightarrow -\mathbf{r}$.

The parity operator is used extensively to explore the nature of possible solutions.

28 The Correspondence Principle

The Correspondence Principle states that ‘quantum mechanics predicts the same observational result as classical physics in the limit of large quantum numbers’.

Originally, it was used to select ‘correct’ quantum models from ‘incorrect’ ones.

It essentially requires a smooth transition to classical behaviour at high energies.



Niels Bohr
(1885-1962)

Consider the case of an infinite square potential well.

Eq. (210) states that the odd-order, even parity modes within the square well are

$$\psi_n(x) = B \cos\left(\frac{n\pi x}{a}\right) \propto [e^{-in\pi x/a} + e^{+in\pi x/a}]. \quad (233)$$

In the momentum representation

$$\begin{aligned}\phi_n(p) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \psi_n(x) e^{-ipx/\hbar} dx \\ &\propto \int_{-a/2}^{+a/2} \left[e^{in\pi x/a} + e^{-in\pi x/a} \right] e^{-ipx/\hbar} dx \\ &= \int_{-\infty}^{+\infty} \left[e^{in\pi x/a} + e^{-in\pi x/a} \right] S(x) e^{-ipx/\hbar} dx,\end{aligned}\quad (234)$$

where $S(x)$ is a top-hat function between $-a/2$ and $a/2$.

The Fourier transform of a product is equal to the convolution of the individual Fourier transforms, and therefore

$$\phi_n(p) \propto \left[\delta\left(\frac{p}{\hbar} - \frac{n\pi}{a}\right) + \delta\left(\frac{p}{\hbar} + \frac{n\pi}{a}\right) \right] \otimes \text{sinc}\left(\frac{pa}{2\hbar}\right). \quad (235)$$

Alternatively we could have derived the same result using the integral:

$$\int_{-\infty}^{+\infty} S(x) e^{-i(\frac{p}{\hbar} \mp \frac{n\pi}{a})x} dx = a \text{sinc}\left[\frac{1}{2}\left(\frac{pa}{\hbar} \mp n\pi\right)\right].$$

$\phi(p)$ is shown in Fig. 44.

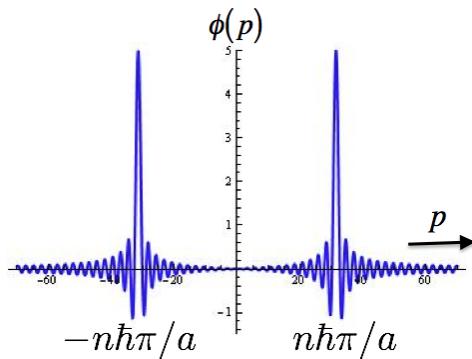


Figure 44: Momentum wave function for a high-energy state in an infinitely deep square potential well.

- The spectrum comprises a sinc function, whose width is inversely proportional to the width of the well, convolved with two delta functions at $p = \pm n\hbar\pi/a$.
- There is an equal chance of measuring the particle travelling in each direction.

- As n increases, the momentum increases, and the uncertainty decreases.
- In this case, $\phi_n(p) \rightarrow \delta(p - p_n) + \delta(p + p_n)$, where $p_n = \pm\sqrt{2mE_n}$, which is the classical limit.

29 Vibrational specific heat of gases

The potential energy associated with the stretching of a diatomic molecule can, for low energies, be modelled as a quadratical potential: Fig. 45.

In this sense, it may be regarded as two masses, m_1 and m_2 , joined by a spring having elastic constant α representing the stiffness of the chemical bond.

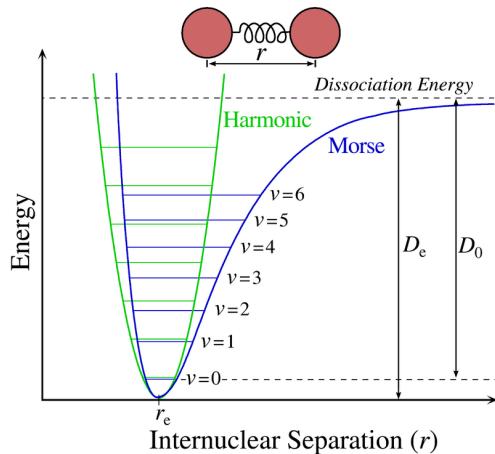


Figure 45: A quadratic well is a reasonable approximation to the Morse potential of a diatomic molecule, although it does not allow for unbound states.

The behaviour can be described by a single harmonic oscillator of reduced mass

$$\mu = \frac{m_1 m_2}{m_1 + m_2}. \quad (236)$$

Classically, each degree of freedom (and therefore each vibrational mode) has thermal energy $\frac{1}{2}k_B T$.

Quantum mechanically, an oscillator has a range of discrete vibrational energy states with $E_n = (n + 1/2)\hbar\omega$,

where $\omega = \sqrt{\alpha/\mu}$.

Classically, each state is excited with a probability distribution given by the Boltzmann distribution:

$$P_n \propto e^{-E_n/k_B T} = e^{-\beta E_n} \quad \text{where } \beta = \frac{1}{k_B T}. \quad (237)$$

Notice that we are now mixing classical thermal statistics and quantum mechanics in a rather casual way. We shall look at a better way of describing this mixture later.

The classical probabilities must sum to unity to be quantitatively meaningful, and normalizing gives

$$P_n = \frac{e^{-\beta E_n}}{\sum_{n=0}^{\infty} e^{-\beta E_n}}. \quad (238)$$

The average vibrational energy per molecule is then

$$\begin{aligned} \langle E \rangle &= \frac{\sum_{n=0}^{\infty} (n + \frac{1}{2}) \hbar \omega e^{-(n+\frac{1}{2}) \hbar \omega \beta}}{\sum_{n=0}^{\infty} e^{-(n+\frac{1}{2}) \hbar \omega \beta}} \\ &= \frac{1}{2} \hbar \omega + \frac{\sum_{n=0}^{\infty} n \hbar \omega e^{-n \hbar \omega \beta}}{\sum_{n=0}^{\infty} e^{-n \hbar \omega \beta}}. \end{aligned} \quad (239)$$

Now let

$$Z = \sum_{n=0}^{\infty} e^{-n \hbar \omega \beta} = \frac{1}{1 - e^{-\hbar \omega \beta}}, \quad (240)$$

which is the partition function, where the second equality follows by summing the geometrical series.

Also

$$-\frac{\partial Z}{\partial \beta} = \sum_{n=0}^{\infty} n \hbar \omega e^{-n \hbar \omega \beta} = \frac{\hbar \omega e^{-\hbar \omega \beta}}{(1 - e^{-\hbar \omega \beta})^2}. \quad (241)$$

$$\langle E \rangle = \frac{1}{2} \hbar \omega + \frac{\hbar \omega}{e^{\hbar \omega \beta} - 1}. \quad (242)$$

Ignoring the zero-point-energy, which is tiny compared with the total energy,

$$\langle E \rangle = \frac{\hbar\omega}{e^{\hbar\omega/k_B T} - 1}, \quad (243)$$

which is the Planck distribution.

It gives the average energy of a quantum, simple harmonic oscillator lightly coupled to a heat bath having temperature T .

The vibrational heat capacity of N diatomic molecules is given by

$$C_{\text{vib}} = N \frac{\partial \langle E \rangle}{\partial T} = N k_B \left(\frac{\hbar\omega}{k_B T} \right)^2 \frac{e^{\hbar\omega/k_B T}}{(e^{\hbar\omega/k_B T} - 1)^2}, \quad (244)$$

which is plotted in Fig. 46.

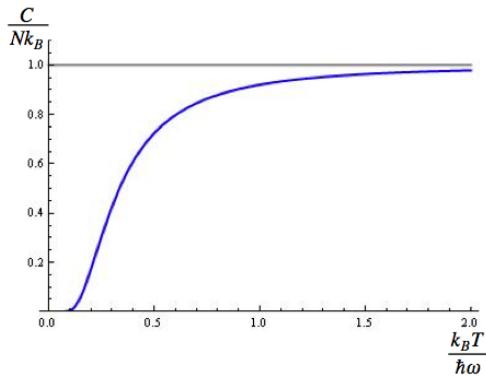


Figure 46: Vibrational heat capacity of N diatomic molecules.

The following is seen:

- C_{vib} tends to the classical limit Nk_B as $T \rightarrow \infty$.
- It is considerably smaller than the classical limit for small values of T .
- The heat capacity essentially becomes the classical value at $\Theta_{\text{vib}} = \hbar\omega/k_B$. Θ_{vib} is sometimes called, confusingly, the *characteristic temperature of vibration*.

- This behaviour is seen in experimental measurements.

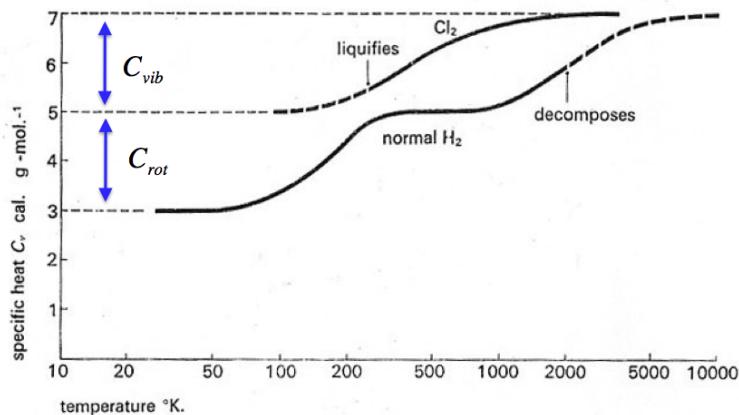


Figure 47: Heat capacities of hydrogen and chlorine as a function of temperature.

The measured heat capacities of hydrogen and chlorine are shown in Fig. 47.

Figure 47 also shows structure arising from rotational (discussed later in the course) and translational behaviour.

The characteristic temperatures of various molecules are listed below. Note that the energies are generally quite high, and vibrational modes are only excited at relatively high temperatures.

Molecule	H ₂	O ₂	N ₂	HCl	CO	NO	Cl ₂
$\Theta_{\text{vib}}(K)$	6140	2239	3352	4150	3080	2690	810

The temperature gives an indication of the stiffness of the bond.

Rotational motion gives rise to much lower energy states.

VI Operator Algebra

In the previous chapters we developed an initial understanding of quantum mechanics, starting from the fundamental physics questions that it aimed to address and continuing to propose a formalism based on wave functions and the Schrödinger equation. We then ventured into solving a range of problems concerning the motion of a particle in 1D, encompassing piece-wise constant potentials and the simple harmonic oscillator. In the process, we encountered and discussed several new and important concepts brought about by quantum mechanics. In this chapter, we shall move beyond the wave function formulation and introduce a new, more powerful description based on vector spaces and operators – a description that has come to take centre stage and is widely used in modern physics. In this language, we shall revisit the notion of position, momentum and energy operators, and discuss the concepts of eigenvalues and eigenvectors, of expectation values and the uncertainty principle (finally formulated in a rigorous way). The full force of quantum mechanics only emerges when the tools of linear operators are brought to bear.

We have hitherto limited ourselves to measurements of position and momentum, but one wonders whether there is a more general formalism that applies to all measurable quantities.

30 Vector spaces

Operators are mathematical procedures for mapping one function onto another, one vector onto another, one ma-

trix onto another, etc.

They are denoted by hats, but sometimes the hat is omitted, and it is left to the context to indicate that the object is an operator.

In these lecture notes, we shall always use hats.

An operator \hat{A} might carry out one of the following mappings:

$$\begin{aligned} \psi(x) &\xrightarrow{\hat{A}} \phi(x) \\ \mathbf{a} &\xrightarrow{\hat{A}} \mathbf{b} \\ \mathbf{A} &\xrightarrow{\hat{A}} \mathbf{B} \\ \hat{A}|a\rangle &= |b\rangle \end{aligned} \tag{245}$$

In the first case, \hat{A} is a recipe for operating on functions to create new functions: differentiation and integration are examples. The Fourier transform is a linear operator.

In the second case, \hat{A} is a mechanism for operating on column vectors to produce new column vectors. Multiplication by a matrix is an example.

In the third case, \hat{A} generates one matrix from another: swapping the columns is an example.

The fourth case is a generic representation that is used to encompass all cases.

Usually, but not always, an operator maps from one kind of object onto an identical kind of object. For example, a square matrix maps a column vector having N elements onto another column vector having N elements.

In this case, the operators are said to map from one vector space onto *the same vector space*.

\hat{A} could, however, map column vectors having M elements onto column vectors having N elements, and in this case the mapping is between *different vector spaces*.

An operator that maps from one vector space to a vector space having a lower dimension, leads to a loss of information.

Thus, the matrix

$$\mathbf{A} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} \times \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix} \quad (246)$$

must be singular because it maps a 3-dimensional space down onto a 2-dimensional space and then back up again.

The matrix

$$\mathbf{A} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \times \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix} \quad (247)$$

need not be singular because it first maps up into a higher dimensional space, and then back down to a lower dimensional space.

There is a remarkable similarity between operators that act on functions and operators that act on discrete vectors.

Consider sampling a function for numerical purposes and listing the samples in a column vector:

$$\psi(x) \equiv \begin{bmatrix} \psi(x_1) \\ \psi(x_2) \\ \vdots \\ \psi(x_N) \end{bmatrix}. \quad (248)$$

We can certainly generate matrices that carry out differentiation, integration, etc., on sampled functions, and the only apparent difference between functions and column vectors is that the former have an infinite number of elements.

One could, conceptually at least, imagine a function to be a column vector in an infinite-dimensional space, and this leads to the notion of function spaces. Different vectors in the space correspond to different functions.

Strictly speaking, column vectors and row vectors are in different vector spaces (you can't add them), and indeed bra ($\langle a |$) and ket ($|a \rangle$) vectors are in different vector spaces, they are called *dual spaces*, but we shall not dwell on the difference.

A general column vector can be written as a weighted linear combination of basis vectors:

$$\mathbf{a} = a_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + a_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad = \sum_{n=1}^N a_n \mathbf{a}_n. \quad (249)$$

Can a function be represented by a weighted linear combination of basis functions?

For certain classes of functions, having certain characteristics, this can be done in a mathematically precise way, but not surprisingly, an infinite number of basis functions is needed:

$$\psi(x) = \sum_{n=1}^{\infty} a_n \psi_n(x). \quad (250)$$

Like discrete column vectors, it is only possible to combine functions of the ‘same kind.’

Eq. (250) raises several issues:

- Why is $\psi(x)$ represented by a discrete sum rather than a continuous superposition of basis functions?
- How do we know that the summation converges?
- What is meant by ‘functions of the same kind’?

The answers to these questions require an understanding of the theory of *Hilbert spaces*, which is beyond the scope of this course, but we can nevertheless build the algebra knowing that the hard work has already been done.



David Hilbert
(1862-1943)

An important class of functions is those that are square integrable:

$$\int_{\mathcal{S}} |\psi(x)|^2 dx < \infty; \quad (251)$$

these can be represented by summations of the kind shown in Eq. (250).

In the case of quantum mechanics, wavefunctions must be square integrable to give finite probabilities, and therefore they can be expanded as infinite sums, and regarded as vectors.

The domain of integration \mathcal{S} is not specified in Eq. (251).

It can have some finite domain, or may extend to infinity, but once the domain is fixed, the basis functions are also defined over the same domain.

The same analysis, and conditions, apply to functions of two variables (functions on 2D surfaces) and of three

variables (functions in 3D volumes).

A set of basis functions for the domain -1 to $+1$ constitutes a different vector space for those covering the domain $-\infty$ to $+\infty$.

A set of 2-dimensional basis functions for a square of side length a constitutes a different vector space than those for a rectangle having side lengths a and b .

Examples of basis functions, and vector spaces, are Legendre polynomials, Hermite functions, Laguerre functions, Bessel functions, spherical harmonics.

In all cases, weighted linear superpositions of vectors can be formed,

$$\begin{aligned}\psi(x) &= \sum_{n=1}^{\infty} a_n \psi_n(x) \\ \mathbf{a} &= \sum_{n=1}^N a_n \mathbf{a}_n \\ \mathbf{A} &= \sum_{n=1}^N a_n \mathbf{A}_n \\ |\mathbf{a}\rangle &= \sum_n a_n |a_n\rangle,\end{aligned}\tag{252}$$

where enough basis vectors must be used to *span* the vector space of interest.

This will ensure that the representation is exact, and the use of '=' is legitimate.

For example, the Pauli spin matrices form an orthonormal basis for 2×2 complex matrices.

It is not surprising that notions such as *inner product*, *orthogonality*, *outer product* can be created for functions on some specified domain.

Inner products are defined through a set of axioms, and different forms of inner product exist depending on the application.

A particularly important kind of *inner product* between two ‘abstract vectors’ is given by

$$\begin{aligned} c &= \int_{\mathcal{S}} \psi^*(x)\phi(x)dx \\ c &= \mathbf{a}^{*T}\mathbf{b} \quad (\text{compare with } \mathbf{a} \cdot \mathbf{b}) \\ c &= \text{Tr} [\mathbf{A}^{*T}\mathbf{B}] \\ c &= \langle a|b \rangle. \end{aligned} \tag{253}$$

where c is a complex number.

The top line is the inner product between two functions; the second line between two column vectors, where $(\dots)^{*T}$ denotes the conjugate transpose; the third line between two matrices; and the bottom line is a generic notation. The conjugate transpose operation is called adjoint or Hermitian conjugate, and it is conventionally denoted by a *dagger*: $(\dots)^{*T} \equiv (\dots)^\dagger$.

$|b\rangle$ is a generic vector, called a *ket vector*.

$\langle a|$ is a generic dual vector, called a *bra vector*. For every ket there is a unique bra.

Together they form a *bra-ket*, $\langle a|b \rangle$, which is a complex number, called the *inner product*.

The precise nature of the vectors and the dual vectors depends on the problem of interest, but the generic notation frees us from worrying about the exact forms.



Paul Dirac
(1902-1984)

Notice that

$$\begin{aligned}\int_{\mathcal{S}} \psi^*(x)\phi(x)dx &= \left(\int_{\mathcal{S}} \phi(x)^*\psi(x)dx \right)^* \\ \mathbf{a}^\dagger \mathbf{b} &= (\mathbf{b}^\dagger \mathbf{a})^* \\ \text{Tr} [\mathbf{A}^\dagger \mathbf{B}] &= \text{Tr} [\mathbf{B}^\dagger \mathbf{A}]^* \\ \langle a|b\rangle &= \langle b|a\rangle^*. \end{aligned}\quad (254)$$

On the basis of the inner product, a meaning can be attributed to *orthogonality*.

Vectors for which the following are true are called *orthogonal*:

$$\begin{aligned}\int_{\mathcal{S}} \psi^*(x)\phi(x)dx &= 0 \\ \mathbf{a}^\dagger \mathbf{b} &= 0 \\ \langle a|b\rangle &= 0. \end{aligned}\quad (255)$$

Likewise the lengths, or *norms*, are given by

$$\begin{aligned}\left[\int_{\mathcal{S}} \psi^*(x)\psi(x)dx \right]^{1/2} &= |\psi(x)| \\ [\mathbf{a}^\dagger \mathbf{a}]^{1/2} &= |\mathbf{a}| \\ [\langle a|a\rangle]^{1/2} &= ||a\rangle|, \end{aligned}\quad (256)$$

where, from Eq. (254), the inner product of an element $|a\rangle$ with itself must be real.

Orthogonal vector space bases can be normalised to give *orthonormal* bases:

$$\begin{aligned}\int_{\mathcal{S}} \psi_n^*(x)\psi_m(x)dx &= \delta_{nm} \\ \mathbf{a}_n^\dagger \mathbf{a}_m &= \delta_{nm} \\ \langle a_n|a_m\rangle &= \delta_{nm}. \end{aligned}\quad (257)$$

If the following are true, a basis set is said to be *complete*, because there are enough basis vectors in the set to span all possible vectors in the space:

$$\sum_{n=1}^{\infty} \psi_n(x) \psi_n^*(y) = \delta(x - y) \quad (258)$$

$$\begin{aligned} \sum_{n=1}^N a_n a_n^\dagger &= I \\ \sum_n |a_n\rangle \langle a_n| &= \hat{I}, \end{aligned}$$

where $\delta(x - y)$ is the Dirac delta function, I is the $N \times N$ identity matrix, and \hat{I} is a generic identity operator.

Eqs. (258) are called *completeness relations*.

Say that we wish to represent a vector in terms of an orthonormal basis set,

$$|a\rangle = \sum_n a_n |a_n\rangle. \quad (259)$$

We can find the expansion coefficients in the following way:

$$\begin{aligned} \langle a_m | a \rangle &= \sum_n a_n \langle a_m | a_n \rangle, \\ &= \sum_n a_n \delta_{mn} \\ &= a_m. \end{aligned} \quad (260)$$

Once the inner product is defined, and a complete set of basis vectors known, we can create linear expansions for any vector in the vector space.

The equality sign in Eq. (260) requires, in general, the basis set to be complete.

The inner product $\langle a_m | a \rangle$ projects the general vector $|a\rangle$ onto the basis vector $|a_m\rangle$ to give a coefficient.

The above formalism applies to systems where only a discrete set of basis functions is needed in order to described precisely any vector in the space. Although, an infinite number of basis functions is often required.

In some cases, a continuous superposition of basis functions is needed.

A bound particle can be in any one of a discrete system of states, whereas an unbound particle can be in any one of a continuous superposition of states.

Operations for dealing with the wavefunctions of unbound systems are essentially devised to deal with mathematically convenient but strictly unphysical situations.

They choose to represent the limit as the region of constraint is made increasingly large. Strictly, this requires a different branch of mathematics called *generalised functions*.

In the case of unbound systems, Eq. (252) no longer applies and a continuous superposition of basis functions must be used:

$$\psi(x) = \int_{-\infty}^{\infty} \phi(k)\chi(k, x) dk, \quad (261)$$

where $\chi(k, x)$ are the basis functions, labelled through the continuous variable k , and weighted by the factor $\phi(k)dk$.

Eq. (261) is still a weighted linear combination of basis functions, and should be compared with Eq. (252).

The orthogonality condition (255) between two different

basis functions labelled k_1 and k_2 , becomes

$$\int_{-\infty}^{+\infty} \chi^*(k_1, x) \chi(k_2, x) dx = \delta(k_2 - k_1). \quad (262)$$

The norm should be given by setting $k_1 = k_2$, but this is not possible, and so the basis functions cannot be normalised.

The completeness relation (258) becomes

$$\int_{-\infty}^{+\infty} \chi^*(k, x_1) \chi(k, x_2) dk = \delta(x_2 - x_1). \quad (263)$$

An obvious solution to Eqs. (262) and (263) is

$$\chi(k, x) = \frac{1}{\sqrt{2\pi}} e^{ikx}. \quad (264)$$

Eq. (264) is a suitable distribution of basis functions to describe the behaviour of unbound systems.

Then, Eq. (261) gives

$$\psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(k) e^{ikx} dk, \quad (265)$$

and we have generated the Fourier transform.

Note that the basis functions $e^{ikx}/\sqrt{2\pi}$ are simultaneous eigenfunctions of the operators

$$-i \frac{d}{dx} \quad \text{and} \quad -i \frac{d}{dk}. \quad (266)$$

It seems reasonable that we could represent any function as a weighted linear combination of continuously displaced delta functions $\psi(x) = \int \phi(x') \delta(x' - x) dx'$. Certainly the ‘coefficients’ are simply $\phi(x') = \psi(x')$. Attempt to apply the orthogonality and completeness relations. Is this straightforward, or are there any problems?

31 Operators

In addition to vectors, we need *operators*.

Operators map one vector onto another vector:

$$\begin{aligned}\psi(y) &= \int_S A(y, x)\phi(x)dx & (267) \\ \mathbf{b} &= \mathbf{A}\mathbf{a} \\ |b\rangle &= \hat{A}|a\rangle.\end{aligned}$$

Operators act on vectors to generate new vectors, but they ‘must’, for our purposes, map onto a vector in the same vector space.

For example, in the space of square integrable functions over some domain, the operator ‘must’ create another square integrable function over the *same domain*.

If we mapped a function onto a smaller domain, and then back up again onto a larger domain, information may be lost. Such operators certainly exist, but we shall not consider them.

The inner product ‘projects’ one vector onto another, but we can also form *outer* products:

$$\begin{aligned}\psi(x)\phi^*(y) & & (268) \\ \mathbf{b}\mathbf{a}^\dagger & \\ |b\rangle\langle a|, &\end{aligned}$$

which are operators.

The outer products between higher-order objects such as matrices are called *tensor products*.

We can add operators,

$$\hat{A}|a\rangle + \hat{B}|a\rangle = [\hat{A} + \hat{B}]|a\rangle \equiv \hat{C}|a\rangle, \quad (269)$$

and $\hat{A} + \hat{B} = \hat{B} + \hat{A}$.

We can also create composite operators,

$$\hat{A}\hat{B}|a\rangle \equiv \hat{C}|a\rangle, \quad (270)$$

but in general they do not commute,

$$\hat{A}\hat{B}|a\rangle \neq \hat{B}\hat{A}|a\rangle, \quad \forall|a\rangle, \quad (271)$$

which is consistent with the observation that the order of two experiments matters.

A convenient tool for determining the degree to which operators commute is called the *commutator*,

$$\boxed{[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}}, \quad (272)$$

where square brackets is the commonly adopted notation.

The commutator is itself an operator, and equal to the null operator if the operators commute.

Suppose that we have some operator \hat{A} , and a complete set of vectors $\{|u_n\rangle\}$. Then,

$$\begin{aligned} \hat{A}|a\rangle &= \hat{I}\hat{A}\hat{I}|a\rangle = \sum_{n,m} |u_n\rangle\langle u_n|\hat{A}|u_m\rangle\langle u_m|a\rangle \\ &= \sum_{n,m} \langle u_n|\hat{A}|u_m\rangle|u_n\rangle\langle u_m|a\rangle \\ &= \sum_{n,m} A_{nm}|u_n\rangle\langle u_m|a\rangle \\ \hat{A} &\equiv \sum_{n,m} A_{nm}|u_n\rangle\langle u_m|, \end{aligned} \quad (273)$$

where $A_{nm} \equiv \langle u_n|\hat{A}|u_m\rangle$ are called the *matrix elements* of \hat{A} .

They constitute the expansion coefficients of an operator, as distinct from the expansion coefficients of a vector.

The last line in Eq. (273) follows from the fact that the previous set of equalities hold for any vector $|a\rangle$, and therefore they must hold at the operator level.

For example, in the case of square integrable functions

$$\begin{aligned} A_{nm} &= \langle u_n | \hat{A} | u_m \rangle \\ &= \int_S \psi_n^*(x) \hat{A} \psi_m(x) dx, \end{aligned} \quad (274)$$

and

$$\hat{A} \equiv \sum_{n,m} A_{nm} \psi_n(x) \psi_m^*(x). \quad (275)$$

Thus we can describe any operator in terms of a linear combination of the outer products of a complete set of basis vectors.

Eq. (273) simply projected the general vector $|a\rangle$ onto each of the basis vectors $|u_m\rangle$ in a complete set, and then reconstruct the answer through the appropriate weighted linear combination of the same basis vectors using the coefficients A_{nm} :

$$\hat{A}|a\rangle = \sum_{n,m} |u_n\rangle A_{nm} \langle u_m|a\rangle. \quad (276)$$

Eq. (258) gave the generalised completeness relationship as $\sum_n |a_n\rangle\langle a_n| = \hat{I}$. From an operator perspective why is it obvious that this expression must hold if the basis set $\{|a_n\rangle, \forall n\}$ is to be complete?

Any operator having the form

$$\hat{P} = |\psi\rangle\langle\psi|, \quad (277)$$

is called a *projection operator*:

$$\begin{aligned}\hat{P}|a\rangle &= |\psi\rangle\langle\psi|a\rangle \\ &= a|\psi\rangle.\end{aligned} \quad (278)$$

It projects the operand $|a\rangle$ onto $|\psi\rangle$, and returns $|\psi\rangle$ multiplied by the appropriate weighting $a \equiv \langle\psi|a\rangle$.

If $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ are Cartesian unit vectors, explain why $\hat{\mathbf{i}}\hat{\mathbf{i}}^T$, $\hat{\mathbf{j}}\hat{\mathbf{j}}^T$ and $\hat{\mathbf{k}}\hat{\mathbf{k}}^T$ are each projection operators, whereas terms such as $\hat{\mathbf{i}}\hat{\mathbf{j}}^T$ are not. Operators of this kind are called Dyads.

The concept of operators as mappings between basis vectors means that we can construct operators having almost any behaviour.

Construct an operator \hat{A} for which repeated application generates the following sequence:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \xrightarrow{\hat{A}} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \xrightarrow{\hat{A}} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \xrightarrow{\hat{A}} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \dots \quad (279)$$

The operator counts down through orthogonal vectors, until it becomes ‘stuck’ at the bottom.

From a mapping perspective, it is clear that

$$A = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} [1 \ 0 \ 0] + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} [0 \ 1 \ 0] + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} [0 \ 0 \ 1]. \quad (280)$$

In each term, the row vector identifies the input argument, and the column vector returns the appropriate answer.

Eq. (280) can be written as

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}. \quad (281)$$

Repeated application of \mathbf{A} verifies the behaviour.

Operators of this kind exist in quantum mechanics. They are called *ladder operators* or *shift operators* depending on whether they act on discrete vectors or continuous functions.

Define the *adjoint* of an operator:

$$\langle u_m | \hat{A} | u_n \rangle = \langle u_n | \hat{A}^\dagger | u_m \rangle^* \quad \forall m, n \quad (282)$$

Note the change in the order of the indices, and the complex conjugate.

The motivation to introduce the adjoint operator is that we know how to generate a ket from the operation $|a\rangle = \hat{A}|u_n\rangle$, and we need to know what operator we must apply to $\langle u_n|$ to generate the associated bra $\langle a|$, whilst ensuring the integrity of the inner product.

\hat{A}^\dagger is called the adjoint of \hat{A} .

The matrix elements of \hat{A} are

$$[\hat{A}]_{mn} = \langle u_m | \hat{A} | u_n \rangle = A_{mn} \quad (283)$$

whereas the matrix elements of \hat{A}^\dagger are, from Eq. (282),

$$[\hat{A}^\dagger]_{mn} = \langle u_m | \hat{A}^\dagger | u_n \rangle = A_{nm}^*. \quad (284)$$

Thus we know the matrix elements of the adjoint once the matrix elements of the operator itself are known.

Given that the matrix elements of an operator in some

complete basis fully define the operator, and that the matrix elements of the adjoint can be found, the adjoint is fully defined.

Every linear operator has a unique adjoint.

For example, the adjoint of the ‘multiplication by a complex number’ is the ‘multiplication by the conjugate of the complex number’.

The adjoint of a general matrix is the conjugate transpose of the matrix.

Other operators have more complicated adjoints, but nevertheless they can be found from the definition above.

Consider an operator acting on $|\psi\rangle$ to produce $|\phi\rangle$:

$$\boxed{\hat{A}|\psi\rangle = |\phi\rangle.} \quad (285)$$

Then for any $|\chi\rangle$,

$$\begin{aligned} \langle\chi|\hat{A}|\psi\rangle &= \langle\chi|\phi\rangle \\ \langle\chi|\hat{A}|\psi\rangle^* &= \langle\phi|\chi\rangle \\ \langle\psi|\hat{A}^\dagger|\chi\rangle &= \langle\phi|\chi\rangle, \end{aligned} \quad (286)$$

and therefore we conclude that

$$\boxed{\langle\psi|\hat{A}^\dagger = \langle\phi|,} \quad (287)$$

which is the dual relationship to Eq. (285).

This equivalence can be appreciated by taking the operators to be matrices, the ket vectors to be column vectors, and the bra vectors to be row vectors.

An important identity for all operators, which we shall use later, is

$$\boxed{(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger,} \quad (288)$$

which follows the same form as the conjugate transpose of a product of matrices.

We can prove Eq. (288) as follows.

According to the definition of the inner product

$$\begin{aligned}\langle u_n | (\hat{A}\hat{B})^\dagger | u_m \rangle^* &= \langle u_m | (\hat{A}\hat{B}) | u_n \rangle \quad \forall m, n \quad (289) \\ &= \sum_r \langle u_m | \hat{A} | u_r \rangle \langle u_r | \hat{B} | u_n \rangle \quad \forall m, n \\ &= \sum_r [\langle u_n | \hat{B}^\dagger | u_r \rangle \langle u_r | \hat{A}^\dagger | u_m \rangle]^* \quad \forall m, n \\ &= \langle u_n | \hat{B}^\dagger \hat{A}^\dagger | u_m \rangle^* \quad \forall m, n,\end{aligned}$$

and we conclude that

$$\langle u_n | (\hat{A}\hat{B})^\dagger | u_m \rangle = \langle u_n | \hat{B}^\dagger \hat{A}^\dagger | u_m \rangle \quad \forall m, n. \quad (290)$$

Because the matrix elements of an operator completely define the action of the operator, and because the matrix elements of $[\hat{A}\hat{B}]^\dagger$ and $\hat{B}^\dagger \hat{A}^\dagger$ are the same, they must be the same operator, yielding Eq. (288).

Operators can have unexpected behaviour; for example, we can create an operator that cycles around a set of orthogonal vectors:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \xrightarrow{\hat{A}} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \xrightarrow{\hat{A}} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \xrightarrow{\hat{A}} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \dots. \quad (291)$$

From a mapping perspective, it is clear that

$$A = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} [1 \ 0 \ 0] + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} [0 \ 1 \ 0] + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} [0 \ 0 \ 1], \quad (292)$$

which becomes

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}. \quad (293)$$

Repeated application of \mathbf{A} verifies the behaviour.

Now form the adjoint of \mathbf{A} by taking the conjugate transpose:

$$A^\dagger = A^T = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}. \quad (294)$$

Repeated application for the adjoint forms the sequence

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \xrightarrow{\hat{A}^\dagger} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \xrightarrow{\hat{A}^\dagger} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \xrightarrow{\hat{A}^\dagger} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \dots \quad (295)$$

The sequence ‘cycles’ in the opposite direction.

There is a displacement operator that shifts continuous functions by some amount along the x axis, and the adjoint moves the function back in the other direction! Equivalently, a shift operator can shift states spatially.

In quantum mechanics there is a *time evolution operator* that moves states forwards in time; its adjoint moves states backwards in time: effectively bringing about a time reversal.

For good physical reasons, which will be discussed later, quantum mechanics is often associated with operators that are *self adjoint*, which means

$$\langle u_m | \hat{A} | u_n \rangle = \langle u_n | \hat{A} | u_m \rangle^*, \quad (296)$$

or

$$\hat{A}^\dagger = \hat{A}. \quad (297)$$

In other words all the matrix elements of \hat{A} and \hat{A}^\dagger are identical.

Multiplication by a real number is a Hermitian operator, and a Hermitian matrix (one that is equal to its

own conjugate transpose) is a Hermitian operator.

There are many others.

The theory of linear operators develops into a large and elegant study.

The crucial point is that all of the key elements of quantum mechanics can be cast into the language of linear operators, and therefore can be applied to continuous systems and systems where there is only a finite number of discrete possible states.

In this section we have seen that there are essentially three parallel mathematical schemes.

They give rise to three different flavours of quantum mechanics:

- Continuous functions, and integral and differential operators, give rise to *Schrödinger's Wave Mechanics*.
- Discrete column and row vectors, and matrices, give rise to *Heisenberg's Matrix Mechanics*.
- Bra and Ket vectors, and abstract operators, give rise to *Dirac's Quantum Mechanics*.

To a large extent they can be regarded as different mathematical representations of the same physical phenomena.

32 Hermitian operators: position, momentum, energy

In quantum mechanics, the state of a system is characterised by its wave function.

Dynamical variables are associated with operators.

Unlike classical physics, the state is not characterised by the variables themselves.

In Handout III, we met several one-dimensional operators:

$$\hat{x} \equiv x \quad \text{Position Operator} \quad (298)$$

$$\hat{p} \equiv -i\hbar \frac{\partial}{\partial x} \quad \text{Momentum Operator} \quad (299)$$

$$\frac{\hat{p}^2}{2m} \equiv -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \quad \text{Kinetic Energy Operator} \quad (300)$$

$$\hat{E} \equiv i\hbar \frac{\partial}{\partial t} \quad \text{Total Energy Operator} \quad (301)$$

$$\hat{H} \equiv -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \quad \text{Hamiltonian.} \quad (302)$$

These can operate on the position wavefunction $\Psi(x, t)$ to give the expectation values of outcomes.

Wavefunctions are vectors in the space of square-integrable functions, and so we must use the inner product given in Eq. (253),

$$\langle \Phi | \Psi \rangle = \int_{-\infty}^{+\infty} \Phi^*(x, t) \Psi(x, t) dx, \quad (303)$$

where time, t , is regarded as a parameter.

The inner product is a function of time, and time is a parameter in quantum mechanics.

In contemporary formulations of quantum mechanics, there is no operator that represents the measurement of time. Time is not a dynamical variable; it is a parameter that provides a backdrop against which the dynamics of a system plays out.

This distinction between space and time in non-relativistic quantum mechanics causes a problem when formulating relativistic quantum mechanics, where space-time is a unified entity. We shall not consider relativistic quantum mechanics in this course.

A fundamental tenet of quantum mechanics is that the act of measuring a physical quantity is represented by a self-adjoint, or equivalently, *Hermitian* operator.

Let us check whether the one-dimensional momentum operator \hat{p} is indeed Hermitian.

For notational brevity we shall assume time independence.

According to Eq. (296), an operator \hat{A} is Hermitian if

$$\int_{-\infty}^{+\infty} \phi^*(x) [\hat{A}\psi(x)] dx = \left[\int_{-\infty}^{+\infty} \psi^*(x) [\hat{A}\phi(x)] dx \right]^*. \quad (304)$$

In the case of the momentum operator:

$$\begin{aligned}
 \int_{-\infty}^{+\infty} \phi^*(x) [\hat{p}\psi(x)] dx &\equiv -i\hbar \int_{-\infty}^{+\infty} \phi^*(x) \frac{d\psi(x)}{dx} dx \\
 &= -i\hbar [\phi(x)^* \psi(x)]_{-\infty}^{+\infty} + i\hbar \int_{-\infty}^{+\infty} \psi(x) \frac{d\phi^*(x)}{dx} dx \\
 &= i\hbar \int_{-\infty}^{+\infty} \psi(x) \frac{d\phi^*(x)}{dx} dx \\
 &= \int_{-\infty}^{+\infty} \psi(x) \left[-i\hbar \frac{d\phi(x)}{dx} \right]^* dx \\
 &= \int_{-\infty}^{+\infty} \psi(x) [\hat{p}\phi(x)]^* dx \\
 &= \left[\int_{-\infty}^{+\infty} \psi^*(x) [\hat{p}\phi(x)] dx \right]^*,
 \end{aligned} \tag{305}$$

where the third line follows because $\psi(x) \rightarrow 0$ and $\phi(x) \rightarrow 0$ for $x \rightarrow \infty$, for square-integrable functions.

Eq. (305) is true for all square-integrable functions, and so \hat{p} is Hermitian: $\hat{p} = -i\hbar\partial/\partial x$ and $\hat{p}^\dagger = -i\hbar\partial/\partial x$.

Consider the position operator $\hat{x} = x$, where x is real.

$$\begin{aligned}
 \int_{-\infty}^{+\infty} \phi^*(x) [\hat{x}\psi(x)] dx &= \int_{-\infty}^{+\infty} \phi^*(x) [x\psi(x)] dx \\
 &= \int_{-\infty}^{+\infty} \psi(x) x \phi(x)^* dx \\
 &= \left[\int_{-\infty}^{+\infty} \psi^*(x) x \phi(x) dx \right]^* \\
 &= \left[\int_{-\infty}^{+\infty} \psi^*(x) [\hat{x}\phi(x)] dx \right]^*,
 \end{aligned} \tag{306}$$

which shows that \hat{x} is Hermitian: $\hat{x} = x$ and $\hat{x}^\dagger = x$.

It also follows straightforwardly that the potential operator is Hermitian:

$$\hat{V}(x) = \hat{V}^\dagger(x) = V(x). \tag{307}$$

Can we show that the Hamiltonian is Hermitian?

Note that

$$\begin{aligned} [\hat{A}\hat{B}]^\dagger &= \hat{B}^\dagger\hat{A}^\dagger \\ [\hat{p}\hat{p}]^\dagger &= \hat{p}^\dagger\hat{p}^\dagger \\ &= \hat{p}\hat{p}, \end{aligned} \tag{308}$$

and therefore \hat{p}^2 is Hermitian.

Because \hat{p}^2 is Hermitian and $V(x)$ is Hermitian, \hat{H} is Hermitian,

$$\hat{H} \equiv \frac{\hat{p}^2}{2m} + V(x) = \hat{H}^\dagger, \tag{309}$$

as expected for an observable.

This whole scheme can be extended to three dimensions with identical results, but now the inner product must relate to integration over a volume.

The reason why observables are represented by Hermitian operators is that their expectation values, the measured values, are always real.

For example,

$$\begin{aligned} \langle A \rangle &= \int_{-\infty}^{+\infty} \psi^*(x)\hat{A}\psi(x)dx \\ &= \left[\int_{-\infty}^{+\infty} \psi^*(x)\hat{A}^\dagger\psi(x)dx \right]^* \\ &= \left[\int_{-\infty}^{+\infty} \psi^*(x)\hat{A}\psi(x)dx \right]^* \\ &= \langle A \rangle^*, \end{aligned} \tag{310}$$

and $\langle A \rangle$ must be real.

Rather than writing out the integrals, it is more efficient, and general, to employ Dirac's notation:

$$\begin{aligned}\langle A \rangle &= \langle \psi | \hat{A} | \psi \rangle \\ &= \langle \psi | \hat{A}^\dagger | \psi \rangle^* \\ &= \langle \psi | \hat{A} | \psi \rangle^* \\ &= \langle A \rangle^*.\end{aligned}\tag{311}$$

33 Expectation values and uncertainty

If a position measurement is carried out multiple times on a system in some pre-prepared quantum state, the average value of the measured quantity, or the *expectation value*, is given by

$$\langle x \rangle = \int_{-\infty}^{+\infty} x |\psi(x)|^2 dx.\tag{312}$$

We also have expressions such as

$$\langle x^2 \rangle = \int_{-\infty}^{+\infty} x^2 |\psi(x)|^2 dx.\tag{313}$$

Be careful, $\langle \dots \rangle$ without a vertical bar corresponds to an expectation value not an inner product.

In a more general form, quantum mechanics has at its heart the following postulate:

For any observable A there is an associated operator \hat{A} which denotes the act of measurement.

If a system is in quantum state $\psi(x)$, then the expectation value of the measurement is

$$\langle A \rangle = \int_{-\infty}^{+\infty} \psi^*(x) \hat{A} \psi(x) dx.\tag{314}$$

or in Dirac notation

$$\langle A \rangle = \langle \psi | \hat{A} | \psi \rangle.\tag{315}$$

The notation for the expectation value changes between texts. Sometimes it is written $\langle A \rangle$, and sometimes $\langle \hat{A} \rangle$.

Each notation has its advantages and disadvantages, but they both mean the average of a real number: the ensemble average of a set of experiments.

We also have expressions such as

$$\boxed{\langle A^2 \rangle = \int_{-\infty}^{+\infty} \psi^*(x) \hat{A}^2 \psi(x) dx,} \quad (316)$$

or in Dirac notation

$$\boxed{\langle A^2 \rangle = \langle \psi | \hat{A}^2 | \psi \rangle.} \quad (317)$$

By \hat{A}^2 we mean that the operator must be applied twice.

The uncertainty in any measurement is then

$$\Delta A = \sqrt{\langle A^2 \rangle - \langle A \rangle^2}. \quad (318)$$

34 Eigenvectors and eigenstates

The eigenvectors of an operator are those vectors that remain unchanged in form, or equivalently in ‘direction’, when the operator is applied:

$$\begin{aligned} \int_S A(y, x) \phi_n(x) dx &= a_n \phi_n(y) \\ \hat{A} \phi_n(x) &= a_n \phi_n(x) \\ \mathbf{A} \mathbf{a}_n &= a_n \mathbf{a}_n \\ \hat{A} |a_n\rangle &= a_n |a_n\rangle, \end{aligned} \quad (319)$$

where a_n is the eigenvalue associated with the n ’th eigenvector $|a_n\rangle$.

The eigenvectors of operators have a special role to play in quantum mechanics.

In general, an operator has multiple eigenvectors and complex eigenvalues.

When an operator acts on one of its eigenvectors, the eigenvector is returned, multiplied by some complex factor.

Notice that the eigenvector is undefined with respect to multiplication by a complex factor.

Only its ‘direction’ remains unchanged.

Explain, by expanding a vector in terms of a complete set of basis states, what is meant by the statement ‘the direction remains unchanged?’

In other words eigenvectors are unchanged in direction, and only change in magnitude and phase, when the operator is applied.

The *state* of a quantum system only depends on the *direction* of the state vector, and the state is not changed by multiplying by a constant.

Indeed the length of the vector is usually normalised to unity.

It turns out that the eigenvectors form a complete set of orthogonal vectors for the space, and therefore can be used as a convenient basis.

If we use the eigenvectors as the basis for describing

an operator,

$$\begin{aligned}
 \hat{A}|a\rangle &= \hat{I}\hat{A}\hat{I}|a\rangle = \sum_{n,m} |a_n\rangle\langle a_n|\hat{A}|a_m\rangle\langle a_m|a\rangle & (320) \\
 &= \sum_{n,m} \langle a_n|\hat{A}|a_m\rangle|a_n\rangle\langle a_m|a\rangle \\
 &= \sum_{n,m} a_m \langle a_n|a_m\rangle|a_n\rangle\langle a_m|a\rangle \\
 &= \sum_{n,m} a_m \delta_{mn} |a_n\rangle\langle a_m|a\rangle \\
 &= \sum_m a_m |a_m\rangle\langle a_m|a\rangle \\
 \hat{A} &= \sum_m a_m |a_m\rangle\langle a_m|,
 \end{aligned}$$

and the operator is said to be in its *diagonal form*.

The process works by projecting the original vector $|a\rangle$ onto the eigenvectors, scaling by the eigenvalues, and then reconstructing the result as a weighted linear combination of eigenvectors. Of course it does!

We can now readily write the adjoint of \hat{A} as

$$\hat{A}^\dagger = \sum_m a_m^* |a_m\rangle\langle a_m|, \quad (321)$$

as can be obtained by noting that the matrix elements of \hat{A} are the a_m 's, and using Eq. (282).

In the case of a self-adjoint operator, this is equal to \hat{A} , and the *eigenvalues must be real*.

Operators often have a number of eigenvectors that have the same eigenvalue, and the vectors, or states, are called *degenerate*.

This has significant consequences in quantum mechanics.

Suppose that we have two *degenerate* states, $|a_1\rangle$ and $|a_2\rangle$.

Any combination $|a\rangle = c_1|a_1\rangle + c_2|a_2\rangle$ of these states is also an eigenstate

$$\begin{aligned}\hat{A}|a\rangle &= \hat{A}(c_1|a_1\rangle + c_2|a_2\rangle) & (322) \\ &= c_1a_1|a_1\rangle + c_2a_2|a_2\rangle \\ &= a_1(c_1|a_1\rangle + c_2|a_2\rangle) \\ &= a_1|a\rangle,\end{aligned}$$

where the second-to-last line follows because $a_1 = a_2$.

If N vectors are degenerate, other sets of N orthogonal eigenvectors can be constructed.

Be careful: software, such as MatLab, calculates the eigenvectors of matrices, but if any of these are degenerate, they can be *rotated* into other perfectly good eigenvectors, and the spatial forms that you calculate may not look like the ones you expect!

Previously, see Eqs. (285) and (287), we commented on the duality of the mappings

$$\hat{A}|\psi\rangle = |\phi\rangle \text{ and } \langle\psi|\hat{A}^\dagger = \langle\phi|, \quad (323)$$

which in the context of eigenvectors become

$$\hat{A}|a_n\rangle = a_n|a_n\rangle \text{ and } \langle a_n|\hat{A}^\dagger = a_n^*\langle a_n|. \quad (324)$$

According to Eq. (315), if the system is in state $|a_n\rangle$ then the expectation value of \hat{A} is

$$\begin{aligned}\langle A \rangle &= \langle a_n|\hat{A}|a_n\rangle & (325) \\ &= a_n\langle a_n|a_n\rangle \\ &= a_n,\end{aligned}$$

where we have assumed that the eigenvector is normalised, $\langle a_n | a_n \rangle = 1$.

If a system is in one of the eigenvectors of some operator, which is associated with some observable, then the expectation value of the measured observable is the eigenvalue.

What is the spread of a measurement?

According to Eq. (318),

$$\begin{aligned} \langle \Delta A \rangle^2 &= \langle \psi | \hat{A}^2 | \psi \rangle - \left(\langle \psi | \hat{A} | \psi \rangle \right)^2 \\ &= \langle a_n | \hat{A}^2 | a_n \rangle - \left(\langle a_n | \hat{A} | a_n \rangle \right)^2 \\ &= a_n^2 - a_n^2 \\ &= 0. \end{aligned} \quad (326)$$

Thus there is no uncertainty in the measurement.

For example, suppose that we measure the momentum of a particle having the wavefunction

$$\psi(x) = e^{ikx}, \quad (327)$$

then

$$\hat{p}|\psi\rangle = -i\hbar \frac{d}{dx} e^{ikx} = \hbar k |\psi\rangle = p |\psi\rangle. \quad (328)$$

Indeed, e^{ikx} is an eigenfunction of the momentum operator $-i\hbar d/dx$.

$$\begin{aligned} (\Delta p)^2 &= \langle \psi | \hat{P}^2 | \psi \rangle - \left(\langle \psi | \hat{P} | \psi \rangle \right)^2 \\ &= 0, \end{aligned} \quad (329)$$

although we have swept under the carpet the problem of normalising the wavefunction.

For this wavefunction, a measurement of momentum yields a particular value with certainty.

In general, if a system is in one of the eigenstates of an observable, the numerical value of a measurement is the eigenvalue, and this value is measured with certainty.

35 Ordered measurements

The order in which measurements are made, or equivalently the order in which operators are applied, matters.

For example,

$$\begin{aligned}\hat{p}\hat{x}|\psi\rangle &\equiv -i\hbar \frac{d}{dx} [x\psi(x)] & (330) \\ &= -i\hbar \left[\psi(x) + x \frac{d\psi(x)}{dx} \right] \\ &= -i\hbar\psi(x) + \hat{x}\hat{p}\psi(x). \\ &\equiv -i\hbar|\psi\rangle + \hat{x}\hat{p}|\psi\rangle.\end{aligned}$$

Thus

$$\hat{x}\hat{p} - \hat{p}\hat{x} = [\hat{x}, \hat{p}] = i\hbar. \quad (331)$$

For any pair of operators \hat{A} and \hat{B} , the commutator \hat{C} , which is itself an operator, is defined by

$$\boxed{\hat{C} = [\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}.} \quad (332)$$

Because $[\hat{x}, \hat{p}] \neq 0$ position and momentum do not commute.

Eq. (331) is important because it shows that it is not possible to measure position and momentum without one measurement perturbing the other.

The degree to which they do not commute is ‘quantified’ by \hbar .

The order in which operators are applied is crucial, and to a large extent is responsible for the difference between

quantum mechanics and classical mechanics.

Because the order of operators largely determines the difference between quantum and classical measurements, a whole algebra has been built up around commutators.

Note that

$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger \begin{cases} \neq \hat{B}\hat{A} \neq \hat{A}\hat{B} & \text{in general} \\ = \hat{B}\hat{A} & \text{if } \hat{A} \text{ and } \hat{B} \text{ are Hermitian, but } \neq \hat{A}\hat{B} \\ = \hat{B}\hat{A} = \hat{A}\hat{B} & \text{if } \hat{A} \text{ and } \hat{B} \text{ additionally commute} \end{cases} \quad (333)$$

It follows that

$$(\hat{A}\hat{B})^\dagger = \hat{A}\hat{B} \text{ only if } \hat{A}^\dagger = \hat{A}, \hat{B}^\dagger = \hat{B}, \text{ and } [\hat{A}, \hat{B}] = \hat{0}.$$

(334)

For a composite operator to be Hermitian, its individual parts must be Hermitian, and its individual parts must commute.

For a composite operator to be an observable, its individual parts must be observables, *and* they must commute.

So, the composite operator $\hat{x}\hat{p}$ is not an observable, as we have seen.

There is an interesting twist, because if \hat{A} and \hat{B} are Hermitian, but do not commute, we can create a composite operator that does commute, and so is an observable.

This gives rise to the concept of the *anti-commutator*:

$$\hat{C} = \{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}.$$

(335)

This can be shown as follows:

$$\begin{aligned}\hat{C}^\dagger &= \hat{B}^\dagger \hat{A}^\dagger + \hat{A}^\dagger \hat{B}^\dagger \\ &= \hat{B} \hat{A} + \hat{A} \hat{B} \\ &= \hat{C}.\end{aligned}\tag{336}$$

It follows that $\hat{x}\hat{p} + \hat{p}\hat{x}$ is an observable, but what is it?

If \hat{A} and \hat{B} are two non-commuting observables, we can create an operator, through the anticommutator, that is an observable.

The commutator of two non-commuting Hermitian operators is not itself Hermitian:

$$\begin{aligned}[\hat{A}, \hat{B}]^\dagger &= (\hat{A}\hat{B} - \hat{B}\hat{A})^\dagger \\ &= (\hat{A}\hat{B})^\dagger - (\hat{B}\hat{A})^\dagger \\ &= \hat{B}^\dagger \hat{A}^\dagger - \hat{A}^\dagger \hat{B}^\dagger \\ &= \hat{B}\hat{A} - \hat{A}\hat{B} \\ &= - [\hat{A}, \hat{B}] \\ &\neq [\hat{A}, \hat{B}]\end{aligned}\tag{337}$$

But, the commutator of two non-commuting Hermitian operators can be made Hermitian by multiplying by i :

$$\begin{aligned}\hat{C}^\dagger &= (i [\hat{A}, \hat{B}])^\dagger \\ &= -i [\hat{A}, \hat{B}]^\dagger \\ &= i [\hat{A}, \hat{B}] \\ &= \hat{C}.\end{aligned}\tag{338}$$

Much of the mathematical structure of quantum mechanics comes from the constraints imposed by the commutator and anticommutator; indeed, their differences account for the differences in behaviour of two major classes of particles: bosons and fermions.

36 Energy eigenstates

Return to the time-dependent Schrödinger equation

$$i\hbar \frac{\partial \Psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi(x, t)}{\partial x^2} + V(x, t)\Psi(x, t). \quad (339)$$

The time-independent Schrödinger equation is applicable when the potential is independent of time:

$$i\hbar \frac{\partial \Psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi(x, t)}{\partial x^2} + V(x)\Psi(x, t), \quad (340)$$

which in state-vector notation becomes

$$i\hbar \frac{\partial}{\partial t} |\Psi(x, t)\rangle = \hat{H} |\Psi(x, t)\rangle. \quad (341)$$

However we know that in this case

$$|\Psi(x, t)\rangle = e^{-iEt/\hbar} |\Psi(x)\rangle. \quad (342)$$

and so

$$\hat{H} |\Psi(x, t)\rangle = E |\Psi(x, t)\rangle. \quad (343)$$

The measured energy E is an eigenvalue of the Hamiltonian operator \hat{H} . For this state, the total energy E is a constant.

A measurement of energy will definitely return E without uncertainty. Because the potential is independent of time, energy is conserved and its value can be measured without uncertainty.

The eigenvalues of the Hamiltonian are the energy levels of the system. The eigenvectors are called *stationary states* because they are states having a definite energy, and therefore state vectors of the form of (342).

Also note that in this case the probability density is independent of time:

$$|\Psi(x, t)|^2 = |\Psi(x)|^2. \quad (344)$$

37 Repeated measurements

Eigenvector analysis leads to an important idea in quantum mechanics.

Suppose that some observable is measured and then measured again immediately afterwards.

Once the first measurement has been made, we know what the value of the second measurement will be with certainty. Thus, after the first measurement, the system must be in the state represented by the eigenvector corresponding to the eigenvalue just measured.

When an observable is first measured, the wavevector, or equivalently wavefunction, collapses into the eigenstate of the value measured. Multiple repeated measurements yield the same measured value with certainty.

This reasoning strongly suggests that when an observable is measured, the outcome can only be one of the eigenvalues of the operator.

For this reasoning to be plausible, the eigenvectors must form a complete set of vectors that span any general vector within the space, and can thereby account for all

possible experimental outcomes.

The *collapse of the wavefunction* led to many disputes in the development of quantum mechanics.

The state vector or wavefunction evolves deterministically according to Schrödinger's equation, but then collapses when a measurement is made. The collapse is not, however, described by a differential equation.

This is a central feature of the Copenhagen interpretation of quantum mechanics.

It is often said that uncertainty arises as a consequence of the measurement interfering with the system. But if this is true, why should the second and subsequent measurements, carried out immediately after the first, return the same value with certainty? Surely the uncertainty should be the same each time? Or is this about knowledge, not measurement? In addition, if a measurement simply interferes with the experiment in a classical way, why do we need quantum mechanics at all; surely, classical physics would suffice?

The collapse of the wavefunction, which is an accepted part of the Copenhagen interpretation, remains a contentious topic. Various interpretations exist: (i) collapse is due to the measurement apparatus interfering with the quantum state in an irreversible way; (ii) collapse is taking place continuously, even when measurements are not being performed; (iii) nonlinear terms are missing in Schrödinger's equation, which would describe the collapse.

Remember, quantum mechanics still remains the best model we have for describing the many counterintuitive behaviours seen in experiments, which cannot be described classically, regardless of the conceptual interpretation of collapse!

38 Orthogonality of eigenstates

Suppose that $|a_m\rangle$ and $|a_n\rangle$ are two eigenstates of a Hermitian operator \hat{A} having corresponding eigenvalues a_m and a_n , respectively:

$$\hat{A}|a_m\rangle = a_m|a_m\rangle \quad \text{and} \quad \hat{A}|a_n\rangle = a_n|a_n\rangle. \quad (345)$$

Then

$$\langle a_m|\hat{A}|a_n\rangle = a_n\langle a_m|a_n\rangle. \quad (346)$$

Since \hat{A} is Hermitian

$$\begin{aligned} \langle a_m|\hat{A}|a_n\rangle &= \langle a_n|\hat{A}^\dagger|a_m\rangle^* \\ &= \langle a_n|\hat{A}|a_m\rangle^* \\ &= a_m^*\langle a_n|a_m\rangle^* \\ &= a_m^*\langle a_m|a_n\rangle. \end{aligned} \quad (347)$$

Subtracting gives

$$(a_m^* - a_n)\langle a_m|a_n\rangle = 0. \quad (348)$$

Setting $m = n$, we see that $a_m^* = a_m$ because $\langle a_m|a_m\rangle \neq 0$. Confirming that the eigenvalues of a Hermitian operator are purely real.

Eq. (348) thus becomes

$$(a_m - a_n)\langle a_m|a_n\rangle = 0. \quad (349)$$

Now suppose that $a_m \neq a_n$, then it must follow that

$$\langle a_m|a_n\rangle = 0 \quad \text{for } m \neq n. \quad (350)$$

The eigenstates of a Hermitian operator corresponding to different eigenvalues are orthogonal.

39 Degeneracy of eigenstates

Two distinct eigenstates $|a_m\rangle$ and $|a_n\rangle$ of a Hermitian operator \hat{A} are called *degenerate* if their eigenvalues are the same: $a_m = a_n = a$.

It means that when an observable is measured two different states can yield the same experimental result.

This situation occurs frequently in practice, for example two different quantum states can have the same energy.

It is always possible, however, to create a pair of orthogonal states from any pair of (linearly independent) states having the same eigenvalue.

40 Completeness of eigenstates

The eigenvectors of an operator form a *complete* set with respect to the vector space under consideration, and can be conveniently used as a basis set.

For any vector $|\psi\rangle$ in the space, we can then write

$$|\psi\rangle = \sum_n c_n |a_n\rangle, \quad (351)$$

where the c_n are general expansion coefficients.

Note the strict equality sign, because the basis set is complete.

The completeness of the eigenvectors allows the expectation value of an operator to be found easily once the eigenvalues are known.

For any state $|\psi\rangle$, the expectation value of \hat{A} is

$$\begin{aligned}\langle A \rangle &= \langle \psi | \hat{A} | \psi \rangle \\ &= \sum_n c_n \langle \psi | \hat{A} | a_n \rangle \\ &= \sum_n c_n a_n \langle \psi | a_n \rangle \\ &= \sum_n a_n |c_n|^2.\end{aligned}\tag{352}$$

We conclude that

$$\boxed{\langle A \rangle = \sum_n a_n |c_n|^2.}\tag{353}$$

If the weighting coefficients are known, the expectation value of an operator can be written down easily.

For any general state, the expectation value is the sum over all possible results of the measurement weighted by the square of the expansion coefficients. If the system is in an eigenstate, say the N 'th eigenstate, then $\langle A \rangle = a_N$, as required.

It is apparent from Eq. (353) that the probability of measuring the eigenvalue a_n is

$$\boxed{P_n = |c_n|^2 = |\langle a_n | \psi \rangle|^2.}\tag{354}$$

We can take any general vector $|\psi\rangle$ and then project it onto the eigenstates, which form a basis, to get a set of complex coefficients. The square-magnitude of each of these coefficients gives the probability that the associated eigenvalue will be measured.

Every abstract vector space has a zero vector defined as:

$\psi(x) = 0, \forall x$	Zero function
$\mathbf{a} = \mathbf{0}$	Zero vector
$\mathbf{A} = 0$	Zero matrix
$ a\rangle = 0\rangle$	Zero ket.

The norm of the zero vector is the zero scalar. The projection of any state vector onto the zero vector is zero. In quantum mechanics, what is the physical interpretation of the zero state vector?

Previously we saw that a wave function can be represented by

$$\phi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \psi(x) e^{-ipx/\hbar} dx, \quad (355)$$

which is nothing more than a weighted linear combination of basis functions.

Eq. (355) is called the momentum representation.

In this representation, the position operator is $\hat{x} = i\hbar\partial/\partial p$, and the basis functions are the eigenfunctions of this operator.

Eq. (352) then gives

$$\begin{aligned} \langle x \rangle &= \int_{-\infty}^{+\infty} \phi^*(p) \hat{x} \phi(p) dp \\ &= \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dx' \int_{-\infty}^{+\infty} dp x \psi(x) \psi^*(x') e^{-ip(x-x')/\hbar}. \end{aligned} \quad (356)$$

Consider the integral,

$$\frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} e^{-ip(x-x')/\hbar} dp = \delta(x - x'), \quad (357)$$

which is nothing more than the orthogonality condition for the basis functions.

Then

$$\begin{aligned}\langle x \rangle &= \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dx' x\psi(x)\psi^*(x')\delta(x-x') \\ &= \int_{-\infty}^{+\infty} x|\psi(x)|^2dx,\end{aligned}\tag{358}$$

and $|\psi(x)|^2dx$ is the probability of finding the particle between x and $x + dx$, which agrees with Eqs. (353) and (354).

41 Compatible and incompatible observables

Under what circumstances is it possible to measure two observables ‘simultaneously’ and precisely?

Consider two commuting observables \hat{A} and \hat{B} :

$$[\hat{A}, \hat{B}] = \hat{0}. \tag{359}$$

First assume that none of the eigenstates $|a_n\rangle$ of \hat{A} are degenerate. In other words, all of the eigenvalues are different.

Notice then that

$$\begin{aligned}\hat{A}\hat{B}|a_n\rangle &= \hat{B}\hat{A}|a_n\rangle \\ &= a_n\hat{B}|a_n\rangle \\ \hat{A}(\hat{B}|a_n\rangle) &= a_n(\hat{B}|a_n\rangle),\end{aligned}\tag{360}$$

and $\hat{B}|a_n\rangle$ is also an eigenstate of \hat{A} with eigenvalue a_n .

Since there is no degeneracy

$$\hat{B}|a_n\rangle \propto |a_n\rangle \quad \text{i.e.,} \quad \hat{B}|a_n\rangle = b_n|a_n\rangle, \tag{361}$$

which shows that $|a_n\rangle$ is also an eigenstate of \hat{B} , albeit with a possibly different eigenvalue.

Even in the degenerate case, as long as an orthonormal set of eigenvectors has been constructed through Gram-Schmidt orthogonalisation, the same reasoning applies.

We conclude that if \hat{A} and \hat{B} commute, they share common eigenvectors, and it does not matter in which order the operators are applied.

If the system is in an eigenstate, precise values are measured for each of the operators, and it does not matter in which order the measurements are made.

Because \hat{A} and \hat{B} can be measured in any order, simultaneously, they are said to be *compatible*.

Observables are *compatible* if their operators commute.

For example, the wavefunctions

$$e^{ikx} \text{ and } e^{-ikx}, \quad (362)$$

are eigenstates of the momentum operator \hat{p} , having eigenvalues $\pm\hbar k$, and of the free-space Hamiltonian \hat{H}_0 , having eigenvalues $\hbar^2 k^2 / 2m$.

This is possible because the two operators commute:

$$[\hat{H}, \hat{p}] = \hat{0}. \quad (363)$$

The measurement of the second quantity is not strictly needed because its value can be calculated once the first measurement has been made.

Observables are said to be *incompatible* if their corresponding operators do not commute:

$$[\hat{A}, \hat{B}] \neq \hat{0}. \quad (364)$$

For example, \hat{x} and \hat{p} do not commute, and in fact

$$[\hat{x}, \hat{p}] = i\hbar. \quad (365)$$

They cannot be measured simultaneously with complete accuracy, and ‘simultaneous’ measurements are subjected to the inequality

$$\Delta x \Delta p \geq \frac{\hbar}{2}. \quad (366)$$

42 State vectors

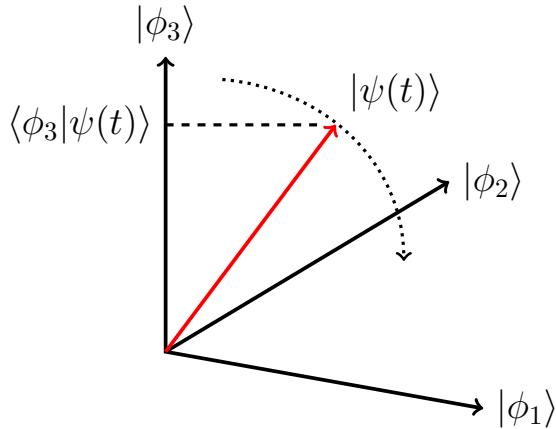


Figure 48: A general state vector $|\psi(t)\rangle$ in terms of the eigenstates $|\phi_n\rangle$ of some operator. The state vector moves around the vector space as time progresses, but then jumps to one of the eigenstates, say $|\phi_3\rangle$, when a measurement is made. The probability of the jump is given by $|\langle\psi(t)|\phi_3\rangle|^2$. The recorded value is the eigenvalue, say a_3 .

The overall picture is conveniently summarised in Fig. 48:

- The quantum state of a system at a point in time is described by a vector $|\psi\rangle$ in a finite or infinite-dimensional space.

- The eigenvectors of an operator $|a_n\rangle$ form a complete set, and so the direction of $|\psi\rangle$ can be described in terms of the coordinates, or inner products with the $|a_n\rangle$ eigenvectors.
- If left on its own, the state vector moves around the space in time according to Schrodinger's wave equation, or whatever other differential equation describes the time evolution.
- The eigenvectors of the operator do not change with time, and so the coefficients of the expansion must be functions of time.
- When a measurement of A is performed, the state vector collapses onto one of the eigenvectors of \hat{A} .
- The state vector preferentially collapses in the direction of the largest projection, because the probability of the outcome is greatest. This is the action of the projection operator.
- If A is measured again, immediately, the same result is found.
- If a measurement of a new variable B is performed, and if \hat{B} commutes with \hat{A} , the result is certain because the eigenvectors of \hat{A} and \hat{B} are the same: the axes point in the same direction. In fact B could have been calculated directly once A is known. These are compatible observables.
- If a measurement of C is made, and if \hat{C} does not commute with \hat{A} , the axes lie in a different direction, and a new projected collapse takes place.
- If the system is left without any further measurements taking place, the system evolves and the state vector can again start to move around the space.
- In some cases, the state vector does not move away from the collapsed eigenvector (e.g., if the latter happens to be an eigenvector of the Hamiltonian operator).

43 Position and momentum wavefunctions revisited

It is now possible to ‘visualise’ the position and momentum wavefunctions in a different way.

Suppose that we have some state, say an energy eigenstate $|E\rangle$.

There is also a set of states, let us call them position states, $|\mathbf{r}\rangle$, that correspond to finding the particle at specific locations. These are the eigenstates of the position operator.

When a measurement of position is made, the state vector collapses onto one of the position eigenstates, and its position is known precisely.

The probability of position \mathbf{r} being recorded is $|\langle \mathbf{r}|E\rangle|^2$, but this must be equal to $|\psi(\mathbf{r})|^2$.

The position wave function $\psi(\mathbf{r})$ is nothing more than the complex valued projection $\langle \mathbf{r}|E\rangle$.

Alternatively, we may measure the momentum of the particle, and then the state vector will collapse onto one of the momentum eigenstates.

The probability of momentum \mathbf{p} being recorded is $|\langle \mathbf{p}|E\rangle|^2$, but this must be equal to $|\phi(\mathbf{p})|^2$.

The momentum wave function $\phi(\mathbf{p})$ is nothing more than the complex valued projection $\langle \mathbf{p}|E\rangle$.

Because position and momentum cannot be measured simultaneously, their axes are effectively non-aligned in our multi-dimensional state vector space: Fig. 49.

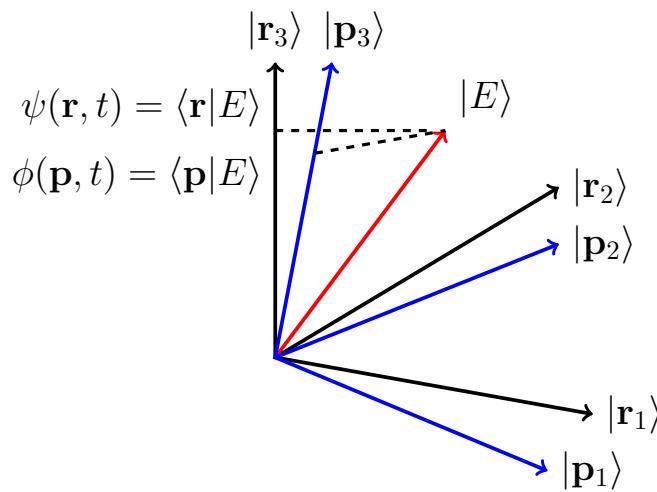


Figure 49: A general state vector $|E\rangle$ (red) in terms of the position eigenstates (black) gives the complex valued position wavefunction. The same state vector $|E\rangle$ in terms of the momentum eigenstates (blue) gives the complex valued momentum wavefunction. The position and momentum wavefunctions are different representations of the same state.

The following terminology is commonly used when referring to wave functions:

1. $\delta(\mathbf{r} - \mathbf{r}')$ is the position state vector in the position representation, which is continuous. \mathbf{r}' is the parameter that labels the particular state, and \mathbf{r} is the variable that spans the vector space. Denote this state $|\mathbf{r}'\rangle$.
2. $1/\sqrt{V} \exp[i\mathbf{k}' \cdot \mathbf{r}]$ is the momentum state vector in the position representation. \mathbf{k}' is the parameter that labels the state, and \mathbf{r} is the vector space. Denote this state $|\mathbf{k}'\rangle$.
3. $1/\sqrt{V} \exp[-i\mathbf{k} \cdot \mathbf{r}']$ is the position state vector in the momentum representation. \mathbf{r}' is the parameter that labels the state, and \mathbf{k} is the vector space. Denote this state $|\mathbf{r}'\rangle$.
4. $\delta_{\mathbf{k}\mathbf{k}'}$ is the momentum state vector in the momentum representation, which for a bound particle is discrete. \mathbf{k}' is the parameter that labels the state, and \mathbf{k} is the vector space. Denote this state $|\mathbf{k}'\rangle$.

(1) and (3) are the same vector expressed in different coordinate systems, and (2) and (4) are the same vector expressed in different coordinate systems.

Calculate each of $\langle \mathbf{r}' | \mathbf{r}' \rangle$, $\langle \mathbf{k}' | \mathbf{k}' \rangle$, $\langle \mathbf{k}' | \mathbf{r}' \rangle$ in the space and momentum domains and comment on the results.

From a vector-space perspective, explain how the above terminology relates to the function it represents.

44 Generalised uncertainty relationships

Although we have evaluated specific commutation relationships, and placed a limit on the simultaneous measurement of position and momentum, we can generalise the Uncertainty Principle in a formal way to any non-commuting (incompatible) observables.

Define the operator \hat{A}_d , where \hat{A}_d^2 measures the deviation of A from its mean value $\bar{A} \equiv \langle A \rangle$:

$$\begin{aligned}\hat{A}_d &= \hat{A} - \bar{A} \\ \hat{A}_d^2 &= (\hat{A} - \bar{A})(\hat{A} - \bar{A}) \\ &= \hat{A}^2 - 2\bar{A}\hat{A} + \bar{A}^2,\end{aligned}\tag{367}$$

which gives

$$\begin{aligned}\langle \hat{A}_d^2 \rangle &= \langle \hat{A}^2 \rangle - \bar{A}^2 \\ &= (\Delta A)^2.\end{aligned}\tag{368}$$

The expectation value of \hat{A}_d^2 is the squared uncertainty $(\Delta A)^2$.

There is a similar expression for $\langle \hat{B}_d^2 \rangle$.

For any state $|\psi\rangle$ consider the state vector

$$|\phi\rangle = (\hat{A}_d + i\lambda\hat{B}_d)|\psi\rangle,\tag{369}$$

where λ is an arbitrary real number.

The functional form of Eq. (369) may seem to be picked out of the blue, but it recurs frequently in quantum mechanics.

We require the norm of $|\phi\rangle$ to be positive (as it must

generally be the case)

$$\begin{aligned}\langle\phi|\phi\rangle &= \langle\psi|\left(\hat{A}_d - i\lambda\hat{B}_d\right)\left(\hat{A}_d + i\lambda\hat{B}_d\right)|\psi\rangle \quad (370) \\ &= (\Delta A)^2 + \lambda^2(\Delta B)^2 + \lambda\langle i\left[\hat{A}_d, \hat{B}_d\right]\rangle \\ &\geq 0,\end{aligned}$$

where we have used the shorthand notation

$$\langle i\left[\hat{A}_d, \hat{B}_d\right]\rangle = \langle\psi|i\left[\hat{A}_d, \hat{B}_d\right]|\psi\rangle. \quad (371)$$

\hat{A} , \hat{A}^2 , \hat{A}_d , \hat{B} , \hat{B}^2 , \hat{B}_d are all Hermitian, and so $(\Delta A)^2$ and $(\Delta B)^2$ are positive real.

$i\left[\hat{A}_d, \hat{B}_d\right]$ is also Hermitian, and so $\langle i\left[\hat{A}_d, \hat{B}_d\right]\rangle$ is real.

$\langle\phi|\phi\rangle$ is at a minimum, as we vary λ , when

$$2\lambda(\Delta B)^2 + \langle i\left[\hat{A}_d, \hat{B}_d\right]\rangle = 0 \quad (372)$$

or

$$\lambda = -\frac{\langle i\left[\hat{A}_d, \hat{B}_d\right]\rangle}{2(\Delta B)^2}. \quad (373)$$

Substitution gives

$$\begin{aligned}\langle\phi|\phi\rangle_{\min} &= (\Delta A)^2 + \frac{\langle i\left[\hat{A}_d, \hat{B}_d\right]\rangle^2}{4(\Delta B)^2} - \frac{\langle i\left[\hat{A}_d, \hat{B}_d\right]\rangle^2}{2(\Delta B)^2}, \quad (374) \\ &= (\Delta A)^2 + \frac{(\Delta B)^2}{4} - \frac{(\Delta B)^2}{2},\end{aligned}$$

which we require to be non-negative. That is,

$$(\Delta A)^2(\Delta B)^2 \geq \frac{1}{4}\langle i\left[\hat{A}_d, \hat{B}_d\right]\rangle^2. \quad (375)$$

Now $\langle A \rangle$ and $\langle B \rangle$ are just numbers, and so

$$\begin{aligned}\left[\hat{A}_d, \hat{B}_d\right] &= \hat{A}_d\hat{B}_d - \hat{B}_d\hat{A}_d \quad (376) \\ &= (\hat{A} - \langle A \rangle)(\hat{B} - \langle B \rangle) - (\hat{B} - \langle B \rangle)(\hat{A} - \langle A \rangle) \\ &= [\hat{A}, \hat{B}].\end{aligned}$$

We conclude that

$$\boxed{\Delta A \Delta B \geq \frac{1}{2} \left| \langle i [\hat{A}, \hat{B}] \rangle \right|,} \quad (377)$$

and there is a limit to the accuracy with which *any* pair of non-commuting observables can be measured simultaneously.

The limit is known once the commutator is known.

For example, $[\hat{x}, \hat{p}] = i\hbar$ and so $\Delta x \Delta p \geq \hbar/2$.

45 Minimum uncertainty states

For the state $|\psi\rangle$ to yield the minimum product of uncertainties, $\Delta A \Delta B$ is a minimum, $\langle \phi | \phi \rangle$ must vanish:

$$|\phi\rangle = (\hat{A}_d + i\lambda \hat{B}_d) |\psi\rangle = 0, \quad (378)$$

which leads to a differential equation in the case of wavefunctions.

For example, the minimum uncertainty product $\Delta x \Delta p = \hbar/2$, requires

$$(\hat{x}_d + i\lambda \hat{p}_d) |\psi\rangle = 0 \quad (379)$$

$$\begin{aligned} & \left[(x - \langle x \rangle) + i\lambda \left(\frac{\hbar}{i} \frac{\partial}{\partial x} - \langle p \rangle \right) \right] \psi(x) = 0 \\ & \frac{\partial \psi(x)}{\partial x} = \left[-\frac{(x - \langle x \rangle)}{\lambda \hbar} + \frac{i \langle p \rangle}{\hbar} \right] \psi(x). \end{aligned}$$

The solution to this differential equation is the familiar Gaussian wavepacket

$$\boxed{\psi(x) = e^{-(x-\langle x \rangle)^2/2\hbar\lambda} e^{i\langle p \rangle x/\hbar}.} \quad (380)$$

which leads to

$$\Delta x = \sqrt{\frac{\hbar\lambda}{2}} \text{ and } \Delta p = \sqrt{\frac{\hbar}{2\lambda}}. \quad (381)$$

Increasing λ increases the uncertainty in x and decreases the uncertainty in p , while keeping the product of the uncertainties constant at the minimum value $\hbar/2$.

Eq. (380) is a *minimum uncertainty state*.

46 A footnote

In this section, we have introduced terms such as ‘Hamiltonian’ and ‘commutator’, and these may seem peculiar. In fact, quantum mechanics, which *is* a mechanics, because it describes the dynamical behaviour of all systems, can be built using the techniques of classical Lagrangian mechanics.

The Lagrangian leads to a formal definition of the Hamiltonian.

The Lagrangian can also be used to find conjugate variables, such as position and momentum, which leads to non-commuting dynamical variables in quantum mechanics.

The Lagrangian also leads to the notion of the Poisson bracket, which plays a similar role in classical mechanics as the commutator does in quantum mechanics.

Formally, a system is best quantised by first writing down the classical Lagrangian, and building the quantum model from there. This is the approach taken when calculating the behaviour of relativistic systems.

We shall not pursue these topics in this course.

47 Degeneracy of eigenstates - not examinable

Consider two distinct eigenstates $|a_m\rangle$ and $|a_n\rangle$ of a Hermitian operator \hat{A} that are *degenerate*. Namely, their eigenvalues are the same: $a_m = a_n = a$.

It is always possible to create a pair of orthogonal states from any pair of states having the same eigenvalues.

Create a new state vector by forming a weighted linear combination of these two states:

$$|\phi\rangle = \alpha_m|a_m\rangle + \alpha_n|a_n\rangle, \quad (382)$$

where α_m and α_n are some general complex coefficients.

Form the inner product with $|a_m\rangle$

$$\langle a_m|\phi\rangle = \alpha_m\langle a_m|a_m\rangle + \alpha_n\langle a_m|a_n\rangle. \quad (383)$$

This result will be zero if we ensure that the coefficients obey the relationship

$$\frac{\alpha_m}{\alpha_n} = -\frac{\langle a_m|a_n\rangle}{\langle a_m|a_m\rangle}. \quad (384)$$

In which case we have created a new vector $|\phi\rangle$ that is orthogonal to $|a_m\rangle$:

$$|\phi\rangle = \alpha_n \left(-\frac{\langle a_m|a_n\rangle}{\langle a_m|a_m\rangle} |a_m\rangle + |a_n\rangle \right), \quad (385)$$

and the value of α_n can be chosen to normalise $|\phi\rangle$.

This procedure is known as *Gram-Schmidt orthonormalisation* and can be carried out even when a large number of eigenvectors have the same eigenvalue.

It is easy to verify that $|\phi\rangle$ is also an eigenvector of \hat{A} having the eigenvalue a .

Consider two non-orthogonal Cartesian vectors in a plane. By sketching a diagram give a geometrical meaning to Eq. (385).

VII Advanced Operators

In the preceding chapter, we introduced a description of quantum mechanics based on vector spaces and operators. We further revisited the notion of position, momentum and energy operators in the new language, and we discussed the concepts of eigenvalues and eigenvectors, of expectation values and the uncertainty principle. Operator algebra is central to quantum mechanics. Its advanced forms pervade many concepts that are now regarded as ‘physical reality’. We have already seen how one can construct operators, albeit artificially, that step through sequences of states, sometimes stopping naturally when the operator is applied multiple times. In this chapter, we will demonstrate the substantial importance that this type of operators have in the physical world by considering again the simple harmonic oscillator. We shall review it firstly at the classical level, and then we will study it at the quantum mechanical level in the operator formalism – to be contrasted with the Schrödinger description in earlier chapters. The quantisation of the harmonic oscillator is central to quantum mechanics because it highlights the relationship between particle-like and wave-like behaviour. Ultimately, it defines what one means by elementary excitations such as photons, phonons, magnons, etc., and it provides an elegant scheme for describing their behaviour and interactions. We will close the chapter by introducing the notion of density operator and by discussing how to form functions of operators.

48 Classical analysis of the harmonic oscillator

Let us revisit the analysis of a classical harmonic oscillator.

The behaviour is described in terms of a pair of coupled equations in the position x and momentum p :

$$\begin{aligned}\frac{dx}{dt} &= \frac{p}{m} \\ \frac{dp}{dt} &= -\alpha x = -\omega^2 mx,\end{aligned}\tag{386}$$

where we have expressed the ‘spring constant’ α in terms of the natural frequency of oscillation ω ; namely, $\omega^2 = \alpha/m$, where m is the mass.

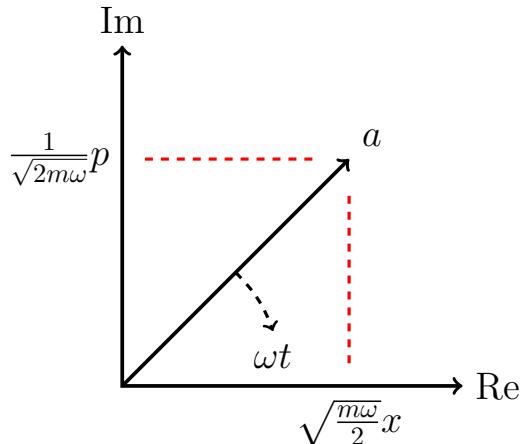


Figure 50: Complex phase-space representation of the mode amplitude. The real part is proportional to position. The imaginary part is proportional to momentum. The complex quantity a rotates in the complex plane as time progresses.

A solution can be found by differentiating the first equation again and substituting into the second equation:

$$\frac{d^2x}{dt^2} = -\omega^2 x,\tag{387}$$

giving a single second-order differential equation that must be solved.

Alternatively, Eq. (386) can be solved by defining

$$\begin{aligned} a &= \sqrt{\frac{m\omega}{2}}x + i\frac{1}{\sqrt{2m\omega}}p \\ a^* &= \sqrt{\frac{m\omega}{2}}x - i\frac{1}{\sqrt{2m\omega}}p, \end{aligned} \quad (388)$$

which is a vector in the complex plane, illustrated in Fig. 50.

Then

$$\begin{aligned} x &= \frac{1}{\sqrt{2m\omega}}(a^* + a) = \sqrt{\frac{2}{m\omega}} \operatorname{Re}[a] \\ p &= i\sqrt{\frac{m\omega}{2}}(a^* - a) = \sqrt{2m\omega} \operatorname{Im}[a]. \end{aligned} \quad (389)$$

Eqs. (386) and (389) combine to give the single, complex, equation

$$\frac{da}{dt} = -i\omega a. \quad (390)$$

The solution is

$$\begin{aligned} a(t) &= a(0)e^{-i\omega t} \\ &= \left[\sqrt{\frac{m\omega}{2}}x(0) + i\frac{1}{\sqrt{2m\omega}}p(0) \right] e^{-i\omega t}. \end{aligned} \quad (391)$$

The modulus $|a(t)|^2$ is independent of time, and becomes

$$|a(t)|^2 = \frac{m\omega}{2}x(0)^2 + \frac{1}{2m\omega}p(0)^2, \quad (392)$$

and therefore

$$\omega aa^* = \frac{m\omega^2}{2}x^2 + \frac{1}{2m}p^2 = E, \quad (393)$$

which is the sum of the potential and kinetic energies.

$a(t)$, which is a complex number, is called the *amplitude of the normal mode*.

It rotates in the complex plane at the angular rate ω such that its projection onto the real axis essentially gives $x(t)$, and its projection onto the imaginary axis essentially gives $p(t)$.

It is a *phase-space* representation of the SHM on the complex plane: Fig. 50.

The length of the vector remains constant, and is proportional to the total energy E .

The energy of the system, or the length of the vector, can be increased by any continuous amount, with an associated increase in the amplitudes of x and p .

49 Quantum analysis of the harmonic oscillator

It is possible to quantise a simple harmonic oscillator using the position and momentum operators.

The Hamiltonian for a harmonic oscillator is

$$\hat{H} = \frac{1}{2m}\hat{p}^2 + \frac{m\omega^2}{2}\hat{x}^2. \quad (394)$$

We have gone from the classical expression for the total energy to the quantum expression by swapping measurable variables by operators.

Crucially, in this case, there is no ambiguity with regards to the order of the operators, which are not constrained classically (namely, all products of observables – x^2 and p^2 – involve at the quantum level operators that trivially commute – $\hat{x}\hat{x} = \hat{x}\hat{x}$ and $\hat{p}\hat{p} = \hat{p}\hat{p}$).

In correspondence with Eq. (388), we introduce two new

operators

$$\begin{aligned}\hat{a} &= \sqrt{\frac{m\omega}{2\hbar}}\hat{x} + i\frac{1}{\sqrt{2m\hbar\omega}}\hat{p} \\ \hat{a}^\dagger &= \sqrt{\frac{m\omega}{2\hbar}}\hat{x} - i\frac{1}{\sqrt{2m\hbar\omega}}\hat{p}.\end{aligned}\quad (395)$$

This functional form has already been used in Handout VI.

$\hat{a}^\dagger \neq \hat{a}$ is not Hermitian, and so does not correspond to an observable.

Nevertheless, \hat{a} and \hat{a}^\dagger have remarkable properties.

$$\begin{aligned}\hat{a}\hat{a}^\dagger &= \frac{m\omega}{2\hbar}\hat{x}^2 + \frac{1}{2m\hbar\omega}\hat{p}^2 - \frac{i}{2\hbar}[\hat{x}, \hat{p}] \\ \hat{a}^\dagger\hat{a} &= \frac{m\omega}{2\hbar}\hat{x}^2 + \frac{1}{2m\hbar\omega}\hat{p}^2 + \frac{i}{2\hbar}[\hat{x}, \hat{p}],\end{aligned}\quad (396)$$

and therefore

$$\begin{aligned}\hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a} &= \frac{1}{i\hbar}[\hat{x}, \hat{p}] = 1 \\ \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a} &= \frac{2}{\hbar\omega}\hat{H}.\end{aligned}\quad (397)$$

The first equation can be conveniently written as

$$[\hat{a}, \hat{a}^\dagger] = 1, \quad (398)$$

and

$$\begin{aligned}\hat{H} &= \frac{\hbar\omega}{2}(\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}) \\ &= \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) \\ &= \hbar\omega\left(\hat{a}\hat{a}^\dagger - \frac{1}{2}\right).\end{aligned}\quad (399)$$

The most commonly used form is:

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

(400)

Suppose that $|\phi\rangle$ is an eigenstate of \hat{H} with eigenvalue E :

$$\hat{H}|\phi\rangle = E|\phi\rangle. \quad (401)$$

Now consider

$$\begin{aligned}\hat{H}\hat{a}|\phi\rangle &= \hbar\omega \left(\hat{a}\hat{a}^\dagger - \frac{1}{2} \right) \hat{a}|\phi\rangle \quad (402) \\ &= \hbar\omega \left(\hat{a}\hat{a}^\dagger\hat{a} - \frac{1}{2}\hat{a} \right) |\phi\rangle \\ &= \hbar\omega\hat{a} \left(\hat{a}^\dagger\hat{a} - \frac{1}{2} \right) |\phi\rangle \\ &= \hat{a} \left(\hat{H} - \hbar\omega \right) |\phi\rangle \\ &= (E - \hbar\omega) \hat{a}|\phi\rangle.\end{aligned}$$

Thus

$$\boxed{\hat{H}\hat{a}|\phi\rangle = (E - \hbar\omega) \hat{a}|\phi\rangle.} \quad (403)$$

Likewise, it can also be shown that

$$\boxed{\hat{H}\hat{a}^\dagger|\phi\rangle = (E + \hbar\omega) \hat{a}^\dagger|\phi\rangle.} \quad (404)$$

- If $|\phi\rangle$ is an eigenstate of \hat{H} with eigenvalue E , then $\hat{a}|\phi\rangle$ is an eigenstate of \hat{H} with eigenvalue $E - \hbar\omega$.
- If $|\phi\rangle$ is an eigenstate of \hat{H} with eigenvalue E , then $\hat{a}^\dagger|\phi\rangle$ is an eigenstate of \hat{H} with eigenvalue $E + \hbar\omega$.

\hat{a}^\dagger and \hat{a} are called *creation* and *annihilation* operators respectively because they increase, decrease the energy of the system by $\hbar\omega$ (a quantum of energy).

They are also called *raising* and *lowering* operators, or collectively *ladder* operators.

They lead to a different formulation of quantum mechanics called *second quantisation*.

Creation and annihilation operators appear in the quantisation of the electromagnetic field, and are central to the subject of quantum optics.

The creation operator adds and the annihilation operator removes a photon from an electromagnetic field.

They are intimately related to the way in which an electron in an atom emits a photon when it changes state: the electron's energy falls by $\hbar\omega$ and the electromagnetic energy increases by $\hbar\omega$.

Although \hat{a}^\dagger and \hat{a} are not Hermitian, and therefore not observable, $\hat{a}^\dagger\hat{a}$ clearly is Hermitian, and observable.

It is called the *number operator*

$$\hat{N} = \hat{a}^\dagger\hat{a}, \quad (405)$$

and then (400) becomes

$$\hat{H} = \hbar\omega \left(\hat{N} + \frac{1}{2} \right). \quad (406)$$

The number operator corresponds to measuring the number of elementary quantum excitations, photons or phonons.

One can show that its eigenvalues are integers, $\hat{N}|\phi_n\rangle = n|\phi_n\rangle$, and that $n \geq 0, \forall n$. (Indeed, $\langle\phi|a^\dagger a|\phi\rangle \geq 0$ because it is the norm squared of the vector $a|\phi\rangle$; and the lower bound is met when $a|\phi\rangle = 0$, giving the ground state of the system). Therefore the minimum possible energy is $\hbar\omega/2$, which is called the *ground-state energy* or *vacuum energy*.

It is called the ‘vacuum energy’ because it is present even when there are no photons! ($n = 0$ in $|\phi_0\rangle$)

Creation and annihilation operators raise, or lower, a state by one quantum,

$$\begin{aligned}\hat{a}^\dagger |\phi_n\rangle &\propto |\phi_{n+1}\rangle \\ \hat{a} |\phi_n\rangle &\propto |\phi_{n-1}\rangle,\end{aligned}\tag{407}$$

and therefore the energy is increased, decreased by $\hbar\omega$. We shall find out the correct proportionality constants in the above equations shortly.

The number operator, Eq. (405), consists of removing a photon and then putting it back. This may seem peculiar, but the number of unique times particles can be removed and put back is simply the number of particles!

This formalism describes *particle counting* in a specific way.

Why does the operator $\hat{a}\hat{a}^\dagger$ not count particles?

Suppose that the eigenvector corresponding to the ground state is $|\phi_0\rangle$; then there is no state below it.

Eq. (403) shows that if we keep on applying the lowering operator to some general state, the energy keeps on being reduced in units of $\hbar\omega$, but the energy cannot go negative, and so the lowering sequence must terminate when $E = \hbar\omega/2$.

If we call the ground state $|\phi_0\rangle$, then

$$\hat{a}|\phi_0\rangle = |0\rangle\tag{408}$$

is a suitable way of terminating the process.

$|0\rangle$ is the zero vector in the space, and it has a norm

of zero. Every vector space has a zero vector.

This is a suitable termination because in the ground state

$$\begin{aligned}\langle \hat{H} \rangle &= \langle \phi_0 | \hat{H} | \phi_0 \rangle \\ &= \hbar\omega \langle \phi_0 | \hat{a}^\dagger \hat{a} | \phi_0 \rangle + \frac{1}{2}\hbar\omega \\ &= \frac{1}{2}\hbar\omega\end{aligned}\tag{409}$$

if (408) is satisfied.

We can generate the first excited state through

$$|\phi_1\rangle = \hat{a}^\dagger |\phi_0\rangle,\tag{410}$$

which according to Eq. (404) is associated with an eigenvalue of $\hbar\omega + \hbar\omega/2$. It is easy to show (using for instance the commutation relation) that $|\phi_1\rangle$ is properly normalised.

It follows from (406) that the expectation value of \hat{N} when the system is in this state is 1.

Continuing,

$$|\phi_n\rangle \propto (\hat{a}^\dagger)^n |\phi_0\rangle,\tag{411}$$

which is associated with an eigenvalue of $n\hbar\omega + \hbar\omega/2$. The introduction of a proportionality constant is however required in this case to ensure that the state is properly normalised, as we shall discuss further below.

The expectation value of \hat{N} , when the system is in eigenstate $|\phi_n\rangle$ is

$$\begin{aligned}\langle N \rangle &= \langle \phi_n | \hat{N} | \phi_n \rangle \\ &= \frac{1}{\hbar\omega} \langle \phi_n | \hat{H} | \phi_n \rangle - \frac{1}{2} \\ &= \frac{E}{\hbar\omega} - \frac{1}{2} \\ E &= \langle N \rangle \hbar\omega + \frac{\hbar\omega}{2}.\end{aligned}\tag{412}$$

The scheme is summarised in Fig. 51.

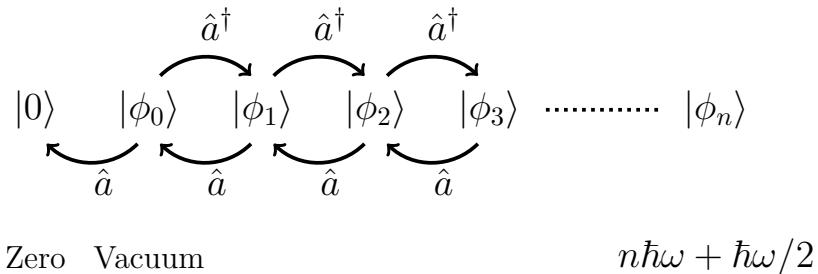


Figure 51: The creation operator \hat{a}^\dagger increases the excitation state by a quantum of energy. The annihilation operator \hat{a} decreases the excitation state by a quantum of energy. The series terminates at the bottom by generating the zero vector $|0\rangle$.

If $|\phi\rangle$ is an eigenstate of \hat{H} with eigenvalue E , then $|\phi\rangle$ is also an eigenstate of \hat{N} with eigenvalue $n = E/\hbar\omega - 1/2$.

Show that \hat{H} and \hat{N} commute, and therefore if the energy is known, the number of photons is known.

In general

$$\hat{H}|\phi_n\rangle = \left(n + \frac{1}{2}\right)\hbar\omega|\phi_n\rangle \quad (413)$$

where the energy of the state is $E_n = (n + 1/2)\hbar\omega$.

The scheme is nearly complete, but in generating high-order states from low-order states, we need to ensure that the states are normalised.

$\hat{a}|\phi_n\rangle$ is an eigenket of \hat{N} having eigenvalue $n - 1$.

It can only differ from $|\phi_{n-1}\rangle$ by some factor:

$$\hat{a}|\phi_n\rangle = c_n|\phi_{n-1}\rangle. \quad (414)$$

The norm of this vector is

$$\begin{aligned} \langle\phi_n|\hat{a}^\dagger\hat{a}|\phi_n\rangle &= n\langle\phi_n|\phi_n\rangle \\ &= |c_n|^2\langle\phi_{n-1}|\phi_{n-1}\rangle. \end{aligned} \quad (415)$$

If $\langle \phi_{n-1} | \phi_{n-1} \rangle = 1$ and we choose $c_n = \sqrt{n}$, then $\langle \phi_n | \phi_n \rangle = 1$, and

$$\hat{a}|\phi_n\rangle = \sqrt{n}|\phi_{n-1}\rangle. \quad (416)$$

Likewise, one can show that

$$\hat{a}^\dagger|\phi_n\rangle = \sqrt{n+1}|\phi_{n+1}\rangle. \quad (417)$$

If $\langle \phi_0 | \phi_0 \rangle = 1$ then we can generate a sequence of normalised states.

Collecting these results:

$\hat{N} \phi_n\rangle$	$=$	$n \phi_n\rangle$	(418)
$\hat{a} \phi_0\rangle$	$=$	$ 0\rangle$	
$ \phi_{n+1}\rangle$	$=$	$\frac{1}{\sqrt{n+1}}\hat{a}^\dagger \phi_n\rangle$	
$ \phi_{n-1}\rangle$	$=$	$\frac{1}{\sqrt{n}}\hat{a} \phi_n\rangle$	
$ \phi_n\rangle$	$=$	$\frac{(\hat{a}^\dagger)^n}{\sqrt{n!}} \phi_0\rangle.$	

The creation, annihilation and number operators \hat{a}^\dagger , \hat{a} and \hat{N} act on an infinite dimensional space, but it is still possible to get a flavour of their matrix representations. Considering only the 4 lowest-order states explicitly, write down the matrix representations of \hat{a}^\dagger , \hat{a} and \hat{N} . In the matrix representation, show explicitly that $\hat{a}^\dagger = (\hat{a})^{*T}$, and $\hat{N} = \hat{a}^\dagger\hat{a}$. Show how the sequence terminates when the annihilation operator is applied to the lowest-order state.

This scheme generates an orthonormal set of states, which correspond to discrete increases in the total energy.

This structure has been generated simply from knowing the commutation relationship between \hat{a} and \hat{a}^\dagger .

Because the *number states* form a complete set, other states can be generated by weighted superposition, leading to other quantum excitations where the number of

photons is uncertain!

Coherent states, which are outside the scope of this course, have particular significance. They are the states that are most similar to classical electromagnetic fields. They comprise a weighted linear combination of number states, and as such a measurement of the energy or the number of photons is quantum mechanically uncertain. Even the number of particles in a resonator can be fundamentally uncertain until observed!

Write down the matrix that represents the behaviour of the creation operator \hat{a}^\dagger (show only the first 4 rows and columns). Write down the matrix that represents the behaviour of the annihilation operator \hat{a} (show only the first 4 rows and columns). By multiplication, derive a matrix representation of the number operator \hat{N} , and calculate its eigenvalues.

The fact that it is only possible to increase, or decrease, the energy of a simple harmonic oscillator in units of $\hbar\omega$, or quanta, is of substantial importance.

- It shows that the stored energy cannot be changed in a continuous way.
- It shows that there is a deep relationship between wavelike behaviour and particle-like behaviour, which is where the problems with classical mechanics started.

In elastic systems, these quanta are called *phonons*, whereas in electromagnetic, or equivalently electrical, systems they are known as *photons*.

Magnetic spin waves are also quantised, and their quanta are called *magnons*.

(Such general relevance derives from the fact that most potentials close to a minimum allow for a quadratic, or harmonic, approximation and therefore the system can be modelled as a simple harmonic oscillator.)

Notice that each of the high-order states have multiple quanta ‘in them’; in other words many photons can ‘exist’ within a single (energy) eigenstate. Particles, or equivalently quanta, for which this is true are called *bosons*.

Later we will talk about *fermions*, which have a different character.

50 Wavefunction analysis

Consider the simple harmonic oscillator from the perspective of wavefunction analysis.

Can we find the lowest-order wavefunction corresponding to the ground state?

Using $\hat{a}|\phi_0\rangle = |0\rangle$ expressed in terms of position and momentum operators acting on a wavefunction,

$$\begin{aligned} \left[\sqrt{\frac{m\omega}{2\hbar}}\hat{x} + i\frac{1}{\sqrt{2m\hbar\omega}}\hat{p} \right] \phi_0(x) &= 0 \quad (419) \\ \sqrt{\frac{m\omega}{2\hbar}}x\phi_0(x) + \sqrt{\frac{\hbar}{2m\omega}}\frac{d\phi_0(x)}{dx} &= 0 \\ \frac{d\phi_0(x)}{dx} + \frac{m\omega}{\hbar}x\phi_0(x) &= 0, \end{aligned}$$

we find the solution

$$\phi_0(x) = c_0 e^{-m\omega x^2/2\hbar}. \quad (420)$$

We need to normalise the wavefunction

$$\int_{-\infty}^{+\infty} c_0^2 e^{-m\omega x^2/\hbar} dx = 1. \quad (421)$$

Using the substitution $u = x\sqrt{m\omega/\hbar}$, we obtain

$$\begin{aligned} \int_{-\infty}^{+\infty} e^{-m\omega x^2/\hbar} dx &= \sqrt{\frac{\hbar}{m\omega}} \int_{-\infty}^{+\infty} e^{-u^2} du \quad (422) \\ &= \sqrt{\frac{\pi\hbar}{m\omega}}, \end{aligned}$$

and we conclude that

$$\boxed{\phi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-m\omega x^2/2\hbar}.} \quad (423)$$

This ground state can be compared to our previous wavefunction analysis of a particle in a quadratic potential.

We can now generate the higher-order states using Eq. (418)

$$\begin{aligned} |\phi_n\rangle &= \frac{(\hat{a}^\dagger)^n}{\sqrt{n!}} |\phi_0\rangle \quad (424) \\ \phi_n(x) &= \frac{1}{\sqrt{n!}} \left[\sqrt{\frac{m\omega}{2\hbar}} x - \sqrt{\frac{\hbar}{2m\omega}} \frac{d}{dx} \right]^n \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-m\omega x^2/2\hbar} \\ &= c_n (2\hbar m\omega)^{-n/2} \left[m\omega x - \hbar \frac{d}{dx}\right]^n e^{-m\omega x^2/2\hbar}, \end{aligned}$$

where the normalisation constant is

$$c_n = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \frac{1}{\sqrt{n!}}. \quad (425)$$

Evaluating these derivatives leads to precisely the same set of functions, the Hermite polynomials, that were calculated in Handout V for the wavefunctions of a quadratic potential.

Indeed the Hermite functions can be generated recursively, corresponding to the ladder operators introduced here.

One can show that, for these number states,

$$\Delta x = \sqrt{\left(n + \frac{1}{2}\right) \frac{\hbar}{m\omega}} \quad (426)$$

$$\Delta p = \sqrt{\left(n + \frac{1}{2}\right) \hbar m\omega},$$

and therefore

$$\boxed{\Delta x \Delta p = \left(n + \frac{1}{2}\right) \hbar.} \quad (427)$$

Only the ground state is a minimum uncertainty state, the others are not.

It is the minimum uncertainty state for this general class of problem.

51 Density operator

Another important operator is the *density operator*.

It can be used to describe the state of a system comprising elements that behave quantum mechanically at the individual level, but which are combined classically.

Consider two boxes, each of which holds an electron. Should the wavefunctions be combined to form a single, large quantum mechanical system, or should they be regarded as independent?

Should we form weighted linear combinations using all of the states, or should we calculate the probabilistic behaviour of the two systems separately, and then combine the probabilities?

The answer lies in whether the systems are quantum mechanically coupled or not, or equivalently whether the

wavefunctions are *coherent* or not.

More general examples occur in the case of thermodynamics, where a large system is in some quantum state, but because of ignorance, it is not known precisely which state the system is in.

How does one describe a classical ensemble of quantum systems?

If a system is in quantum state $|\psi\rangle$, then the expectation value of a general operator \hat{A} is

$$\begin{aligned}\langle A \rangle &= \langle \psi | \hat{A} | \psi \rangle \\ &= \sum_r \langle \psi | u_r \rangle \langle u_r | \hat{A} | \psi \rangle \\ &= \sum_r \langle u_r | \hat{A} | \psi \rangle \langle \psi | u_r \rangle \\ &= \text{Tr} [\hat{A} |\psi\rangle\langle\psi|],\end{aligned}\tag{428}$$

where $\{|u_r\rangle\}$ is any complete orthonormal basis set.

To this point we simply have

$$\boxed{\langle A \rangle = \text{Tr} [\hat{A} |\psi\rangle\langle\psi|]}\tag{429}$$

as a notationally different, but mathematically identical, way of writing

$$\langle A \rangle = \langle \psi | \hat{A} | \psi \rangle.\tag{430}$$

Now suppose that we have a classical ensemble of such systems, and that each member of the ensemble can be in one of a number of different quantum states,

$$|\psi_1\rangle, \dots |\psi_i\rangle, \dots |\psi_I\rangle,\tag{431}$$

with probabilities P_1 , P_i and P_I respectively.

Then

$$\begin{aligned}\langle A \rangle &= \sum_{i=1}^I P_i \langle \psi_i | \hat{A} | \psi_i \rangle \\ &= \text{Tr} \left[\sum_{i=1}^I P_i |\psi_i\rangle\langle\psi_i| \hat{A} \right],\end{aligned}\quad (432)$$

which gives

$$\boxed{\langle A \rangle = \text{Tr} [\hat{O} \hat{A}]} \quad (433)$$

where

$$\boxed{\hat{O} = \sum_{i=1}^I P_i |\psi_i\rangle\langle\psi_i|} \quad (434)$$

is called the *density operator*.



John von Neumann
(1903-1957)

It allows both quantum mechanical and classical expectations to be handled by a single operator.

Notice that one adds quantum states *coherently* by forming linear combinations of state vectors, whereas one adds states *incoherently* by forming linear combinations of density operators.

A state is called *pure* if $\hat{O} = |\psi\rangle\langle\psi|$, for some quantum state $|\psi\rangle$.

The density operator has various properties:

$$\begin{aligned}\text{Tr} [\hat{O}] &= 1 \quad \text{normalisation condition} \\ \hat{O}^\dagger &= \hat{O} \quad \text{self adjoint} \\ \hat{O}\hat{O} &= \hat{O} \quad \text{idempotent (projector) for a pure state only}\end{aligned}$$

Consider a weighted linear combination of two quantum states:

$$\begin{aligned} |\psi\rangle &= c_1|\psi_1\rangle + c_2|\psi_2\rangle & (435) \\ &= \sqrt{P_1}e^{i\phi_1}|\psi_1\rangle + \sqrt{P_2}e^{i\phi_2}|\psi_2\rangle \\ &= \sqrt{P_1}e^{i\phi_1}\left[|\psi_1\rangle + \frac{\sqrt{P_2}}{\sqrt{P_1}}e^{i(\phi_2-\phi_1)}|\psi_2\rangle\right]. \end{aligned}$$

The resultant vector, and therefore the outcome of any experiment, depends on the relative quantum phase difference between the two states.

Indeed this allows for interference phenomena.

Consider instead a weighted linear combination of two density operators based on the same states:

$$\begin{aligned} |\psi\rangle\langle\psi| &= c_1c_1^*|\psi_1\rangle\langle\psi_1| + c_2c_2^*|\psi_2\rangle\langle\psi_2| & (436) \\ &= P_1\left[|\psi_1\rangle\langle\psi_1| + \frac{P_2}{P_1}|\psi_2\rangle\langle\psi_2|\right]. \end{aligned}$$

The resultant density matrix does not depend on the relative phase difference between those states, and therefore the outcome of an experiment is the same as if two systems, one in state $|\psi_1\rangle$ and one in state $|\psi_2\rangle$, were experimented on separately, with probabilities P_1 and P_2 .

This general idea gives rise to the notion of *entanglement* where one distinguishes between systems that are quantum mechanically entangled and systems that are not.

We shall see later that a pure state can only evolve with time into another pure state, retaining the basic character of the scheme.

If however the system interacts with some external environment, then phase coherence can be lost, a process known as *decoherence*.

A state having the form of Eq. (434) where only one term is present,

$$\hat{O} = |\psi_i\rangle\langle\psi_i|, \quad (437)$$

is called a *pure state*, whereas a state having two or more terms present,

$$\hat{O} = \sum_{i=1}^I P_i |\psi_i\rangle\langle\psi_i| \quad \text{with } I \geq 2, \quad (438)$$

is called a *mixed state*.

Interestingly, the notion of absolute phase disappears when the density operator is used, even in the case of a pure state, and so this formulation has a certain elegance.

52 Functions of operators

Given that measurable quantities correspond to operators, we have to be concerned with functions of operators.

For example, we know that energy can be measured through the Hamiltonian operator, but a measurement might measure some function of the energy, rather than the energy directly itself.

What if we want to determine

$$e^{-E/kT}; \quad (439)$$

surely this can be measured, but what is the corresponding operator?

It seems natural to write

$$e^{-\hat{H}/kT}, \quad (440)$$

but what does it mean to have a function of an operator?

Indeed, what is

$$e^{-\mathbf{H}/kT}, \quad (441)$$

where \mathbf{H} is a matrix?

The definition is that one uses the power-series of the function, and replaces the variable with the operator.

Thus, because

$$e^{-x} = 1 - x + \frac{x^2}{2} - \frac{x^3}{3!} + \dots, \quad (442)$$

then

$$e^{-\hat{X}} = \hat{I} - \hat{X} + \frac{\hat{X}^2}{2} - \frac{\hat{X}^3}{3!} + \dots. \quad (443)$$

Suppose that we express some general operator \hat{X} in terms of its eigenvectors x_i , or equivalently in diagonal form:

$$\hat{X} = \sum_i x_i |\psi_i\rangle\langle\psi_i|. \quad (444)$$

Then

$$\begin{aligned} \hat{X}^2 &= \sum_{ij} x_i x_j |\psi_i\rangle\langle\psi_i|\langle\psi_j|\psi_j| \\ &= \sum_i x_i^2 |\psi_i\rangle\langle\psi_i|, \\ \hat{X}^n &= \sum_i x_i^n |\psi_i\rangle\langle\psi_i|, \end{aligned} \quad (445)$$

and finally

$$e^{-\hat{X}} = \sum_i e^{-x_i} |\psi_i\rangle\langle\psi_i|. \quad (446)$$

In general terms,

$$F(\hat{X}) = \sum_i F(x_i) |\psi_i\rangle\langle\psi_i|. \quad (447)$$

The operator $F(\hat{X})$ has the same eigenvectors as \hat{X} , but it has eigenvalues $F(x_i)$ rather than x_i .

This is indeed what we would have intuitively required, because a measurement returns one of the eigenvalues of the associated operator, and therefore the measurement of some function of an operator should return the same function of the eigenvalues.

From a mapping perspective, the RHS of Eq. (447) provides an interpretation for the action of $F(\hat{X})$ on any general vector $|\eta\rangle$. Namely, project $|\eta\rangle$ onto each of the eigenvectors of \hat{X} , scale by the corresponding $F(x_i)$, and reconstruct the result by a weighted linear combination of eigenvectors.

This same reasoning is used for all functions even though in many cases it is not clear whether a convergent power-series expansion exists!

As an example, consider the density operator

$$\hat{\rho} = \frac{1}{Z} e^{-\hat{H}/kT}, \quad (448)$$

where \hat{H} is the Hamiltonian of a simple harmonic oscillator,

$$\hat{H}|\phi_n\rangle = E_n|\phi_n\rangle = \left(n + \frac{1}{2}\right)\hbar\omega|\phi_n\rangle, \quad (449)$$

and Z is a normalisation factor given by

$$Z = \text{Tr} \left[e^{-\hat{H}/kT} \right]. \quad (450)$$

These are called *thermal states* because they represent thermodynamic ensembles of quantum systems.

Indeed this opens the door to combining the principles

of thermodynamics and quantum mechanics, but this topic is outside of the scope of this course.

What is the expectation value of \hat{H} for a thermal state?

$$\langle H \rangle = \text{Tr} [\hat{\rho} \hat{H}] . \quad (451)$$

To make the process easy, represent $\hat{\rho}$ in terms of the eigenvectors of \hat{H} :

$$\hat{\rho} = \frac{1}{Z} \sum_{n=0}^{\infty} e^{-E_n/kT} |\phi_n\rangle\langle\phi_n|, \quad (452)$$

which follows from (446).

In other words the classical probability of the system occupying state n is proportional to $e^{-E_n/kT}$.

We need to normalise the density operator such that $\text{Tr} [\hat{\rho}] = 1$:

$$\begin{aligned} \text{Tr} [\hat{\rho}] &= \frac{1}{Z} \sum_{n=0,m=0}^{\infty} e^{-E_n/kT} \langle \phi_m | \phi_n \rangle \langle \phi_n | \phi_m \rangle \\ &= \frac{1}{Z} \sum_{n=0}^{\infty} e^{-E_n/kT} \\ &= \frac{1}{Z} \sum_{n=0}^{\infty} e^{-(n+\frac{1}{2})\hbar\omega/kT} \\ &= 1. \end{aligned} \quad (453)$$

This is indeed consistent with the choice in Eq. (450), and Z could have been calculated directly from it.

To evaluate the sum use

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}, \quad (454)$$

for $x < 1$ to give

$$\frac{1}{Z} = 2\sinh(\hbar\omega/2kT), \quad (455)$$

and

$$\hat{\rho} = 2\sinh(\hbar\omega/2kT) \sum_{n=0}^{\infty} e^{-E_n/kT} |\phi_n\rangle\langle\phi_n|. \quad (456)$$

Now

$$\begin{aligned} \langle \hat{H} \rangle &= \text{Tr} [\hat{\rho} \hat{H}] \\ &= 2\sinh(\hbar\omega/2kT) \text{Tr} \left[\sum_{n=0, m=0}^{\infty} e^{-E_n/kT} E_m |\phi_n\rangle\langle\phi_n| \langle\phi_m|\right] \\ &= 2\sinh(\hbar\omega/2kT) \text{Tr} \left[\sum_{n=0}^{\infty} e^{-E_n/kT} E_n |\phi_n\rangle\langle\phi_n| \right] \\ &= 2\sinh(\hbar\omega/2kT) \sum_{n=0}^{\infty} E_n e^{-E_n/kT}, \end{aligned} \quad (457)$$

$$\langle E \rangle = \hbar\omega \left[\frac{1}{2} + \frac{1}{e^{\hbar\omega/kT} - 1} \right], \quad (458)$$

where the summation was evaluated using

$$\sum_{n=0}^{\infty} nx^n = x \frac{d}{dx} \sum_{n=0}^{\infty} x^n = \frac{x}{(1-x)^2}. \quad (459)$$

Eq. (458) is the well-known expression for the average energy of a simple harmonic oscillator weakly coupled to a heat bath having temperature T , and the occupancies of the states in the ensemble have the known thermodynamic form (i.e., proportional to the corresponding Boltzmann factor).

As another example, consider the matrix elements of the density operator \hat{O} in Eq. (438) relative to some general basis $\{|a_m\rangle\}$:

$$A_{mn} = \langle a_m | \hat{O} | a_n \rangle \quad (460)$$

$$\begin{aligned} A_{mn} &= \sum_{i=1}^I P_i \langle a_m | \psi_i \rangle \langle \psi_i | a_n \rangle, \\ \hat{O} &= \sum_{mn} A_{mn} |a_m\rangle\langle a_n|, \end{aligned}$$

which is not in *diagonal form*.

Consider the trace of the density operator itself:

$$\begin{aligned}\text{Tr} [\hat{O}] &= \sum_j \sum_{i=1}^I P_i \langle \psi_j | \psi_i \rangle \langle \psi_i | \psi_j \rangle \\ &= \sum_{i=1}^I P_i \\ \text{Tr} [\hat{O}] &= 1,\end{aligned}\tag{461}$$

which is the necessary *normalisation condition* for the density operator.

Also, the matrix elements of the density operator are the same as those of its adjoint; in other words, density operators are self adjoint:

$$\hat{O}^\dagger = \hat{O}.\tag{462}$$

Consider the composite operator

$$\begin{aligned}\hat{O}\hat{O} &= \sum_{i=1, j=1}^I P_i P_j |\psi_i\rangle \langle \psi_i| \langle \psi_j|\langle \psi_j| \\ &= \sum_{i=1}^I P_i^2 |\psi_i\rangle \langle \psi_i|.\end{aligned}\tag{463}$$

If, and only if, the system is in a pure state, only one term is present in the summation, and P_1 must equal unity. Then

$$\hat{O}\hat{O} = |\psi_i\rangle \langle \psi_i| = \hat{O}.\tag{464}$$

An operator for which this is true is called *idempotent*, and once the density operator is known it can be used as a check as to whether the system is in a pure state.

When two or more states are present there are techniques for exploring how many pure states are present, and their relative weightings, but we shall not cover that topic here.

VIII Time-Dependent Quantum Mechanics

In the preceding chapters, we introduced two formulations of quantum mechanics, in terms of Schrödinger's equation and wave functions, and in terms of vector spaces and operators. In both cases, time is a parameter. For example, the operator algebra we have discussed applies at an instant in time, and at every time. We do *not* integrate over t to find the probability that an event happened within some time window T . The notion that the event happened, but that we simply did not know about until the observation was made, is not an appropriate way to think. We can only calculate the probability of observing some outcome at a chosen time t . This begs the question as to how states, and perhaps operators, evolve in time. In this chapter we shall discuss time evolution in quantum mechanics. Like many other aspects, we shall see that its dynamical behaviour can be described concisely in terms of commutation relations. We shall introduce the so-called *Schrödinger picture*, where observables such as position \mathbf{r} and momentum \mathbf{p} are associated with operators, such as $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}} = -i\hbar\nabla$, that do not change with time. This is very different from classical mechanics where position and momentum are dynamical variables that vary with time in familiar ways. In the *Schrödinger picture*, quantum states vary with time, and therefore the likely outcome of an experiment varies with time. We shall also introduce the *Heisenberg picture*, where quantum states do not change with time, but operators do. In many ways, this is closer to the classical viewpoint.

53 Postulates of quantum mechanics

The standard interpretation of quantum mechanics is called the *Copenhagen Interpretation*, following the famous Solvay Conference of 1927.

It is defined by the following five postulates, which we have already encountered:

1. The most complete knowledge of a system is represented by a state vector $|\psi\rangle$.
2. For every observable A there exists a linear Hermitian operator \hat{A} , and the result of a measurement of A must be one of the eigenvalues of \hat{A} .
3. If eigenvalue a_n corresponds to eigenstate $|a_n\rangle$, then the probability of obtaining result a_n when the system is in state $|\psi\rangle$ is $|\langle a_n | \psi \rangle|^2$.
4. As a result of a measurement of \hat{A} in which the value a_n is obtained, the state of the system is changed discontinuously to the corresponding eigenstate $|a_n\rangle$.
5. Between measurements, the state vector $|\psi\rangle$ evolves deterministically in time according to the differential equation $i\hbar\partial|\psi\rangle/\partial t = \hat{H}|\psi\rangle$, where \hat{H} is the operator corresponding to the total energy.

Following a measurement, because of the collapse of the wavefunction, the time evolution must be $i\hbar\partial|a_n\rangle/\partial t = \hat{H}|a_n\rangle$, where $|a_n\rangle$ is the eigenvector corresponding to the eigenvalue last recorded.

Quantum mechanics seems to comprise periods of deterministic evolution separated by discontinuous jumps.

This idea caused much controversy, and still does!

54 State propagator

Suppose that a system is in some general state $|\psi(t)\rangle$. It can be expanded as a weighted linear combination of the eigenvectors of \hat{H} :

$$|\psi(t)\rangle = \sum_n c_n(t) |\phi_n\rangle. \quad (465)$$

The time dependence appears in the coefficients because the eigenvectors form a complete basis for any t (the possible outcomes do not change with time); it is their relative weights that change with time.

The time-dependent Schrödinger equation is

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle. \quad (466)$$

Substituting Eq. (465) into it, we find

$$\begin{aligned} i\hbar \sum_n \frac{dc_n(t)}{dt} |\phi_n\rangle &= \sum_n c_n(t) \hat{H} |\phi_n\rangle \quad (467) \\ &= \sum_n c_n(t) E_n |\phi_n\rangle \\ \sum_n \left[i\hbar \frac{dc_n(t)}{dt} - c_n(t) E_n \right] |\phi_n\rangle &= 0 \\ \left[i\hbar \frac{dc_n(t)}{dt} - c_n(t) E_n \right] &= 0 \quad \forall n, \end{aligned}$$

where the last line follows from the orthogonality of the eigenvectors.

Solving the differential equation gives

$$c_n(t) = c_n(0) e^{-iE_n t / \hbar}. \quad (468)$$

The phases of energy eigenstates evolve at different rates:

$|\psi(t)\rangle = \sum_n c_n(0) e^{-iE_n t / \hbar} |\phi_n\rangle. \quad (469)$

Remember that the $|\phi_n\rangle$ are energy eigenstates.

The time-dependence is ‘separate’ from the state.

In the case of wavefunctions

$$\psi(x, t) = \sum_n e^{-iE_n t/\hbar} c_n(0) \phi_n(x). \quad (470)$$

General time dependent behaviour is described by a superposition of ‘frequency’ components, each one of which corresponds to an energy that might be recorded when a measurement of energy takes place: $\omega_n = E_n/\hbar$.

As time progresses, the frequency components change in phase relative to each other, giving rise to an evolving wavevector.

If only one frequency component is present, the energy is certain, and the *form* of the state does not change with time. These are known as *stationary* states.

Notice that the relative amplitudes of the states remain constant.

Eq. (469) applies when the general time-dependent state vector $|\psi(t)\rangle$ is expressed in terms of a weighted linear combination of the eigenvectors of \hat{H} , but one wonders if there is a more compact expression that can be applied to state vectors generally, without the decomposition.

The exponential prefactor in (469) is simply a function of each of the eigenvalues; but we know from Handout VII that

$$F(\hat{X}) |\psi_n\rangle = F(x_n) |\psi_n\rangle, \quad (471)$$

and therefore

$$\begin{aligned} |\psi(t)\rangle &= \sum_n c_n(0) e^{-i\hat{H}t/\hbar} |\phi_n\rangle \\ &= e^{-i\hat{H}t/\hbar} \sum_n c_n(0) |\phi_n\rangle \\ &= e^{-i\hat{H}t/\hbar} |\psi(0)\rangle. \end{aligned} \quad (472)$$

We conclude that

$$|\psi(t)\rangle = e^{-i\hat{H}t/\hbar} |\psi(0)\rangle, \quad (473)$$

which is a general expression for the time dependence of any state vector (where one must remember what is meant by a ‘function of an operator’).

Take Eq. (471), form the outer product on each side with $\langle\psi_n|$, sum over n remembering that the basis is complete ($\sum_n |\psi_n\rangle\langle\psi_n| = \hat{I}$), to give:

$$F(\hat{X}) = \sum_n F(x_n) |\psi_n\rangle\langle\psi_n|. \quad (474)$$

(This expression was already discussed in Handout VII.)

In the current context

$$\begin{aligned} e^{-i\hat{H}t/\hbar} &= \sum_n e^{-iE_nt/\hbar} |\phi_n\rangle\langle\phi_n| \\ &= \sum_n e^{-i\omega_nt} |\phi_n\rangle\langle\phi_n| \end{aligned} \quad (475)$$

Thus in (473), the time evolution occurs by projecting a general vector $|\psi(0)\rangle$ onto each of the energy eigenstates, multiplying by $e^{-i\omega_nt}$, and reconstructing the result through a weighted summation of eigenstates.

It is simply an elegant way of applying phase factors to the individual energy components.

The state vector at t can also be related to the state vector at some earlier time t_0 through

$$|\psi(t)\rangle = \hat{U}(t, t_0)|\psi(t_0)\rangle. \quad (476)$$

where the operator

$$\hat{U}(t, t_0) = e^{-i\hat{H}(t-t_0)/\hbar} \quad (477)$$

is called the *time shift* (or *time translation*, or *time evolution*) operator.

It can be easily verified that it is a unitary operator, defined by

$$\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{I}. \quad (478)$$

Show that the time-translation operator obeys a composition law, where a total shift can be constructed from a number of smaller shifts. What does the inverse of the time-translation operator do? What does this say about the evolution of a quantum state, and why is it not possible to accommodate the collapse of a state within this formalism?

Eq. (466) gives

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle &= \hat{H}|\psi(t)\rangle. \\ i\hbar\frac{\partial}{\partial t}\hat{U}|\psi(0)\rangle &= \hat{H}\hat{U}|\psi(0)\rangle. \end{aligned} \quad (479)$$

$$i\hbar\frac{d}{dt}\hat{U} = \hat{H}\hat{U}, \quad (480)$$

where the last line follows because (479) must be true for all wave vectors. Take each of a complete set of eigenvectors in turn, $|\psi_n(0)\rangle$, form the inner product with each of $\langle\psi_m(0)|$, and note that one is then comparing the matrix elements of $i\hbar\partial/\partial t\hat{U}$ and $\hat{H}\hat{U}$.

By expressing each of \hat{H} and \hat{U} in terms of outer products of energy eigenstates, show that

$$[\hat{U}, \hat{H}] = \hat{0}. \quad (481)$$

Indeed, the eigenstates of \hat{U} are the energy eigenstates, as demonstrated by Eq. (475).

\hat{U} has many other properties, such as preserving inner products under time translations, preserving normalisation, preserving the traces of operators, etc.

$e^{-i\hat{H}t/\hbar}$ is a unitary operator, indicating that only the phases of quantities change when it is applied.

55 Ehrenfest's Theorem

State vectors evolve in a fully deterministic way in quantum mechanics, but the outcome of any particular measurement is probabilistic.

This means that one cannot define the trajectory of a particle, which may seem to jump around randomly between subsequent measurements.

Therefore, one cannot define the time derivative of an observable such as position!

However, it is possible to consider the time evolution of the *expectation value* of an operator.

For a system in state $|\psi(t)\rangle$, the expectation value of some general operator \hat{A} at time t is

$$\langle \hat{A} \rangle = \langle \psi(t) | \hat{A} | \psi(t) \rangle, \quad (482)$$

where the time dependence has been included explicitly *as a parameter*.

The expectation value of \hat{A} will generally change with time:

$$\begin{aligned}\frac{d\langle \hat{A} \rangle}{dt} &= \frac{d\langle \psi(t) |}{dt} \hat{A} |\psi(t)\rangle + \langle \psi(t) | \frac{d\hat{A}}{dt} |\psi(t)\rangle + \langle \psi(t) | \hat{A} \frac{d|\psi(t)\rangle}{dt} \\ &= \frac{-1}{i\hbar} \langle \psi(t) | \hat{H}^\dagger \hat{A} |\psi(t)\rangle + \langle \frac{d\hat{A}}{dt} \rangle + \frac{1}{i\hbar} \langle \psi(t) | \hat{A} \hat{H} |\psi(t)\rangle \\ &= \frac{i}{\hbar} \langle \psi(t) | [\hat{H}^\dagger \hat{A} - \hat{A} \hat{H}] |\psi(t)\rangle + \langle \frac{d\hat{A}}{dt} \rangle \\ &= \frac{i}{\hbar} \langle \psi(t) | [\hat{H} \hat{A} - \hat{A} \hat{H}] |\psi(t)\rangle + \langle \frac{d\hat{A}}{dt} \rangle,\end{aligned}\quad (483)$$

giving

$$\boxed{\frac{d\langle \hat{A} \rangle}{dt} = \frac{i}{\hbar} \langle [\hat{H}, \hat{A}] \rangle + \langle \frac{d\hat{A}}{dt} \rangle}, \quad (484)$$

which is called *Ehrenfest's Theorem*.

It is an equation of motion for the expectation value of any operator \hat{A} .



Paul Ehrenfest
(1880-1933)

In the Schrödinger picture, and strictly for conservative systems where \hat{H} is independent of time, operators corresponding to observables are typically independent of time, and we find

$$\boxed{\frac{d\langle \hat{A} \rangle}{dt} = \frac{1}{\hbar} \langle i [\hat{H}, \hat{A}] \rangle}. \quad (485)$$

- Eq. (484) indicates that, for $\hat{A} = \hat{H}$, the expectation value of the total energy is time invariant.

- If \hat{A} commutes with \hat{H} , the eigenstates of \hat{A} and \hat{H} are the same, and the observable is a constant of motion (assuming there is no explicit time dependence in \hat{A} itself).
- if a system is in an energy eigenstate, namely a stationary state, the expectation values of *all* observables are independent of time.
- If, however, \hat{A} does not commute with \hat{H} , then a measurement of \hat{A} will prepare the system in an eigenstate of \hat{A} and not in a stationary state. Therefore, a *later* measurement of \hat{A} , after time evolving the system, may yield a different result as it will generically no longer be in the same eigenstate.

56 Time propagation using matrix elements

The eigenvectors and eigenvalues of \hat{H} play a central role in determining the time evolution of a quantum system.

But how does one use Eq. (473) in practice?

The time dependence of the expectation value an operator can be expressed neatly in terms of the operator's matrix elements with respect to the energy eigenstates.

According to (469), any general state $|\psi(t)\rangle$ evolves according to

$$|\psi(t)\rangle = \sum_n c_n(0) e^{-iE_n t/\hbar} |\phi_n\rangle, \quad (486)$$

where

$$c_n(0) = \langle \phi_n | \psi(0) \rangle. \quad (487)$$

Similarly,

$$\langle \psi(t) | = \sum_m c_m^*(0) e^{+iE_m t/\hbar} \langle \phi_m |, \quad (488)$$

and so

$$\begin{aligned}\langle \hat{A} \rangle &= \langle \psi(t) | \hat{A} | \psi(t) \rangle \\ &= \sum_{mn} c_m^* c_n e^{i(E_m - E_n)t/\hbar} \langle \phi_m | \hat{A} | \phi_n \rangle \\ &= \sum_{mn} c_m^* c_n e^{i(E_m - E_n)t/\hbar} A_{mn}.\end{aligned}\tag{489}$$

Thus

$$\boxed{\langle \hat{A} \rangle = \sum_{mn} c_m^* c_n e^{i(E_m - E_n)t/\hbar} A_{mn}},\tag{490}$$

where

$$A_{mn} = \langle \phi_m | \hat{A} | \phi_n \rangle\tag{491}$$

are the matrix elements formed by projecting the time-independent operator onto the stationary states.

Eq. (490) provides a practical way of calculating the time dependence of expectation values.

Consider why Eq. (490) requires the system to be conservative, with no external sources or sinks of energy.

57 Useful commutation relations

Because much of quantum mechanics can be cast into equations involving commutation relations, it is valuable to have efficient ways of evaluating the commutation relations of various complex expressions.

It can be shown by straightforward manipulation that, for any integer l , the following are true:

$$\begin{aligned}[\hat{x}, \hat{p}^l] &= i\hbar l \hat{p}^{l-1} = i\hbar \frac{\partial}{\partial \hat{p}} \hat{p}^l \\ [\hat{p}, \hat{x}^l] &= -i\hbar l \hat{x}^{l-1} = -i\hbar \frac{\partial}{\partial \hat{x}} \hat{x}^l,\end{aligned}\tag{492}$$

where the derivatives have their ‘obvious’ meaning, but are introduced solely for notational convenience.

For any functions $F(\hat{p})$ and $G(\hat{x})$ that can be expanded in power series, it follows that

$$\boxed{\begin{aligned} [\hat{x}, F(\hat{p})] &= i\hbar \frac{\partial F(\hat{p})}{\partial \hat{p}} \\ [\hat{p}, G(\hat{x})] &= -i\hbar \frac{\partial G(\hat{x})}{\partial \hat{x}} \end{aligned}} \quad (493)$$

These expressions are deemed to be true even if it is not known whether a power series exist for $F(\hat{p})$ and $G(\hat{x})$.

As a matter of fact, even more general expressions hold, based on power series of functions of two independent variables:

$$\boxed{\begin{aligned} [\hat{x}, F(\hat{p}, \hat{x})] &= i\hbar \frac{\partial F(\hat{p}, \hat{x})}{\partial \hat{p}} \\ [\hat{p}, F(\hat{p}, \hat{x})] &= -i\hbar \frac{\partial F(\hat{p}, \hat{x})}{\partial \hat{x}} \end{aligned}} \quad (494)$$

The order of the operators must be preserved in $F(\hat{p}, \hat{x})$ when using (494).

For example,

$$\begin{aligned} [\hat{x}, \hat{p}\hat{x}\hat{p}] &= i\hbar \frac{\partial}{\partial \hat{p}} (\hat{p}\hat{x}\hat{p}) \\ &= i\hbar (\hat{x}\hat{p} + \hat{p}\hat{x}) . \end{aligned} \quad (495)$$

58 Potential as an operator

The Hamiltonian is given by

$$\hat{H} = \frac{\hat{p}^2}{2m} + \hat{V}, \quad (496)$$

but we have not given much thought as to what we mean by the potential operator \hat{V} .

It is an observable, and therefore must take the form of an operator.

In the case of a wavefunction, it is just a scalar function of position (in the cases considered so far), and the problem does not arise.

More generally, it is a function of the operator \hat{x} . In other words $\hat{V} = V(\hat{x})$.

We saw previously that functions of operators are defined in terms of their power series.

In fact, given that an operator is an observable, then a function of an operator must also be an observable.

In general, following Eq. (494),

$$[\hat{A}, F(\hat{A})] = \hat{0}. \quad (497)$$

giving

$$[\hat{x}, F(\hat{x})] = \hat{0}, \quad (498)$$

and

$$[\hat{p}, F(\hat{p})] = \hat{0}. \quad (499)$$

One now sees that

$$[\hat{x}, V(\hat{x})] = \hat{0}, \quad (500)$$

but

$$\begin{aligned} [\hat{p}, V(\hat{x})] &\neq \hat{0} \\ &= -i\hbar \frac{dV(\hat{x})}{d\hat{x}}. \end{aligned} \quad (501)$$

59 Ehrenfest's Theorem and classical limits

Do classical trajectories make any sense in the context of Ehrenfest's Theorem?

Consider Eq. (485) in the case where $\hat{A} = \hat{x}$:

$$\frac{d\langle \hat{x} \rangle}{dt} = \frac{1}{\hbar} \langle i [\hat{H}, \hat{x}] \rangle. \quad (502)$$

For a particle of mass m moving in potential $V(x)$,

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}), \quad (503)$$

and

$$\frac{d\langle \hat{x} \rangle}{dt} = \frac{i}{\hbar} \left\langle \left[\frac{\hat{p}^2}{2m} + V(\hat{x}), \hat{x} \right] \right\rangle. \quad (504)$$

According to (492)

$$[\hat{p}^2, \hat{x}] = -2i\hbar\hat{p}, \quad (505)$$

and according to (500)

$$[V(\hat{x}), \hat{x}] = \hat{0}, \quad (506)$$

and so

$$\boxed{\frac{d\langle \hat{x} \rangle}{dt} = \frac{\langle \hat{p} \rangle}{m}}, \quad (507)$$

which is a direct analogue of a classical equation of motion.

Now take $\hat{A} = \hat{p}$:

$$\begin{aligned} \frac{d\langle \hat{p} \rangle}{dt} &= \frac{i}{\hbar} \left\langle [\hat{H}, \hat{p}] \right\rangle \quad (508) \\ &= \frac{i}{\hbar} \left\langle \left[\frac{\hat{p}^2}{2m} + V(\hat{x}), \hat{p} \right] \right\rangle \\ &= \frac{i}{\hbar} \langle [V(\hat{x}), \hat{p}] \rangle \\ &= -\langle \frac{dV(\hat{x})}{d\hat{x}} \rangle \\ &= \langle F(\hat{x}) \rangle, \end{aligned}$$

where F is the force.

The correspondence with the classical expression is clear, but be careful: might have we expected $F(\langle \hat{x} \rangle)$?

In general

$$\langle F(\hat{x}) \rangle \neq F(\langle \hat{x} \rangle). \quad (509)$$

Does this make sense at all?

For notational simplicity, suppose that we have some *classical* random variable x , and define $\Delta x = x - \langle x \rangle$:

$$F(x) = F(\langle x \rangle) + F'(\langle x \rangle)\Delta x + F''(\langle x \rangle)\frac{(\Delta x)^2}{2} \dots \quad (510)$$

$$\begin{aligned} \langle F(x) \rangle &= F(\langle x \rangle) + F'(\langle x \rangle)\langle \Delta x \rangle + F''(\langle x \rangle)\frac{\langle (\Delta x)^2 \rangle}{2} \dots \\ &= F(\langle x \rangle) + F''(\langle x \rangle)\frac{\langle (\Delta x)^2 \rangle}{2} \dots \end{aligned}$$

where $F'(\langle x \rangle)$ denotes the derivative of F evaluated at $\langle x \rangle$.

Thus $\langle F(x) \rangle \approx F(\langle x \rangle)$ only when the function changes slowly over the scale of the random variations in x .

Quantum mechanically, and in other words, when $F(\langle x \rangle)$ changes slowly compared with the quantum uncertainty Δx , then $\langle F(x) \rangle \approx F(\langle x \rangle)$; otherwise the expression fails and classical mechanics fails.

Overall, when the quantum mechanical uncertainties are small compared with the scale size of the system, classical mechanics is a good approximation, which is the *Correspondence Principle*.

60 Particle in a well

Let us demonstrate how energy eigenfunctions can be used to trace the time evolution of a quantum state.

Suppose that a particle is in the ground state of an infinite well having width $L/2$, as illustrated in Fig. 52

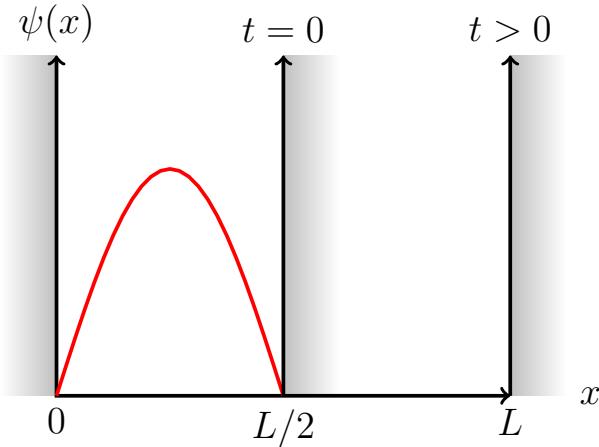


Figure 52: A particle in an infinite potential well. At $t = 0$ the wall at $L/2$ is moved to L .

From Handout V, the normalised wavefunction is

$$\psi(x, t) = \sqrt{\frac{4}{L}} \sin\left(\frac{2\pi x}{L}\right) e^{-i\frac{\hbar}{2m}\left(\frac{2\pi}{L}\right)^2 t}. \quad (511)$$

Suppose that the boundary at $L/2$ is instantaneously shifted to L . In this case, the particle is no longer in an energy eigenstate, and so it evolves in time into something different.

Expand the ‘old’ wavefunction into the energy eigenstates of the ‘new’ well. Because the old wavefunction is contained within the region of the new well, the new eigenfunctions $\phi_n(x, 0)$ must *span* the functional form of the old wavefunction $\psi(x, 0)$:

$$\psi(x, 0) = \sum_{n=1}^{\infty} c_n \phi_n(x, 0), \quad (512)$$

where

$$\phi_n(x, t) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right) e^{-i\frac{\hbar}{2m}\left(\frac{n\pi}{L}\right)^2 t}. \quad (513)$$

The coefficients are given by

$$c_n = \int_0^{L/2} \phi_n^*(x, 0) \psi(x, 0) dx. \quad (514)$$

Note that the integration in the inner product extends from 0 to L , but we made use of the fact that the starting wavefunction only extends from 0 to $L/2$ and vanishes identically between $L/2$ and L .

Thus

$$c_2 = \frac{1}{\sqrt{2}} \text{ for } n = 2, \quad (515)$$

and

$$\begin{aligned} c_n &= \sqrt{\frac{2}{L}} \sqrt{\frac{4}{L}} \int_0^{L/2} \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{2\pi x}{L}\right) dx \quad (516) \\ &= \frac{4\sqrt{2}}{\pi} (-1)^{(n+1)/2} \frac{1}{n^2 - 4} \text{ for } n \text{ odd}, \end{aligned}$$

($c_n = 0$ for all even $n \geq 4$) and so

$$\begin{aligned} \psi(x, t) &= \frac{1}{\sqrt{L}} \sin\left(\frac{2\pi x}{L}\right) e^{-i\frac{\hbar}{2m}\left(\frac{2\pi}{L}\right)^2 t} \quad (517) \\ &+ \frac{8}{\pi\sqrt{L}} \sum_{n \text{ odd}} \frac{(-1)^{(n+1)/2}}{n^2 - 4} \sin\left(\frac{n\pi x}{L}\right) e^{-i\frac{\hbar}{2m}\left(\frac{n\pi}{L}\right)^2 t}. \end{aligned}$$

A number of time captures of the square-amplitude of this function are shown in Fig. 53. The time sequence is red, green, blue and black.

The particle starts off in the LHS of the plot, and sloshes backwards and forwards between the two sides of the well.

This kind of ‘sloshing’ is typical when energy eigenfunctions are combined.

The probability of measuring any one of the energy eigenvalues remains constant even though the overall spatial distribution changes.

The wavefunction behaves in a deterministic way, even though the outcome of any measurement is probabilistic.

More complicated problems can be solved numerically.

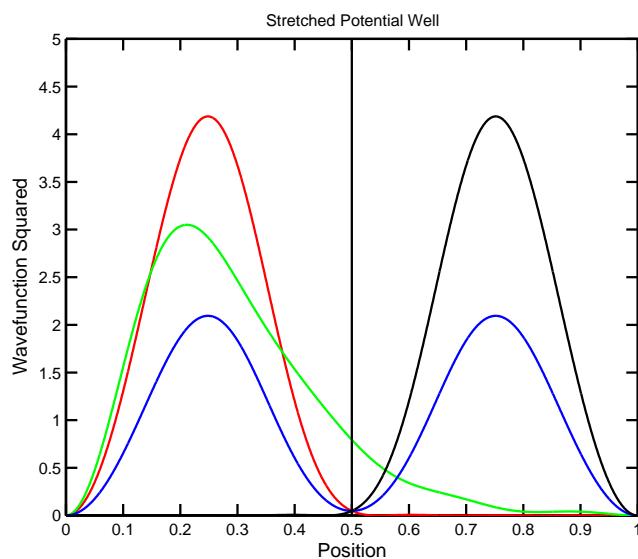


Figure 53: Time evolution of the squared amplitude of the wavefunction discussed in the text, when the width of the infinite potential well is instantaneously stretched from 0.5 to 1.0 at time $t = 0$. Only the eigenstates $n = 1, 2, 3, 5$ were used in the reconstruction. The time sequence is red, green, blue, black.

Just after the wall of the box has been moved, the most probable state to be recorded when a measurement of energy is made is the $n = 2$ state. Show that this energy is the same as the energy of the particle before the wall is moved.

In quantum mechanics, an adiabatic change corresponds to a change that happens sufficiently slowly that if a system is in one of its energy eigenfunctions – a stationary state – the system stays in the same state as time evolves.

In fact, if the system starts in its ground state, it remains in the appropriate ground state as the system is changed.

States do not cross, and this has important consequences for the way in which complicated problems can be handled.

61 Time-energy uncertainty

The general expression for deriving uncertainty relations, discussed in Handout VI, applies also to dynamical variables.

Position and momentum are properties of a dynamical system, and can be measured.

Time on the other hand is a parameter; it identifies the time at which a measurement is made; it is not a measurable property of a system, and so it is not an operator.

Interestingly, however, it satisfies an uncertainty relation when interpreted as the time over which an energy measurement takes place.

Consider the time rate of change of the expectation value of some generic operator \hat{A} ,

$$\frac{d\langle \hat{A} \rangle}{dt} = \frac{i}{\hbar} \langle [\hat{H}, \hat{A}] \rangle. \quad (518)$$

Combining this with the generalised uncertainty relationship between \hat{A} and \hat{H} from Handout VI, we have that

$$\Delta E \Delta A \geq \frac{1}{2} \left| \langle i [\hat{H}, \hat{A}] \rangle \right| = \frac{\hbar}{2} \left| \frac{d\langle \hat{A} \rangle}{dt} \right|, \quad (519)$$

and therefore

$$\boxed{\Delta E \Delta t \geq \frac{\hbar}{2}}, \quad (520)$$

where

$$\Delta t = \Delta A / \left| \frac{d\langle \hat{A} \rangle}{dt} \right|. \quad (521)$$

Δt is the time required for the expectation value of the measured quantity \hat{A} to change by an amount equal to the uncertainty in the same quantity.

Eq. (520) connects the uncertainty in energy to a quantity that is characteristic of the time rate of change of any observable of the system.

If ΔE is small, the time taken for the expectation value of any observable to change *significantly* is *long*.

In the case of a stationary state, $\Delta E \rightarrow 0$ and $\Delta t \rightarrow \infty$, indicating that the expectation value of every observable is constant.

In this sense, there is an uncertainty relationship that applies to energy and time, but its actual meaning must always be remembered.

62 Heisenberg's Picture

Throughout this course we have worked in the so-called *Schrödinger picture* of quantum mechanics.

According to this model, state vectors evolve with time, but for conservative systems operators representing observables do not.

This is at variance with classical mechanics, where dynamical variables change in time and the ‘state’ of the system does not.

There is another approach to quantum mechanics, which is entirely analogous to the Schrödinger picture, but which places all time dependencies on the operators.

Physical quantities, such as probabilities and expectation values, do not depend on the picture, but different pictures have different applications.

We have seen that the expectation value of any observ-

able is given by

$$\langle A(t) \rangle = \langle \psi(t) | \hat{A} | \psi(t) \rangle, \quad (522)$$

where the time dependence is indicated explicitly.

We know, however, that

$$|\psi(t)\rangle = e^{-i\hat{H}t/\hbar} |\psi(0)\rangle, \quad (523)$$

and so (522) can be written as

$$\langle A(t) \rangle = \langle \psi(0) | e^{+i\hat{H}^\dagger t/\hbar} \hat{A} e^{-i\hat{H}t/\hbar} | \psi(0) \rangle. \quad (524)$$

In this expression, $A(t)$ can represent any variable, or indeed function of a variable.

Then

$$\langle A(t) \rangle = \langle \psi(0) | \hat{A}^H(t) | \psi(0) \rangle, \quad (525)$$

where

$$\hat{A}^H(t) = e^{+i\hat{H}t/\hbar} \hat{A} e^{-i\hat{H}t/\hbar}. \quad (526)$$

(Recall that the Hamiltonian is Hermitian: $\hat{H}^\dagger = \hat{H}$.)

For two general times, t and t_0 , where $t \geq t_0$

$$\langle A(t) \rangle = \langle \psi(t_0) | \hat{A}^H(t) | \psi(t_0) \rangle, \quad (527)$$

and

$$\begin{aligned} \hat{A}^H(t) &= e^{+i\hat{H}(t-t_0)/\hbar} \hat{A} e^{-i\hat{H}(t-t_0)/\hbar} \\ &= \hat{U}^\dagger(t, t_0) \hat{A} \hat{U}(t, t_0). \end{aligned} \quad (528)$$

Eq. (525) is called the *Heisenberg picture*, and (526) is the operator \hat{A} in the Heisenberg picture.

It is usually indicated by superscript H , and often the Schrödinger picture is indicated by superscript S .

We shall assume the Schrödinger picture implicitly unless stated otherwise.

They describe exactly the same quantum mechanical processes, but now the time dependence is included in the operators rather than in the state vectors.

In some ways it is closer to the classical situation where dynamical variables, such as position and momentum, change with time.

As we will show below, operators in the Heisenberg picture change with time according to the differential equation

$$i\hbar \frac{d\hat{A}^H}{dt} = [\hat{A}^H, \hat{H}], \quad (529)$$

which is called the *Heisenberg equation*.

It is equivalent to the Schrödinger equation, and is a differential equation that operators must obey in the Heisenberg picture.

Clearly, any operator that commutes with \hat{H} is a constant of motion.

It should be appreciated that operators that commute in the Schrödinger picture also commute in the Heisenberg picture.

For conservative systems, the Hamiltonian has a special characteristic in the two pictures. By definition

$$\hat{H}^H(t) = \hat{U}^\dagger \hat{H} \hat{U} = \hat{H}. \quad (530)$$

Namely, \hat{H}^H is independent of time, and is equal to \hat{H} .

In order to show that operators in the Heisenberg picture

satisfy the differential equation (529), consider

$$\begin{aligned}
 \frac{d\hat{A}^H}{dt} &= \frac{d}{dt}\hat{U}^\dagger \hat{A}\hat{U} \\
 &= \frac{d\hat{U}^\dagger}{dt}\hat{A}\hat{U} + \hat{U}^\dagger \frac{d\hat{A}}{dt}\hat{U} + \hat{U}^\dagger \hat{A} \frac{d\hat{U}}{dt} \\
 i\hbar \frac{d\hat{A}^H}{dt} &= -\hat{U}^\dagger \hat{H}^\dagger \hat{A}\hat{U} + i\hbar \hat{U}^\dagger \frac{d\hat{A}}{dt}\hat{U} + \hat{U}^\dagger \hat{A} \hat{H} \hat{U} \\
 &= -\hat{H}\hat{U}^\dagger \hat{A}\hat{U} + i\hbar \hat{U}^\dagger \frac{d\hat{A}}{dt}\hat{U} + \hat{U}^\dagger \hat{A}\hat{U}\hat{H} \\
 &= [\hat{A}^H, \hat{H}] + i\hbar \hat{U}^\dagger \frac{d\hat{A}}{dt}\hat{U} \\
 &= [\hat{A}^H, \hat{H}]
 \end{aligned} \tag{531}$$

where Eqs. (480) and (481) have been used, and the last line follows for operators in the Schrödinger picture that are independent of time.

63 Time dependence of the density operator

Given that, in the Heisenberg picture, operators rather than state vectors change with time, one wonders whether there is a simple expression for the time evolution of the density operator.

It would then be possible to follow the behaviour of mixed states, not merely pure states.

The density operator has been defined as

$$\hat{O}(t) = \sum_{i=1}^I P_i |\psi_i(t)\rangle \langle \psi_i(t)|, \tag{532}$$

and therefore

$$\begin{aligned}
 i\hbar \frac{d\hat{O}(t)}{dt} &= \sum_{i=1}^I P_i \left(\hat{H} |\psi_i(t)\rangle \langle \psi_i(t)| - |\psi_i(t)\rangle \langle \psi_i(t)| \hat{H}^\dagger \right) \\
 &= \hat{H}\hat{O} - \hat{O}\hat{H}
 \end{aligned} \tag{533}$$

where we used the fact that $i\hbar d|\psi_i(t)\rangle/dt = \hat{H}|\psi_i(t)\rangle$.

In Eq. (533) we have assumed that the classical probabilities P_i do not depend on time, namely stationary classical statistics.

We conclude that

$$i\hbar \frac{d\hat{O}}{dt} = [\hat{H}, \hat{O}], \quad (534)$$

which is called the *von-Neumann* equation.

It is the quantum mechanical analogue of the so-called classical Liouville equation.

It describes the time evolution of a mixed state, and it necessarily includes a wealth of physics.

64 Summary

In summary, the following differential equations must be satisfied:

Schrödinger's equation is a differential equation for the time evolution of state vectors in the Schrödinger picture:

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \hat{H}|\psi(t)\rangle. \quad (535)$$

Heisenberg's equation is a differential equation for the evolution of any observable operator in the Heisenberg picture:

$$i\hbar \frac{d\hat{A}^H}{dt} = [\hat{A}^H, \hat{H}]. \quad (536)$$

Ehrenfest's equation is a differential equation for the time rate of change of the expectation value of any operator:

$$\frac{d\langle \hat{A} \rangle}{dt} = \frac{i}{\hbar} \langle [\hat{H}, \hat{A}] \rangle, \quad (537)$$

(written here in the form where \hat{A} does not intrinsically depends on time).

Overall it is seen that the equations of motion can be written elegantly in terms of commutation relationships.

von-Neumann's equation is a differential equation for the time evolution of the density operator of a mixed state:

$$i\hbar \frac{d\hat{O}}{dt} = [\hat{H}, \hat{O}] . \quad (538)$$

IX Quantum Mechanics in Three Dimensions

In the previous chapter we discussed time evolution in quantum mechanics, starting from the so-called *Schrödinger picture*, where it is quantum states and not observables (operators) that change with time. We introduced the concept of time translation operator and Ehrenfest's theorem describing the evolution of the expectation value of operators. We also discussed the connection to classical mechanics and the time-energy uncertainty principle, and we gave an example involving a particle in an infinite well. We also introduced the so-called *Heisenberg picture*, where quantum states do not change with time, but operators do; in many ways, this is closer to the classical viewpoint. And we gave an example involving the density operator.

In the lecture notes so far we limited our consideration to 1D systems. In this chapter, we shall expand our discussion to higher dimensions, and specifically to 3D systems. This will allow us to introduce the notion of angular momentum (recall that there is no such quantity in 1D), and as we shall see that it plays a fundamental part in quantum mechanics. We will define the angular momentum operator and discuss its properties, commutation relations, eigenstates and eigenvalues. We will encounter again ladder operators in this context, as well as relevant uncertainty relations. We will discuss the connection between angular momentum and magnetic moment, and give examples such as diatomic molecules and the Hydrogen-like atom.

65 Three-dimensional wavefunctions

The extension of wave-mechanics to three dimensions requires the introduction of appropriate wave functions,

$$\Psi(\mathbf{r}, t) = \begin{cases} \Psi(x, y, z, t) \\ \Psi(r, \theta, \phi, t) \end{cases}, \quad (539)$$

and the probability of finding the particle in volume element $d^3\mathbf{r}$ at \mathbf{r} at time t is given by

$$P(\mathbf{r}) = |\Psi(\mathbf{r}, t)|^2 d^3\mathbf{r} = \begin{cases} |\Psi(x, y, z, t)|^2 dx dy dz \\ |\Psi(r, \theta, \phi, t)|^2 r^2 \sin \theta dr d\theta d\phi \end{cases}. \quad (540)$$

The inner product of two state vectors is

$$\langle \Phi | \Psi \rangle = \int_{\mathcal{S}} \Phi^*(\mathbf{r}, t) \Psi(\mathbf{r}, t) d^3\mathbf{r}, \quad (541)$$

where the domain \mathcal{S} must be stated, and the normalisation condition becomes

$$\langle \Psi | \Psi \rangle = \int_{\mathcal{S}} |\Psi(\mathbf{r}, t)|^2 d^3\mathbf{r} = 1. \quad (542)$$

The Cartesian components of the momentum operator are given by

$$\begin{aligned} \hat{p}_x &= -i\hbar \frac{\partial}{\partial x} \\ \hat{p}_y &= -i\hbar \frac{\partial}{\partial y} \\ \hat{p}_z &= -i\hbar \frac{\partial}{\partial z} \end{aligned} \quad (543)$$

or in vector notation

$$\hat{\mathbf{p}} = \hat{p}_x \mathbf{i} + \hat{p}_y \mathbf{j} + \hat{p}_z \mathbf{k} = -i\hbar \nabla. \quad (544)$$

We shall use boldface type to denote spatial vectors; \mathbf{i} , \mathbf{j} , \mathbf{k} to denote unit spatial vectors; and a boldface type

with a hat to denote vector operators, $\hat{\mathbf{p}}$.

A general eigenfunction of the 3D momentum operator $\hat{\mathbf{p}}$ takes the form of a plane-wave

$$\boxed{\Psi(\mathbf{r}, t) \propto e^{i\mathbf{k} \cdot \mathbf{r}} e^{-i\omega t} = e^{i\mathbf{p} \cdot \mathbf{r}/\hbar} e^{-i\omega t}}, \quad (545)$$

where the direction of travel of the phase front, the direction of the wavevector \mathbf{k} , is in the direction of \mathbf{p} .

Also,

$$\boxed{\mathbf{p} = \hbar\mathbf{k}.} \quad (546)$$

Eq. (545) separates into 3 ‘one-dimensional’ functions.

In general, the wavevector \mathbf{k} is a function of ω , or equivalently, the momentum is a function of energy.

It can be shown by direct substitution that

$$-i\hbar\nabla [\Psi(\mathbf{r}, t)] = \mathbf{p}\Psi(\mathbf{r}, t), \quad (547)$$

proving that $\Psi(\mathbf{r}, t)$ is indeed an eigenfunction of the 3D momentum operator $-i\hbar\nabla$, having a vector-valued eigenvalue \mathbf{p} .

The commutation relationships are as usual,

$$\begin{aligned} [\hat{x}, \hat{p}_x] &= i\hbar \\ [\hat{y}, \hat{p}_y] &= i\hbar \\ [\hat{z}, \hat{p}_z] &= i\hbar, \end{aligned} \quad (548)$$

except that

$$\begin{aligned} [\hat{x}, \hat{p}_y] \Psi &\equiv x \frac{\partial \Psi}{\partial y} - \frac{\partial(x\Psi)}{\partial y} \\ &= x \frac{\partial \Psi}{\partial y} - x \frac{\partial \Psi}{\partial y} \\ &= 0. \end{aligned} \quad (549)$$

Position and momentum in orthogonal directions commute, and so they can be measured ‘simultaneously’ to any precision.

The commutation relations can be written in the general form

$$[\hat{\mathbf{r}}_j, \hat{\mathbf{p}}_k] = i\hbar\delta_{jk}. \quad (550)$$

Components of \mathbf{r} and \mathbf{p} in different Cartesian directions can be specified simultaneously with arbitrary precision, but components in the same direction must satisfy the usual uncertainty relation.

Our previous 1D treatment with $\Psi(\mathbf{r}, t)$ independent of y and z gives

$$\begin{aligned} \hat{p}_y\psi(x) &= 0 \\ \hat{p}_z\psi(x) &= 0, \end{aligned} \quad (551)$$

which corresponds to a particle moving in the \mathbf{i} direction, with $p_y = 0$, $p_z = 0$, and with y and z completely unknown.

Also, there is no probability flux in the $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ directions.

The kinetic energy operator, which is a scalar, is

$$\begin{aligned} \hat{T} &= \frac{1}{2m}\hat{\mathbf{p}} \cdot \hat{\mathbf{p}} \\ &= \frac{1}{2m}\hat{p}^2 \\ &= -\frac{\hbar^2}{2m}\nabla^2 \\ &= -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) \end{aligned} \quad (552)$$

so that the time-independent Schrödinger equation becomes

$$\boxed{-\frac{\hbar^2}{2m}\nabla^2\Psi(\mathbf{r}) + V(\mathbf{r})\Psi(\mathbf{r}) = E\Psi(\mathbf{r}),} \quad (553)$$

or

$$\boxed{\nabla^2 \Psi(\mathbf{r}) = \frac{2m}{\hbar^2} [V(\mathbf{r}) - E] \Psi(\mathbf{r})}. \quad (554)$$

For example, in spherical polar coordinates,

$$\begin{aligned} & \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \Psi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Psi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \Psi}{\partial \phi^2} \\ &= \frac{2m}{\hbar^2} [V(\mathbf{r}) - E] \Psi. \end{aligned} \quad (555)$$

Even in the case of a simple Coulomb potential describing the behaviour of an electron in a hydrogen atom, $V(r) = 1/r$, this equation is tedious to solve.

There are many elegant numerical routines for solving expressions of this kind, even when large numbers of particles are involved.

66 Orbital angular momentum

Classically, the orbital angular momentum about point $\mathbf{r} = \mathbf{0}$ is defined by

$$\boxed{\mathbf{L} = \mathbf{r} \times \mathbf{p} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x & y & z \\ p_x & p_y & p_z \end{vmatrix}} \quad (556)$$

and we should expect the rules governing position and linear momentum to be applicable when the system is quantised.

The quantum mechanical operators for the components of \mathbf{L} are taken to be

$$\boxed{\begin{aligned} \hat{L}_x &= \hat{y}\hat{p}_z - \hat{z}\hat{p}_y \\ \hat{L}_y &= \hat{z}\hat{p}_x - \hat{x}\hat{p}_z \\ \hat{L}_z &= \hat{x}\hat{p}_y - \hat{y}\hat{p}_x. \end{aligned}} \quad (557)$$

No additional postulates are needed to quantise \mathbf{L} because spatially orthogonal components of position and momentum commute, and there is no ambiguity in (557).

Often in going from the classical to quantum expressions ambiguities do exist, and this can be highly problematic. In some cases *ordering theorems* exist; for example, for creation and annihilation operators \hat{a} and \hat{a}^\dagger .

Given that $\hat{x}, \hat{y}, \hat{z}, \hat{p}_x, \hat{p}_y, \hat{p}_z$ are all individually Hermitian, the angular momentum operators are also Hermitian,

$$\begin{aligned}\hat{L}_x &= \hat{L}_x^\dagger \\ \hat{L}_y &= \hat{L}_y^\dagger \\ \hat{L}_z &= \hat{L}_z^\dagger,\end{aligned}\tag{558}$$

giving

$$\hat{\mathbf{L}} = \hat{\mathbf{L}}^\dagger.\tag{559}$$

Eq. (559) is an example of a vector quantum operator (like the linear momentum operator).

Angular momentum is an observable.

What are the commutation relations between the different components of the angular momentum operator?

Consider

$$\begin{aligned}[\hat{L}_x, \hat{L}_y] &= [(\hat{y}\hat{p}_z - \hat{z}\hat{p}_y), (\hat{z}\hat{p}_x - \hat{x}\hat{p}_z)] \\ &= (\hat{y}\hat{p}_z - \hat{z}\hat{p}_y)(\hat{z}\hat{p}_x - \hat{x}\hat{p}_z) - (\hat{z}\hat{p}_x - \hat{x}\hat{p}_z)(\hat{y}\hat{p}_z - \hat{z}\hat{p}_y) \\ &= \hat{y}\hat{p}_z\hat{z}\hat{p}_x - \hat{z}\hat{p}_y\hat{z}\hat{p}_x - \hat{y}\hat{p}_z\hat{x}\hat{p}_z + \hat{z}\hat{p}_y\hat{x}\hat{p}_z \\ &\quad - \hat{z}\hat{p}_x\hat{y}\hat{p}_z + \hat{x}\hat{p}_z\hat{y}\hat{p}_z + \hat{z}\hat{p}_x\hat{z}\hat{p}_y - \hat{x}\hat{p}_z\hat{z}\hat{p}_y \\ &= \hat{y}[\hat{p}_z, \hat{z}]\hat{p}_x + \hat{p}_y[\hat{z}, \hat{p}_z]\hat{x} \\ &= i\hbar(\hat{x}\hat{p}_y - \hat{y}\hat{p}_x) \\ &= i\hbar\hat{L}_z.\end{aligned}\tag{560}$$

In fact

$$\begin{aligned} [\hat{L}_x, \hat{L}_y] &= i\hbar \hat{L}_z \\ [\hat{L}_y, \hat{L}_z] &= i\hbar \hat{L}_x \\ [\hat{L}_z, \hat{L}_x] &= i\hbar \hat{L}_y. \end{aligned} \tag{561}$$

Different components of angular momentum do not commute with each other, which is unlike different components of linear momentum.

No two components can be known precisely.

Only one component of $\hat{\mathbf{L}}$ can be measured precisely, by convention usually called the z component.

Classically, the magnitude of the angular momentum is given by $|\mathbf{L}|^2 = L_x^2 + L_y^2 + L_z^2$.

The operator for the square of the total angular momentum is

$$\begin{aligned} \hat{L}^2 &= \hat{\mathbf{L}} \cdot \hat{\mathbf{L}} \\ &= \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2, \end{aligned} \tag{562}$$

which will be used in the context of calculating the expectation value of the magnitude of angular momentum:

$$\langle \hat{L}^2 \rangle = \langle \hat{L}_x^2 \rangle + \langle \hat{L}_y^2 \rangle + \langle \hat{L}_z^2 \rangle. \tag{563}$$

Does \hat{L}^2 commute with the components of $\hat{\mathbf{L}}$?

Consider for example

$$\begin{aligned} [\hat{L}^2, \hat{L}_z] &= [\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2, \hat{L}_z] \\ &= [\hat{L}_x^2, \hat{L}_z] + [\hat{L}_y^2, \hat{L}_z] + [\hat{L}_z^2, \hat{L}_z] \\ &= [\hat{L}_x^2, \hat{L}_z] + [\hat{L}_y^2, \hat{L}_z]. \end{aligned} \tag{564}$$

Now,

$$\begin{aligned}
 [\hat{L}_x^2, \hat{L}_z] &= \hat{L}_x \hat{L}_x \hat{L}_z - \hat{L}_z \hat{L}_x \hat{L}_x \\
 &= \hat{L}_x \hat{L}_x \hat{L}_z - \hat{L}_x \hat{L}_z \hat{L}_x + \hat{L}_x \hat{L}_z \hat{L}_x - \hat{L}_z \hat{L}_x \hat{L}_x \\
 &= \hat{L}_x [\hat{L}_x, \hat{L}_z] + [\hat{L}_x, \hat{L}_z] \hat{L}_x \\
 &= -i\hbar (\hat{L}_x \hat{L}_y + \hat{L}_y \hat{L}_x) ,
 \end{aligned} \tag{565}$$

and also

$$[\hat{L}_y^2, \hat{L}_z] = i\hbar (\hat{L}_x \hat{L}_y + \hat{L}_y \hat{L}_x) , \tag{566}$$

which combine to give

$$[\hat{L}^2, \hat{L}_z] = \hat{0} . \tag{567}$$

In fact, one has in general that

$[\hat{L}^2, \hat{L}_x] = \hat{0}$	(568)
$[\hat{L}^2, \hat{L}_y] = \hat{0}$	
$[\hat{L}^2, \hat{L}_z] = \hat{0} .$	

The magnitude of the total angular momentum, and any one of the components are *compatible* observables.

They have the same eigenstates.

But L_x, L_y, L_z do not commute with each other, and so only L and one of L_x, L_y, L_z can be known simultaneously.

67 Eigenvalues of the angular momentum

Given that \hat{L}^2 and \hat{L}_z are compatible observables, it is natural to enquire about their eigenvectors and eigenvalues.

It is possible to find the solution through the use of ladder operators.

Rather than having incremental steps in energy, as in the case of the harmonic oscillator, they generate the eigenstates of \hat{L}_z through incremental steps in angular momentum.

Consider the operators

$$\boxed{\begin{aligned}\hat{L}_+ &= \hat{L}_x + i\hat{L}_y \\ \hat{L}_- &= \hat{L}_x - i\hat{L}_y,\end{aligned}} \quad (569)$$

from which it follows, using (558), that

$$\boxed{\begin{aligned}\hat{L}_+ &= \hat{L}_-^\dagger \\ \hat{L}_- &= \hat{L}_+^\dagger\end{aligned}} \quad (570)$$

They are not Hermitian, and therefore not observable.

Now

$$\begin{aligned}[\hat{L}_z, \hat{L}_+] &= [\hat{L}_z, \hat{L}_x] + i[\hat{L}_z, \hat{L}_y] \\ &= \hbar(i\hat{L}_y + \hat{L}_x) \\ &= \hbar\hat{L}_+, \end{aligned} \quad (571)$$

and so

$$\hat{L}_z\hat{L}_+ = \hat{L}_+\hat{L}_z + \hbar\hat{L}_+. \quad (572)$$

Suppose that $|\phi_\alpha\rangle$ is an eigenstate of \hat{L}_z having eigenvalue $\alpha\hbar$:

$$\hat{L}_z|\phi_\alpha\rangle = \alpha\hbar|\phi_\alpha\rangle. \quad (573)$$

\hbar has the dimensions of angular momentum, and so α is a dimensionless real number.

Now

$$\begin{aligned}\hat{L}_z\hat{L}_+|\phi_\alpha\rangle &= \hat{L}_+\hat{L}_z|\phi_\alpha\rangle + \hbar\hat{L}_+|\phi_\alpha\rangle \\ &= (\alpha + 1)\hbar\hat{L}_+|\phi_\alpha\rangle.\end{aligned} \quad (574)$$

If $|\phi_\alpha\rangle$ is an eigenstate of \hat{L}_z having eigenvalue $\alpha\hbar$, then $\hat{L}_+|\phi_\alpha\rangle$ is also an eigenstate of \hat{L}_z but having eigenvalue $(\alpha + 1)\hbar$.

The z component of the angular momentum has been increased by \hbar upon applying of the operator \hat{L}_+ .

Likewise, if $|\phi_\alpha\rangle$ is an eigenstate of \hat{L}_z having eigenvalue $\alpha\hbar$, then $\hat{L}_-|\phi_\alpha\rangle$ is also an eigenstate of \hat{L}_z but having eigenvalue of $(\alpha - 1)\hbar$.

\hat{L}_+ and \hat{L}_- are ladder operators increasing/decreasing the z component of the angular momentum by \hbar .

In this way, one can construct the full set of eigenstates and eigenvalues, and therefore the measurement of a single component of angular momentum must be quantised in units of \hbar .

This is the reason why atoms and rotational transitions of molecules radiate and absorb electromagnetic energy in the form of discrete spectral lines.

Since \hat{L}^2 and L_z commute, $|\phi_\alpha\rangle$ is also an eigenstate of \hat{L}^2 having eigenvalue $\Lambda\hbar^2$, for some dimensionless real number Λ :

$$\hat{L}^2|\phi_\alpha\rangle = \Lambda\hbar^2|\phi_\alpha\rangle. \quad (575)$$

Using Eq. (568),

$$\begin{aligned} [\hat{L}^2, \hat{L}_\pm] &= [\hat{L}^2, \hat{L}_x \pm i\hat{L}_y] = 0 \\ \hat{L}^2 \hat{L}_\pm |\phi_\alpha\rangle &= \hat{L}_\pm \hat{L}^2 |\phi_\alpha\rangle \\ &= \Lambda\hbar^2 \hat{L}_\pm |\phi_\alpha\rangle. \end{aligned} \quad (576)$$

$\hat{L}_\pm|\phi_\alpha\rangle$ is an eigenstate of \hat{L}^2 with the same eigenvalue $\Lambda\hbar^2$ as $|\phi_\alpha\rangle$.

\hat{L}_\pm does not change the magnitude of the angular momentum $\Lambda \hbar^2$.

The ladder operators \hat{L}_+ and \hat{L}_- generate the eigenstates of \hat{L}_z , but all of them are eigenstates of \hat{L}^2 having the same eigenvalues $\Lambda \hbar^2$.

The z component of angular momentum changes, but the magnitude of the angular momentum does not.

Conceptually, this is reminiscent of a classical angular momentum vector of fixed magnitude, forming different angles with respect to a reference z axis (see also Fig. 54).

For the moment, let us label these states $|\phi_{\Lambda,\alpha}\rangle$.

We know that L_z must be less than L :

$$\langle \hat{L}_z^2 \rangle \leq \langle \hat{L}^2 \rangle, \quad (577)$$

and

$$\alpha^2 \leq \Lambda. \quad (578)$$

For a given value of Λ there is some maximum absolute value of α , which can have positive and negative values.

Starting with this value, the ladder operator can be applied to generate the sequence

$$\alpha = \ell, \ell - 1, \ell - 2, \dots, 2 - \ell, 1 - \ell, -\ell. \quad (579)$$

There are two possibilities:

$$\begin{aligned} & \ell, \dots, 2, 1, 0, -1, -2, \dots, -\ell \\ & \text{or} \\ & \ell \dots, \frac{3}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{3}{2}, \dots, -\ell, \end{aligned} \quad (580)$$

but it will be shown later that the second is ruled out by the symmetry of the spatial form of the orbital angular

momentum wavefunction.

The total number of eigenvalues is $2\ell + 1$, which must be an integer, and the index is denoted by m_ℓ :

$$m_\ell = \ell, \ell - 1, \ell - 2, \dots, 0, \dots, 2 - \ell, 1 - \ell, -\ell. \quad (581)$$

Note that m_ℓ includes the value 0.

Thus we have

$$\begin{aligned} \hat{L}_z |\phi_{\Lambda, m_\ell}\rangle &= m_\ell \hbar |\phi_{\Lambda, m_\ell}\rangle, \\ \hat{L}^2 |\phi_{\Lambda, m_\ell}\rangle &= \Lambda \hbar^2 |\phi_{\Lambda, m_\ell}\rangle. \end{aligned} \quad (582)$$

$m_\ell \hbar$ are the z components of the angular momentum that could be measured.

$\Lambda \hbar^2$ are the magnitudes of the angular momentum that could be measured.

For a fixed total angular momentum, the z component can be increased in units of \hbar , but eventually the series must terminate because the z component cannot exceed the magnitude.

What is the relationship between the maximum value of m_ℓ and Λ ?

The highest eigenstate of \hat{L}_z , denoted $|\phi_{\Lambda, \ell}\rangle$, has eigenvalue $\ell \hbar$, and from Eq. (582) we have that

$$\hat{L}_z |\phi_{\Lambda, \ell}\rangle = \ell \hbar |\phi_{\Lambda, \ell}\rangle. \quad (583)$$

In order to terminate the sequence we must also have

$$\hat{L}_+ |\phi_{\Lambda, \ell}\rangle = 0. \quad (584)$$

We also have

$$\hat{L}^2 |\phi_{\Lambda, \ell}\rangle = \Lambda \hbar^2 |\phi_{\Lambda, \ell}\rangle. \quad (585)$$

Consider then

$$\begin{aligned}\hat{L}_-\hat{L}_+ &= (\hat{L}_x - i\hat{L}_y)(\hat{L}_x + i\hat{L}_y) \quad (586) \\ &= \hat{L}_x^2 + \hat{L}_y^2 + i\hat{L}_x\hat{L}_y - i\hat{L}_y\hat{L}_x \\ &= \hat{L}^2 - \hat{L}_z^2 - \hbar\hat{L}_z \\ \hat{L}^2 &= \hat{L}_-\hat{L}_+ + \hbar\hat{L}_z + \hat{L}_z^2.\end{aligned}$$

Applying this to Eq. (585) yields

$$\Lambda\hbar^2 = 0 + \ell\hbar^2 + \ell^2\hbar^2, \quad (587)$$

and so

$$\boxed{\Lambda = \ell(\ell + 1)}, \quad (588)$$

and we conclude that

$$\hat{L}^2|\phi_{\Lambda,\ell}\rangle = \ell(\ell + 1)\hbar^2|\phi_{\Lambda,\ell}\rangle. \quad (589)$$

A measurement of the magnitude of the angular momentum yields a value of $\hbar\sqrt{\ell(\ell + 1)}$, and a measurement of the z component of angular momentum yields a value of $\hbar m_\ell$, where m_ℓ is an integer, ℓ is a non-negative integer, and $|m_\ell| \leq \ell$.

The situation is summarised as follows:

$$\boxed{\begin{aligned}\hat{L}^2|\ell, m_\ell\rangle &= \ell(\ell + 1)\hbar^2|\ell, m_\ell\rangle \\ \hat{L}_z|\ell, m_\ell\rangle &= m_\ell\hbar|\ell, m_\ell\rangle \\ m_\ell &\in \ell, \ell - 1, \dots, 1, 0, -1, \dots, -\ell + 1, -\ell\end{aligned}}, \quad (590)$$

where ℓ is a positive integer.

Fig. 54 illustrates the case of $\ell = 2$, corresponding to a magnitude of $\hbar\sqrt{6}$.

ℓ is called the *orbital angular momentum quantum number*.

m_ℓ is called the *azimuthal or magnetic angular momentum quantum number*.

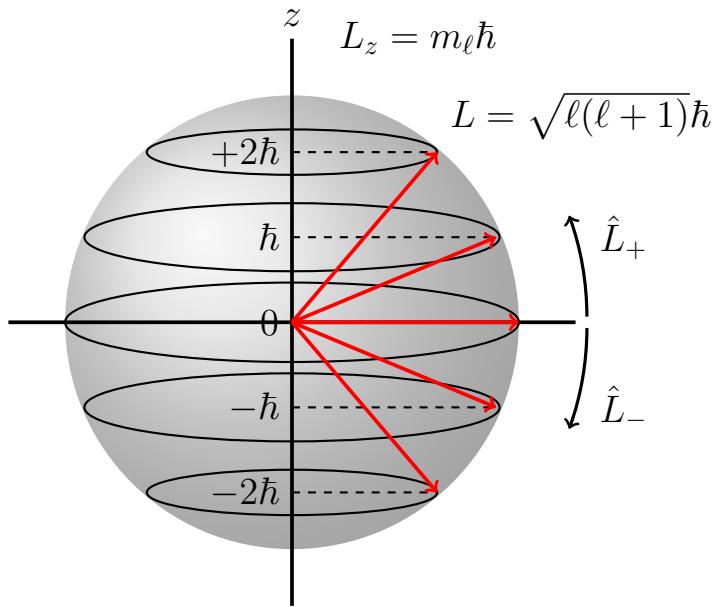


Figure 54: Quantisation of angular momentum. The ladder operators increase/decrease the z -component of angular momentum. The series terminates at the upper and lower ends when L_z becomes greater than the magnitude L by generating the zero vector. Notice that the z component of angular momentum can be zero, and the average value of the angular momentum can be zero.

While the eigenstates $|\ell, m_\ell\rangle$ of \hat{L}^2 and \hat{L}_z have precisely known values of L and L_z , the values of L_x and L_y are completely unknown, apart from the special case $L = 0$.

68 Normalising the Ladder Operators

\hat{L}_+ and \hat{L}_- raise and lower the azimuthal angular momentum, but we have not attempted to normalise the eigenstates they generate.

Introduce normalising coefficients according to

$$\begin{aligned}\hat{L}_-|\ell, m_\ell\rangle &= C_{\ell, m_\ell}|\ell, m_\ell - 1\rangle \\ \hat{L}_+|\ell, m_\ell\rangle &= D_{\ell, m_\ell}|\ell, m_\ell + 1\rangle\end{aligned}\tag{591}$$

and then

$$\begin{aligned} C_{\ell,m_\ell} C_{\ell,m_\ell}^* \langle \ell, m_\ell - 1 | \ell, m_\ell - 1 \rangle &= \langle \ell, m_\ell | \hat{L}_+ \hat{L}_- | \ell, m_\ell \rangle \\ |C_{\ell,m_\ell}|^2 &= \langle \ell, m_\ell | \hat{L}^2 + \hbar \hat{L}_z - \hat{L}_z^2 | \ell, m_\ell \rangle \\ &= \hbar^2 [\ell(\ell + 1) + m_\ell - m_\ell^2] \langle \ell, m_\ell | \ell, m_\ell \rangle \\ &= \hbar^2 [\ell(\ell + 1) - m_\ell(m_\ell - 1)], \end{aligned} \quad (592)$$

where in the first line we used $\hat{L}_-^\dagger = \hat{L}_+$, and in the second line we used a relation equivalent to Eq. (586).

We have also assumed normalised state vectors in order to yield the normalisation constants.

Therefore

$$C_{\ell,m_\ell} = \hbar \sqrt{\ell(\ell + 1) - m_\ell(m_\ell - 1)}. \quad (593)$$

Notice that $C_{\ell,m_\ell} = 0$ when $m_\ell = -\ell$, which is expected since $\hat{L}_- | \ell, -\ell \rangle = 0$.

It can similarly be shown that

$$D_{\ell,m_\ell} = \hbar \sqrt{\ell(\ell + 1) - m_\ell(m_\ell + 1)}, \quad (594)$$

which has the property $D_{\ell,m_\ell} = 0$ when $m_\ell = \ell$, which is expected since $\hat{L}_+ | \ell, \ell \rangle = 0$.

We conclude that

$\hat{L}_\pm | \ell, m_\ell \rangle = \hbar \sqrt{\ell(\ell + 1) - m_\ell(m_\ell \pm 1)} | \ell, m_\ell \pm 1 \rangle.$

(595)

69 General observations

- The angular momentum has been quantised through $\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$ without any knowledge of the nature of the system: for example the potential as a function of radius, $V(r)$.
- The basic principles apply to any rotational system, whatever its physical form.

- The discrete nature of spectral lines, and the stability of atoms, has its origins in the quantisation of angular momentum.
- The eigenvectors and eigenvalues have been deduced solely from operator methods, with no reference to the forms of the wavefunctions that described particular systems, such as an electron in ‘orbit’ around a nucleus.

The angular momentum eigenstates corresponding to $\ell = 0, 1, 2, 3, 4$, are called s, p, d, f, g states respectively.

The terms originate from the historical spectroscopic classification of atomic lines, which were called Sharp, Principal, Diffuse and Fine.

For the s state, all of the components of angular momentum are simultaneously zero. This does not conflict with the uncertainty relationship because, for example,

$$\Delta L_x \Delta L_y \geq \frac{1}{2} \left| \langle i \left[\hat{L}_x, \hat{L}_y \right] \rangle \right| = \frac{\hbar}{2} \left| \langle \hat{L}_z \rangle \right|, \quad (596)$$

and $\langle \hat{L}_z \rangle = 0$ in the s state.

This is different to the situation with the total energy, where a vacuum energy is always present.

70 Magnetic moments

Classically, an electron moving in a circular orbit having radius r and speed v has an orbital period of $T = 2\pi r/v$ and angular momentum $m_e v r$.

It constitutes a magnetic dipole: Fig. 55.

If we interpret the motion as creating a loop of area πr^2

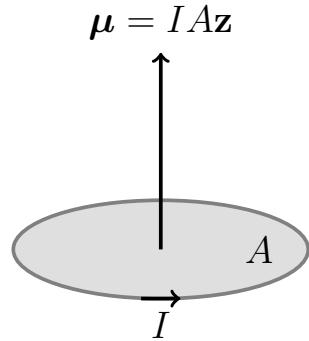


Figure 55: Classical magnetic dipole.

carrying a current $I = -e/T$, then a magnetic dipole moment of $\mu_z = IA$ is produced:

$$\mu_z = -\frac{e}{2m_e} L_z. \quad (597)$$

The ratio between the magnetic dipole moment and the angular momentum is known as the *gyromagnetic ratio*, γ , which for this case is given by

$$\gamma = -\frac{e}{2m_e}. \quad (598)$$

Since $\langle \hat{L}_z \rangle = m_\ell \hbar$, for the state $|\ell, m_\ell\rangle$, the z -component of the magnetic dipole moment of an orbiting electron becomes

$$\mu_z = -\frac{e\hbar}{2m_e} m_\ell. \quad (599)$$

The dipole moment scales inversely with the mass: the more massive the particle, the smaller the dipole moment.

This occurs, classically, because for a fixed angular momentum, a high mass means a lower angular velocity.

Be careful, two particles can have the same angular momentum, but only one of them need contribute to the overall dipole moment.

The factor

$$\mu_B = \frac{e\hbar}{2m_e}, \quad (600)$$

is called the *Bohr magneton*, and then

$$\mu_z = -\mu_B m_\ell. \quad (601)$$

The magnetic dipole moment of any system is quantised.

71 Orbital angular momentum eigenfunctions

It is possible to find the wave functions corresponding to the eigenstates $|\ell, m_\ell\rangle$ through the use of the ladder operators, similarly to the quantum harmonic oscillator.

The eigenfunctions give the probability of finding the system in some particular orientation.

The solution is most easily found through spherical polar coordinates.

The *linear momentum operator* is

$$\hat{\mathbf{p}} = -i\hbar\nabla = -i\hbar \left(\mathbf{r} \frac{\partial}{\partial r} + \boldsymbol{\theta} \frac{1}{r} \frac{\partial}{\partial \theta} + \boldsymbol{\phi} \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} \right), \quad (602)$$

where the boldface unit vectors are understood without ambiguity.

What are the angular momentum operators in spherical coordinates?

Recall that

$$\mathbf{r} = \sin \theta \cos \phi \mathbf{x} + \sin \theta \sin \phi \mathbf{y} + \cos \theta \mathbf{z}, \quad (603)$$

$$\boldsymbol{\theta} = \cos \theta \cos \phi \mathbf{x} + \cos \theta \sin \phi \mathbf{y} - \sin \theta \mathbf{z}, \quad (604)$$

$$\boldsymbol{\phi} = -\sin \phi \mathbf{x} + \cos \phi \mathbf{y}; \quad (605)$$

$$\mathbf{x} = \sin \theta \cos \phi \mathbf{r} + \cos \theta \cos \phi \boldsymbol{\theta} - \sin \phi \boldsymbol{\phi}, \quad (606)$$

$$\mathbf{y} = \sin \theta \sin \phi \mathbf{r} + \cos \theta \sin \phi \boldsymbol{\theta} + \cos \phi \boldsymbol{\phi}, \quad (607)$$

$$\mathbf{z} = \cos \theta \mathbf{r} - \sin \theta \boldsymbol{\theta}. \quad (608)$$

Then,

$$\begin{aligned} L_z &= xp_y - yp_x \\ &= r \sin \theta \cos \phi (p_r \sin \theta \sin \phi + p_\theta \cos \theta \sin \phi + p_\phi \cos \phi) \\ &\quad - r \sin \theta \sin \phi (p_r \sin \theta \cos \phi + p_\theta \cos \theta \cos \phi - p_\phi \sin \phi) \\ &= r \sin \theta p_\phi \end{aligned} \quad (609)$$

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial \phi},$$

where we used

$$\hat{p}_r = \frac{\partial}{\partial r}, \quad \hat{p}_\theta = \frac{1}{r} \frac{\partial}{\partial \theta}, \quad \hat{p}_\phi = \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi}. \quad (610)$$

Eq. (610) could also be found by inspection.

Likewise

$$L_x = yp_z - zp_y \quad (611)$$

$$\begin{aligned} &= r \sin \theta \sin \phi (p_r \cos \theta - p_\theta \sin \theta) \\ &\quad - r \cos \theta (p_r \sin \theta \sin \phi + p_\theta \cos \theta \sin \phi + p_\phi \cos \phi) \\ &= r \sin \theta \sin \phi (-p_\theta \sin \theta) \\ &\quad - r \cos \theta (p_\theta \cos \theta \sin \phi + p_\phi \cos \phi) \\ &= -r \sin \phi p_\theta - r \cos \theta \cos \phi p_\phi \end{aligned}$$

$$\hat{L}_x = i\hbar \left(\sin \phi \frac{\partial}{\partial \theta} + \cot \theta \cos \phi \frac{\partial}{\partial \phi} \right).$$

A similar expression can be derived for \hat{L}_y , giving the

set

$$\boxed{\begin{aligned}\hat{L}_x &= i\hbar \left(\sin \phi \frac{\partial}{\partial \theta} + \cot \theta \cos \phi \frac{\partial}{\partial \phi} \right) \\ \hat{L}_y &= i\hbar \left(-\cos \phi \frac{\partial}{\partial \theta} + \cot \theta \sin \phi \frac{\partial}{\partial \phi} \right) \\ \hat{L}_z &= -i\hbar \frac{\partial}{\partial \phi}.\end{aligned}} \quad (612)$$

It follows that the raising and lowering operators, $\hat{L}_{\pm} = \hat{L}_x \pm i\hat{L}_y$, are given by

$$\boxed{\hat{L}_{\pm} = \hbar e^{\pm i\phi} \left(\pm \frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right)}, \quad (613)$$

and

$$\boxed{\hat{L}^2 = -\hbar^2 \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right].} \quad (614)$$

Notice that all of the operators \hat{L}^2 , \hat{L}_x , \hat{L}_y , \hat{L}_z , \hat{L}_+ , \hat{L}_- depend only on θ and ϕ , and so the eigenfunctions will only be functions of these angles, say $Y_{\ell,m_\ell}(\theta, \phi)$.

In the case of \hat{L}_z ,

$$\hat{L}_z Y_{\ell,m_\ell}(\theta, \phi) = m_\ell \hbar Y_{\ell,m_\ell}(\theta, \phi). \quad (615)$$

Using the polar form of the operator

$$\hat{L}_z Y_{\ell,m_\ell}(\theta, \phi) = -i\hbar \frac{\partial}{\partial \phi} Y_{\ell,m_\ell}(\theta, \phi) = m_\ell \hbar Y_{\ell,m_\ell}(\theta, \phi), \quad (616)$$

we find that

$$\boxed{Y_{\ell,m_\ell}(\theta, \phi) = F_{\ell,m_\ell}(\theta) e^{im_\ell \phi}.} \quad (617)$$

A complete rotation around the z axis must leave the wave function unchanged: $Y_{\ell,m_\ell}(\theta, \phi + 2\pi) = Y_{\ell,m_\ell}(\theta, \phi)$.

$$\begin{aligned}e^{im_\ell \phi} &= e^{im_\ell (\phi + 2\pi)} \\ &= e^{im_\ell \phi} e^{im_\ell 2\pi} \\ 1 &= e^{im_\ell 2\pi}.\end{aligned} \quad (618)$$

The azimuthal quantum number m_ℓ must be an integer, and the orbital angular momentum quantum number ℓ , which is the maximum value of m_ℓ , must also be an integer.

The spatial form of the wavefunction rules out the possibility of half-integer quantum numbers.

\hat{L}_+ must return zero when it operates on the highest eigenstate of \hat{L}_z , which has $m_\ell = \ell$:

$$\hat{L}_+ |\ell, \ell\rangle = \hbar e^{i\phi} \left(\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right) Y_{\ell,\ell}(\theta, \phi) = 0. \quad (619)$$

Using Eq. (617), namely $Y_{\ell,\ell}(\theta, \phi) = F_{\ell,\ell}(\theta) e^{i\ell\phi}$,

$$\begin{aligned} \frac{\partial Y_{\ell,\ell}(\theta, \phi)}{\partial \theta} &= -i \cot \theta \frac{\partial Y_{\ell,\ell}(\theta, \phi)}{\partial \phi} \\ &= \ell \cot \theta Y_{\ell,\ell}(\theta, \phi), \end{aligned} \quad (620)$$

and one obtains the differential equation

$$\sin \theta \frac{\partial F_{\ell,\ell}(\theta)}{\partial \theta} = \ell \cos \theta F_{\ell,\ell}(\theta). \quad (621)$$

Therefore

$$F_{\ell,\ell}(\theta) = C (\sin \theta)^\ell \quad (622)$$

giving

$$Y_{\ell,\ell}(\theta, \phi) = C (\sin \theta)^\ell e^{i\ell\phi}. \quad (623)$$

The eigenfunctions of \hat{L}_z having lower eigenvalues are found using the lowering operator \hat{L}_- .

Using Eq. (613) gives

$$\begin{aligned} C_{\ell,m_\ell} Y_{\ell,m_\ell-1}(\theta, \phi) &= \hbar e^{-i\phi} \left(-\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right) Y_{\ell,m_\ell}(\theta, \phi) \\ C_{\ell,m_\ell} F_{\ell,m_\ell-1} e^{i(m_\ell-1)\phi} &= \hbar e^{-i\phi} \left(-\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right) F_{\ell,m_\ell}(\theta) e^{im_\ell\phi} \\ &= -\hbar \left(\frac{\partial}{\partial \theta} + m_\ell \cot \theta \right) F_{\ell,m_\ell}(\theta) e^{i(m_\ell-1)\phi}, \end{aligned} \quad (624)$$

and one arrives at the expression

$$F_{\ell,m_\ell-1}(\theta) = \frac{-\hbar}{C_{\ell,m_\ell}} \left(\frac{\partial F_{\ell,m_\ell}(\theta)}{\partial \theta} + m_\ell \cot \theta F_{\ell,m_\ell}(\theta) \right). \quad (625)$$

This iteration, when acting on (623), generates the eigenfunctions corresponding to lower values of m_ℓ , which must be normalised to unity when integrated over all angles:

$$\int |Y_{\ell,m_\ell}(\theta, \phi)|^2 d\Omega = \int_0^\pi \int_0^{2\pi} |Y_{\ell,m_\ell}(\theta, \phi)|^2 \sin \theta d\theta d\phi = 1. \quad (626)$$

$|Y_{\ell,m_\ell}(\theta, \phi)|^2$ is the probability density of finding the angular momentum of the system in some particular orientation.

The orthonormal functions satisfying these relationships are called *spherical harmonics*.

The first few Y_{ℓ,m_ℓ} are as follows:

$$\begin{aligned} Y_{0,0} &= \sqrt{\frac{1}{4\pi}} \\ Y_{1,0} &= \sqrt{\frac{3}{4\pi}} \cos \theta \\ Y_{1,\pm 1} &= \mp \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\phi} \\ Y_{2,0} &= \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1) \\ Y_{2,\pm 1} &= \mp \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{\pm i\phi} \\ Y_{2,\pm 2} &= \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{\pm 2i\phi}. \end{aligned} \quad (627)$$

Spherical harmonics have many properties, but two particularly useful ones are

$$\begin{aligned} Y_{\ell,-m_\ell}(\theta, \phi) &= (-1)^{m_\ell} Y_{\ell,m_\ell}(\theta, \phi) = Y_{\ell,m_\ell}(\theta, \pi - \phi), \\ Y_{\ell,m_\ell}(\pi - \theta, \pi + \phi) &= (-1)^\ell Y_{\ell,m_\ell}(\theta, \phi). \end{aligned} \quad (628)$$

72 Parity and angular momentum

In three dimensions, the parity operator inverts the wavefunction through the origin:

$$\begin{aligned}\psi(\mathbf{r}) &\xrightarrow{\hat{P}} \psi(-\mathbf{r}) \\ &\xrightarrow{\hat{P}} \psi(r, \pi - \theta, \pi + \phi).\end{aligned}\quad (629)$$

and so

$$\boxed{\hat{P}Y_{\ell,m_\ell}(\theta, \phi) = (-1)^\ell Y_{\ell,m_\ell}(\theta, \phi)} \quad (630)$$

Any eigenstate of \hat{L}^2 with eigenvalue $\ell(\ell + 1)\hbar^2$ is also an eigenstate of the parity operator \hat{P} with eigenvalue +1 or -1 depending on whether ℓ is odd or even.

73 Diatomic molecules

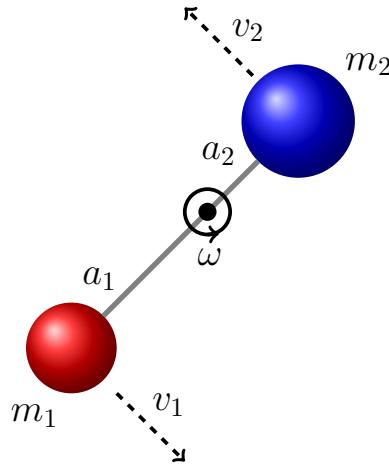


Figure 56: Diatomic molecule modelled as a rigid rotor.

So far we have considered the allowed values of angular momentum of any general system.

The associated *energy* levels depend on the specific system.

Consider the behaviour of a rigid diatomic molecule, rotating about its centre of mass: Fig. 56.

For free rotational motion, there is no potential energy.

The kinetic energy is $E = I\omega^2/2$, where I is the moment of inertia about O , or the Centre of Mass, and ω is the angular velocity.

Since $L = I\omega$,

$$E = \frac{L^2}{2I}. \quad (631)$$

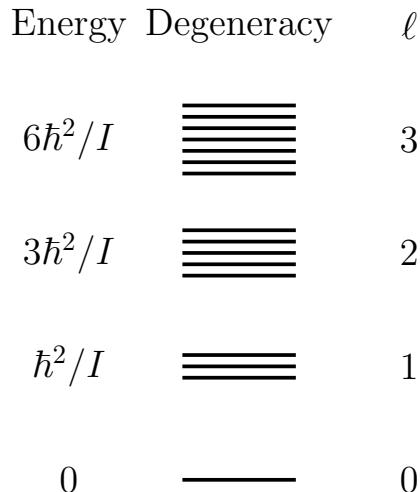


Figure 57: Rotational energy states.

Because of the known angular momentum eigenstates, $|\ell, m_\ell\rangle$, we know the allowed energy levels, which depend on ℓ .

Later, we will see that these lines can split into finer detail due to a phenomenon called *hyperfine splitting*.

The degeneracy is determined by the range of m_ℓ , illustrated in Fig. 57:

$$E_\ell = \frac{\ell(\ell+1)\hbar^2}{2I} \text{ with degeneracy } 2\ell + 1. \quad (632)$$

Consider the rotational specific heat capacity of diatomic gases.

According to Handout VII, the probability of a system, which is weakly coupled to a heat bath having temperature T , to be in the energy state n is

$$P_n = \frac{e^{-\beta E_n}}{\sum_{n=1}^{\infty} e^{-\beta E_n}}, \quad (633)$$

where $\beta = 1/k_B T$.

Taking into account the degeneracy of the states, the mean rotational energy per rotor is

$$\langle E \rangle = \frac{\sum_{\ell} \frac{\ell(\ell+1)\hbar^2}{2I} (2\ell + 1) e^{-\frac{\ell(\ell+1)\hbar^2}{2Ik_B T}}}{\sum_{\ell} (2\ell + 1) e^{-\frac{\ell(\ell+1)\hbar^2}{2Ik_B T}}}. \quad (634)$$

Unlike the vibrational case, this does not simplify easily.

Note that there is a characteristic temperature of rotation (traditionally called θ):

$$\theta_{rot} = \frac{\hbar^2}{2Ik_B}. \quad (635)$$

The rotational heat capacity of N molecules,

$$C_{rot} = N \frac{\partial \langle E \rangle}{\partial T}, \quad (636)$$

can then be derived from (634).

The heat capacity, normalised to the gas constant, is shown in Fig. 58.

Remember that the gas constant is given by $R = N_A k_B$, where N_A is the number of molecules per mole.

Notice that

- $C_{rot} \rightarrow 0$ as $T \rightarrow 0$
- $C_{rot} \rightarrow R$ as $T \rightarrow \infty$, as expected for the classical result involving two degrees of freedom

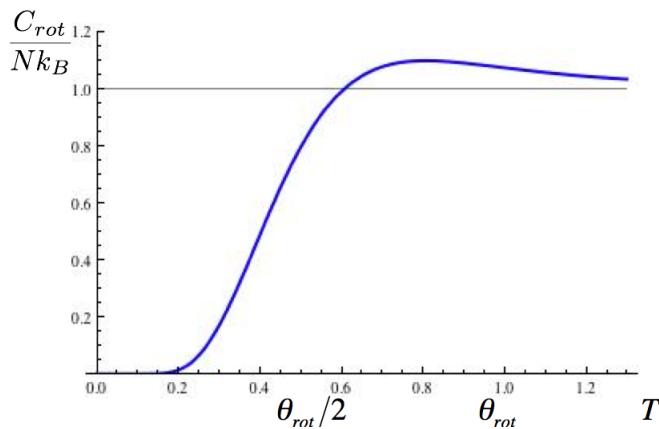


Figure 58: Rotational specific heat normalised to R , where $N = n_{\text{mol}} N_A$ is the number of molecules, and n_{mol} is the number of moles.

- the heat capacity changes rapidly at $\theta_{\text{rot}}/2$.

For example, some typical rotational temperatures are:

Molecule	H_2	HD_2	D_2	HCl	O_2
$\Theta_{\text{rot}}(K)$	85	64	43	15	2

which are much smaller than typical vibration temperatures.

Measured heat capacities of hydrogen and chlorine are shown in Fig. 59.

Notice the following:

- θ_{rot} falls as the molecule's mass, and therefore I , increases
- for molecules heavier than D_2 , θ_{rot} is always lower than the temperature above which the molecules can exist as a gas
- in this case, a rotational contribution to the heat capacity is present.

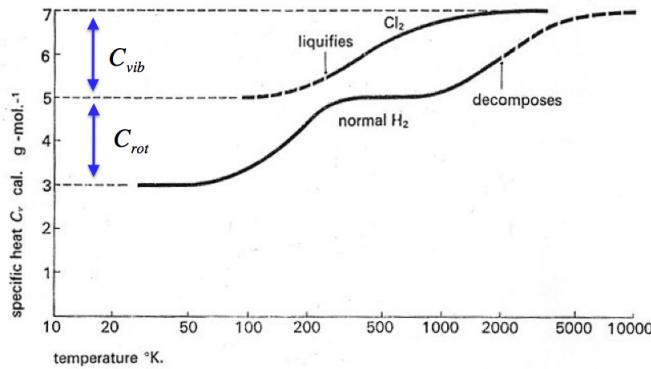


Figure 59: Total heat capacities of hydrogen and chlorine.

74 Separation of radial and angular behaviour

For a particle moving in some potential $V(\mathbf{r})$ the Hamiltonian takes the form

$$\begin{aligned}\hat{H} &= -\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \\ &= -\frac{\hbar^2}{2m} \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) \right. \\ &\quad \left. + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right] + V(\mathbf{r}).\end{aligned}\quad (637)$$

Remembering that $\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$, using (612) and after some algebra, Eq. (637) becomes

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{\hat{L}^2}{2mr^2} + V(\mathbf{r}). \quad (638)$$

For a particle moving in a *central* potential, $V(\mathbf{r}) = V(r, \theta, \phi) = V(r)$, the Hamiltonian takes the form

$$\hat{H} = -\frac{\hbar^2}{2mr^2} \frac{1}{r} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{\hat{L}^2}{2mr^2} + V(r), \quad (639)$$

and the eigenvalue problem becomes

$$-\frac{\hbar^2}{2mr^2} \frac{1}{r} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \psi}{\partial r} \right) + \frac{\hat{L}^2 \psi}{2mr^2} + V(r) \psi = E \psi. \quad (640)$$

Since the operators \hat{L}_x , \hat{L}_y and \hat{L}_z depend only on angular coordinates, they commute with \hat{r} and any function of \hat{r} , and thence they commute with the Hamiltonian in the case of a central potential.

This means that their expectation values do not change with time, and they are *conserved* quantities. (It remains however that L_x , L_y and L_z cannot be measured simultaneously.)

\hat{L}^2 , \hat{L}_z and \hat{H} must have a common set of eigenstates and their eigenvalues are well defined simultaneously.

The eigenstates $Y_{\ell,m_\ell}(\theta, \phi)$ of \hat{L}^2 and \hat{L}_z depend only on θ and ϕ , but we will see that the eigenstates of \hat{H} also depend on r for a central potential.

Eq. (640) is an eigenvalue problem with solutions of the form

$$\psi_{n,\ell,m_\ell}(r, \theta, \phi) = R_{n,\ell}(r) Y_{\ell,m_\ell}(\theta, \phi), \quad (641)$$

which introduces a radial behaviour whilst preserving the angular dependence.

The $R_{n,\ell}(r)$ are radial eigenfunctions, where n is called the *principal quantum number*.

States having the same values of n and ℓ , but different values of m_ℓ , are degenerate in energy.

Now $\hat{L}^2\psi = \ell(\ell + 1)\hbar^2\psi$, and so Eq. (640) becomes

$$\begin{aligned} -\frac{\hbar^2}{2m} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R_{n,\ell}}{\partial r} \right) + \frac{\ell(\ell + 1)\hbar^2}{2mr^2} R_{n,\ell} \\ + V(r)R_{n,\ell} = ER_{n,\ell}, \end{aligned} \quad (642)$$

where we have divided through by Y_{ℓ,m_ℓ} .

It is also convenient to make the substitution $U_{n,\ell}(r) = r R_{n,\ell}(r)$, giving

$$-\frac{\hbar^2}{2m} \frac{\partial^2 U_{n,\ell}}{\partial r^2} + \left[\frac{\ell(\ell+1)\hbar^2}{2mr^2} + V(r) \right] U_{n,\ell} = E U_{n,\ell}. \quad (643)$$

It is just a one-dimensional (radial) Shrödinger equation in r , but with an extra contribution

$$\frac{\ell(\ell+1)\hbar^2}{2mr^2} = \frac{\langle \hat{L}^2 \rangle}{2mr^2} \quad (644)$$

to the potential.

The extra term is the ‘centrifugal potential’ in the radial equation for a classical orbit.

This result is very general.

75 Hydrogen-like atom

A hydrogen-like atom consists of a single electron in ‘orbit’ around a nucleus of atomic number Z .

In this case,

$$V(r) = -\frac{Ze^2}{4\pi\epsilon_0 r}, \quad (645)$$

and the potential is a *central potential*.

From Eq. (643), we can write

$$-\frac{\hbar^2}{2m} \frac{\partial^2 U_{n,\ell}}{\partial r^2} + \left[\frac{\ell(\ell+1)\hbar^2}{2mr^2} - \frac{Ze^2}{4\pi\epsilon_0 r} \right] U_{n,\ell} = E_{n,\ell} U_{n,\ell}, \quad (646)$$

where we have explicitly indicated that the energy is quantised.

For convenience of notation, let us make the substitutions

$$A = \frac{2m}{\hbar^2} \frac{Ze^2}{4\pi\epsilon_0} \quad (647)$$

and

$$\kappa^2 = -\frac{2mE_{n,\ell}}{\hbar^2}, \quad (648)$$

which give

$$\frac{\partial^2 U_{n,\ell}}{\partial r^2} - \left[\frac{\ell(\ell+1)}{r^2} - \frac{A}{r} + \kappa^2 \right] U_{n,\ell} = 0. \quad (649)$$

Consider large r :

$$\frac{\partial^2 U_{n,\ell}}{\partial r^2} - \kappa^2 U_{n,\ell} \simeq 0, \quad (650)$$

and so one can make the informed guess

$$U_{n,\ell} \sim e^{-\kappa r}. \quad (651)$$

Consider small r :

$$\frac{\partial^2 U_{n,\ell}}{\partial r^2} - \frac{\ell(\ell+1)}{r^2} U_{n,\ell} \simeq 0, \quad (652)$$

and so one can similarly make the informed guess

$$U_{n,\ell} \sim r^{\ell+1}. \quad (653)$$

We can combine these functional forms to guess a solution of the form

$$U_{n,\ell} \propto G(r) r^{\ell+1} e^{-\kappa r}. \quad (654)$$

Substituting into Eq. (649), we obtain

$$r \frac{\partial^2 G}{\partial r^2} + 2(\ell+1-\kappa r) \frac{\partial G}{\partial r} + [A - 2\kappa(\ell+1)] G = 0, \quad (655)$$

which is the *associated Laguerre equation*.

Adopt a power series approach to solving the problem:

$$G(r) = \sum_{q=0}^{\infty} c_q r^q, \quad (656)$$

giving,

$$\begin{aligned} \sum_q c_q q(q-1)r^{q-1} + 2(\ell+1-\kappa r) \sum_q c_q qr^{q-1} \\ + [A - 2\kappa(\ell+1)] \sum_q c_q r^q = 0. \end{aligned} \quad (657)$$

This equation must be satisfied for each power of r separately, giving the recursion

$$c_{q+1} = c_q \frac{2\kappa(q+\ell+1-A/2\kappa)}{(q+1)[q+2(\ell+1)]}. \quad (658)$$

Actually, we require the series to terminate (for convergence), in the sense that $c_q = 0$ for all $q > q_{\max}$, where

$$q_{\max} + \ell + 1 - \frac{A}{2\kappa} = 0. \quad (659)$$

Therefore,

$$\frac{A}{2\kappa} = n \quad (660)$$

must be an integer.

n is called the *principal quantum number*, and $q_{\max} = n - (\ell + 1)$.

Because $q_{\max} \geq 0$,

$$\boxed{\ell \leq n - 1;} \quad (661)$$

and since $\ell \geq 0$,

$$\boxed{n \geq 1.} \quad (662)$$

The solution finally becomes

$$\boxed{R_{n,\ell} = \frac{U_{n,\ell}}{r} \propto \sum_{q=0}^{n-(\ell+1)} c_q r^{q+\ell} e^{-\kappa r}.} \quad (663)$$

Using Eqs. (647), (648) and (660) gives

$$E_n = -\frac{Z^2 e^4 m}{2(4\pi\epsilon_0)^2 \hbar^2} \frac{1}{n^2}. \quad (664)$$

The energy does not depend on ℓ , but only on n , which is why it is called the principal quantum number.

The energy levels are thus given by

$$E_n = -\frac{\hbar^2 Z^2}{2m a_0^2} \frac{1}{n^2}, \quad (665)$$

where

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{me^2} = 0.53 \times 10^{-10} \text{ m} \quad (666)$$

is the Bohr radius.

Evaluating the constants, one finds

$$E_n = -\frac{13.6Z^2}{n^2} \text{ eV}. \quad (667)$$

The energy only depends on the principal quantum number n , and not on ℓ , a degeneracy that results from the specific form of the Coulomb potential.

Eq. (664) is identical to the Bohr model encountered at the beginning of the lecture notes.

Bohr recognised the need to associate electrons with wavelike behaviour, and the need to ensure that each orbit must be an integer number of de Broglie wavelengths, but he did not appreciate the quantisation of angular momentum, or the rich structure that follows.



Niels Henrik David Bohr
(1885-1962)

76 Summary

The stationary states of a hydrogen-like atom are labeled by the principal, orbital, and magnetic quantum numbers n , ℓ and m_ℓ respectively.

They correspond to the energy, magnitude of the angular momentum, and z component of the angular momentum.

They are associated with three compatible observables \hat{H} , \hat{L}^2 and \hat{L}_z .

The allowed ranges of the quantum numbers are

$$\boxed{\begin{aligned} n &\geq 1 \\ \ell &= 0, 1, 2, \dots, n-1 \\ m_\ell &= -\ell, -\ell+1, \dots, -1, 0, 1, \dots, \ell-1, \ell. \end{aligned}} \quad (668)$$

The energy degeneracy of state n is

$$\boxed{2 \sum_{\ell=0}^{n-1} (2\ell + 1) = 2n^2,} \quad (669)$$

where the factor of 2 arises from spin, a notion that will be introduced later in the notes.

According to Eq. (641), the overall wavefunctions are given by

$$\psi_{n,\ell,m_\ell}(r, \theta, \phi) = R_{n,\ell}(r)Y_{\ell,m_\ell}(\theta, \phi), \quad (670)$$

where according to Eq. (627) the angular parts have the normalised forms

$$\begin{aligned} Y_{0,0} &= \sqrt{\frac{1}{4\pi}} \\ Y_{1,0} &= \sqrt{\frac{3}{4\pi}} \cos \theta \\ Y_{1,\pm 1} &= \mp \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\phi} \\ Y_{2,0} &= \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1) \\ Y_{2,\pm 1} &= \mp \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{\pm i\phi} \\ Y_{2,\pm 2} &= \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{\pm 2i\phi}. \end{aligned} \quad (671)$$

The radial part is given by

$$R_{n,\ell} \propto \sum_{q=0}^{n-(\ell+1)} c_q r^{q+\ell} e^{-\kappa r}, \quad (672)$$

where we see that only the $\ell = 0$ (s) state can be nonzero at the origin.

Using

$$c_{q+1} = c_q \frac{2\kappa (q + \ell + 1 - n)}{(q + 1)[q + 2(\ell + 1)]}, \quad (673)$$

we can create the first few terms:

$n = 1, \ell = 0$	$G(r) = C_0$	$R_{10} \propto e^{-\kappa r}$
$n = 2, \ell = 0$	$G(r) = C_0(1 - \kappa r)$	$R_{20} \propto (1 - \kappa r)e^{-\kappa r}$
$n = 2, \ell = 1$	$G(r) = C_0 r$	$R_{21} \propto r e^{-\kappa r}$

Note that

$$\kappa = \frac{Z}{na_0}. \quad (674)$$

To normalise the wavefunctions we must use

$$\int_0^\infty \int_0^\pi \int_0^{2\pi} |\psi(r, \theta, \phi)|^2 r^2 \sin \theta dr d\theta d\phi = 1, \quad (675)$$

but since the angular part has already been normalised, only the radial part needs considering:

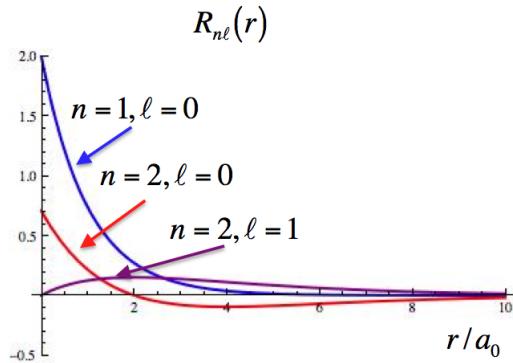
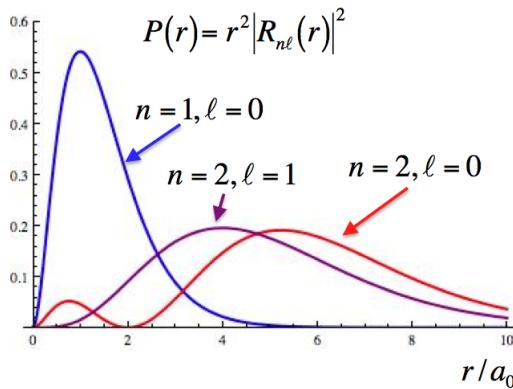
$$\int_0^\infty |R_{n,\ell}(r)|^2 r^2 dr = 1. \quad (676)$$

It is clear that $P(r)$, where

$P(r) dr = |R_{n,\ell}(r)|^2 r^2 dr,$

 (677)

is the probability of finding the electron between r and $r + dr$.

Figure 60: $R_{n,\ell}(r)$ for the $1s$, $2s$ and $2p$ states of hydrogen-like atoms.Figure 61: $P(r)$ for the $1s$, $2s$ and $2p$ states hydrogen-like atoms.

When normalised, the radial wavefunctions become

$$\boxed{\begin{aligned} R_{10}(r) &= 2 \left(\frac{Z}{a_0} \right)^{\frac{3}{2}} e^{-Zr/a_0} \\ R_{20}(r) &= 2 \left(\frac{Z}{2a_0} \right)^{\frac{3}{2}} \left(1 - \frac{Zr}{2a_0} \right) e^{-Zr/2a_0} \\ R_{21}(r) &= \sqrt{\frac{1}{3}} \left(\frac{Z}{2a_0} \right)^{\frac{3}{2}} \frac{Zr}{a_0} e^{-Zr/2a_0}, \end{aligned}} \quad (678)$$

see Figs. 60 and 61.

Consider $n = 1$:

The lowest energy state is called the $1s$ state.

$n = 1, \ell = 0, m_\ell = 0$.

$$|1s\rangle = \psi_{1,0,0} = R_{1,0}Y_{0,0} = \frac{1}{\sqrt{\pi}} \left(\frac{Z}{a_0}\right)^{\frac{3}{2}} e^{-Zr/a_0}. \quad (679)$$

The $1s$ state is spherically symmetric and is non-zero at the origin.

Consider $n = 2$:

$\ell = 0$ is the $2s$ state.

$\ell = 1, m_\ell = -1, 0, +1$ are the three $2p$ states.

Then

$$(680) \quad \begin{aligned} |2s\rangle &= \psi_{2,0,0} = R_{2,0}Y_{0,0} = \frac{1}{\sqrt{\pi}} \left(\frac{Z}{2a_0}\right)^{\frac{3}{2}} \left(1 - \frac{Zr}{2a_0}\right) e^{-Zr/2a_0} \\ |2p, 0\rangle &= \psi_{2,1,0} = R_{2,1}Y_{1,0} = \frac{1}{\sqrt{4\pi}} \left(\frac{Z}{2a_0}\right)^{\frac{3}{2}} \frac{Zr}{a_0} e^{-Zr/2a_0} \cos \theta \\ |2p, \pm 1\rangle &= \psi_{2,1,\pm 1} = R_{2,1}Y_{1,\pm 1} = \mp \frac{1}{\sqrt{8\pi}} \left(\frac{Z}{2a_0}\right)^{\frac{3}{2}} \frac{Zr}{a_0} e^{-Zr/2a_0} \sin \theta e^{\pm i\phi}. \end{aligned}$$

The $2s$ state is spherically symmetric and is non-zero at the origin.

The $2p$ states are zero at the origin.

This general model accounts for many of the essential features of the periodic table. See Mark Winter's web site, <http://winter.group.shef.ac.uk/orbitron/>, for a rather nice library of electronic shells: Fig. 62.

We have derived the wavefunctions of electrons in hydrogen-like atoms.

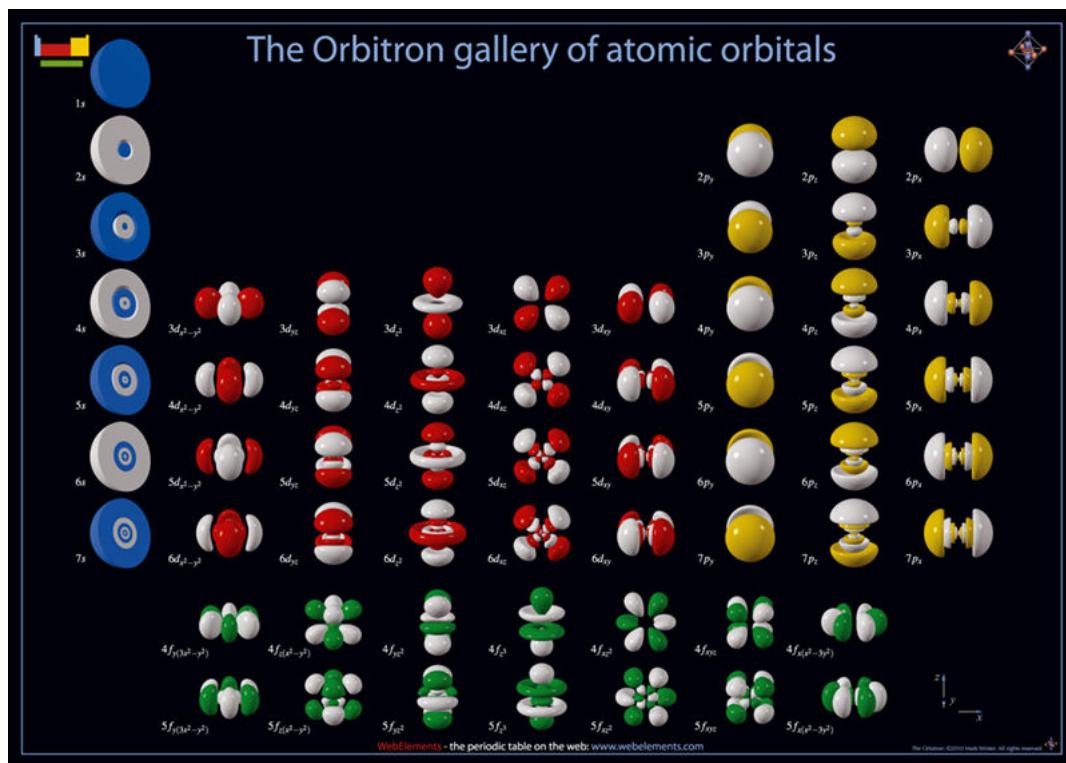


Figure 62: Electron shells: <http://winter.group.shef.ac.uk/orbitron/>.

The analysis was based on the notion of a *central potential*.

In molecules and solids, the neighbouring atoms disturb the potential, and it no longer has a radial form.

Usually the analysis requires numerical simulations, but if certain symmetries are obeyed it is possible to use the eigenstates of free atoms as the basis set for constructing the eigenstates of molecules.

This process is called *hybridisation*. We shall not consider it in this course.

X Two-Particle Systems

In the previous chapter we expanded our discussion of quantum mechanics to higher dimensions, introducing the notion of angular momentum. We defined the angular momentum operator and derived its properties, commutation relations, eigenstates and eigenvalues. We further discussed the connection between angular momentum and magnetic moment, and gave detailed examples such as diatomic molecules and the Hydrogen-like atom. This concludes the part of the lecture notes dedicated to single particle mechanics. The rest of the course will introduce multi-particle systems, starting from two particles in the present chapter. This will set the stage to discuss, in later chapters, the notion of spin and total angular momentum, as well as the role of interactions and identical particles. In this chapter, we shall see how to describe a quantum mechanical system composed of two particles, leading to the notion of entangled states and spooky action at a distance. We will see how to describe the states and their physical properties, and how to solve the two-body problem in quantum mechanics.

77 Two-particle states

A hydrogen atom is a two-particle system: the nucleus and an ‘orbiting’ electron.

To this point we have assumed that the nucleus is stationary, at $r = 0$, and that all available degrees of freedom are those associated with the movement of the electron.

Here we start to tackle the problem as a proper many-particle system with its associated states.

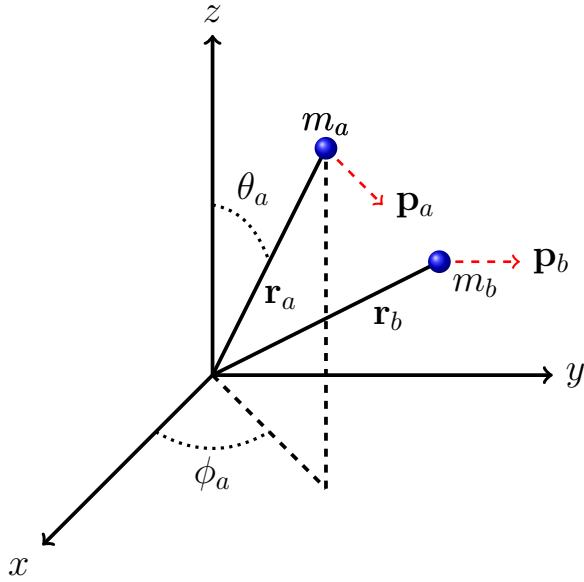


Figure 63: Coordinate system of two particles.

Fig. 63 shows two particles whose dynamical behaviour is described in terms of a spherical coordinate system.

Particle a is at \mathbf{r}_a with mass m_a and momentum \mathbf{p}_a .

Particle b is at \mathbf{r}_b with mass m_b and momentum \mathbf{p}_b .

Now we take a simple step that leads to a profound difference with respect to classical mechanics.

The overall state is written in terms of the positions of both particles:

$$|\psi_{a,b}\rangle \equiv \psi(\mathbf{r}_a, \mathbf{r}_b, t). \quad (681)$$

The probability of finding particle a in volume element $d^3\mathbf{r}_a$ at \mathbf{r}_a and particle b in volume element $d^3\mathbf{r}_b$ at \mathbf{r}_b , at time t , is

$$|\psi(\mathbf{r}_a, \mathbf{r}_b, t)|^2 d^3\mathbf{r}_a d^3\mathbf{r}_b, \quad (682)$$

which is a joint probability distribution.

$|\psi(\mathbf{r}_a, \mathbf{r}_b, t)|^2$ is the joint probability density.

This wavefunction must be normalised according to

$$\iint |\psi(\mathbf{r}_a, \mathbf{r}_b, t)|^2 d^3\mathbf{r}_a d^3\mathbf{r}_b = 1, \quad (683)$$

where both integrals extend over all space.

The probability of finding particle a in region \mathcal{A} and particle b in region \mathcal{B} is

$$P_{\mathcal{A}, \mathcal{B}} = \int_{\mathcal{A}} \int_{\mathcal{B}} |\psi(\mathbf{r}_a, \mathbf{r}_b, t)|^2 d^3\mathbf{r}_a d^3\mathbf{r}_b, \quad (684)$$

where the regions may or may not overlap.

The spatial probability density associated with particle a , regardless of the position of b , is

$$P(\mathbf{r}_a, t) = \int |\psi(\mathbf{r}_a, \mathbf{r}_b, t)|^2 d^3\mathbf{r}_b. \quad (685)$$

Likewise, the spatial probability density associated with particle b , regardless of the position of a , is

$$P(\mathbf{r}_b, t) = \int |\psi(\mathbf{r}_a, \mathbf{r}_b, t)|^2 d^3\mathbf{r}_a. \quad (686)$$

The process of integrating over a discarded variable to produce a new distribution is called *marginalisation*.

If the wavefunction is separable in the two spatial variables, namely $\psi(\mathbf{r}_a, \mathbf{r}_b, t) = \psi_a(\mathbf{r}_a)\psi_b(\mathbf{r}_b)$, then

$$P_{\mathcal{A}, \mathcal{B}} = \int_{\mathcal{A}} |\psi_a(\mathbf{r}_a, t)|^2 d^3\mathbf{r}_a \int_{\mathcal{B}} |\psi_b(\mathbf{r}_b, t)|^2 d^3\mathbf{r}_b = P_{\mathcal{A}} P_{\mathcal{B}}. \quad (687)$$

A separable wave function corresponds to a system where the two particles behave, quantum mechanically, independently of each other.

The probability of finding particle a in region \mathcal{A} and particle b in region \mathcal{B} is simply the product of the two individual probabilities.

The probability distribution $|\psi(\mathbf{r}_a, \mathbf{r}_b, t)|^2$ is associated with a state of the system that can be written as

$$|\psi_{a,b}\rangle \equiv \psi(\mathbf{r}_a, \mathbf{r}_b, t). \quad (688)$$

In the case of a separable state,

$$|\psi_{a,b}\rangle = |\psi_a\rangle|\psi_b\rangle, \quad (689)$$

where the product on the left is called the *direct product*, or *Cartesian product*.

It is still a ‘single’ ket vector, just a vector in a higher-dimensional space than each of $|\psi_a\rangle$ and $|\psi_b\rangle$.

$|\psi_{a,b}\rangle$ is normalised if each of $|\psi_a\rangle$ and $|\psi_b\rangle$ are normalised:

$$\begin{aligned} \langle\psi_{a,b}|\psi_{a,b}\rangle &= \langle\psi_b|\langle\psi_a|\psi_a\rangle|\psi_b\rangle \\ &= \langle\psi_b|\psi_b\rangle\langle\psi_a|\psi_a\rangle \\ &= 1. \end{aligned} \quad (690)$$

We require a *basis* for the composite system:

$$|\psi_{a,b}\rangle = \sum_m \sum_n c_{mn} |a_m, b_n\rangle, \quad (691)$$

where double indexing is required to accommodate all possible basis vectors.

But suppose that

$$|\psi_a\rangle = \sum_m a_m |a_m\rangle \quad (692)$$

and

$$|\psi_b\rangle = \sum_n b_n |b_n\rangle \quad (693)$$

then

$$\begin{aligned} |\psi_a\rangle|\psi_b\rangle &= \sum_m \sum_n a_m b_n |a_m\rangle|b_n\rangle \quad (694) \\ &= \sum_m \sum_n c_{mn} |a_m\rangle|b_n\rangle. \end{aligned}$$

This is an appropriate expansion when the state is separable.

If the $|a_m\rangle$ are eigenstates of some operator \hat{A} , and $|b_n\rangle$ are eigenstates of some operator \hat{B} , then each of the basis vectors $|a_m\rangle|b_n\rangle$ corresponds to some specific and allowable measurement outcome.

The fact that all possible combinations appear in the sum shows that all possible pairs of measurement results are accounted for, as required.

A crucial step is that even non-separable state vectors can be written as a weighted linear combination of separable basis vectors - this is a mathematical step, and also true of, for example, functions.

$|\psi_{a,b}\rangle$ is written as an expansion in terms of $M \times N$ basis vectors $|a_m\rangle|b_n\rangle$:

$$|\psi_{a,b}\rangle = \sum_m \sum_n c_{mn} |a_m\rangle|b_n\rangle, \quad (695)$$

where $\{|a_m\rangle, \forall m\}$ form a complete set for measurements on particle ‘a’, and $\{|b_n\rangle, \forall n\}$ form a complete set for measurements on particle ‘b’. In general they will be the eigenvectors of single-particle operators.

It is a postulate of quantum mechanics that the set $\{|a_m\rangle|b_n\rangle, \forall m, n\}$ is a complete basis to describe all possible outcomes of measurements on the system.

For example, if the states correspond to positions, then $|a_m, b_n\rangle$ is the state that corresponds to measuring particle *a* at the position labelled by *m* and *b* at the position labelled by *n*.

The possible, combined outcome of the measurement of the positions of *a* and *b* constitutes a single state in its own right.

There are $M \times N$ basis vectors *not* $M + N$.

This formulation is a ‘natural’ extension of quantum mechanics to two particles through the use of a combined probability distribution.

However, it allows one particle to ‘influence’ the behaviour of another: *spooky action at a distance* as Einstein called it, and as we shall explain below.

Because the set $\{|a_m\rangle|b_n\rangle\}$ constitutes a basis, all possible states of the system can be written in the form

$$|\psi\rangle = \sum_m \sum_n c_{mn} |a_m\rangle|b_n\rangle. \quad (696)$$

In the case where $|\psi\rangle$ is separable into the Cartesian product $|\psi_a\rangle|\psi_b\rangle$, it is possible to attribute one state solely to a and a different state solely to b .

Then the measurement of the position of particle a and the measurement of the position of particle b may be regarded as independent measurements, and the probabilities multiplied to get the probability of the composite outcome.

Suppose that we wish to find the expectation values of two operators, one relating to a and the other relating to b .

For a separable system

$$\begin{aligned}\langle \hat{O}_a \hat{O}_b \rangle &= \langle \psi_b | \langle \psi_a | \hat{O}_a \hat{O}_b | \psi_a \rangle | \psi_b \rangle \\ &= \langle \psi_a | \hat{O}_a | \psi_a \rangle \langle \psi_b | \hat{O}_b | \psi_b \rangle \\ &= \langle \hat{O}_a \rangle \langle \hat{O}_b \rangle,\end{aligned}\tag{697}$$

and the two systems do not ‘interfere’ with one another.

A related concept is the measurement of the x and y coordinates of a particle in two dimensions.

But, a state of the form

$$|\psi\rangle = C_{1,2}|a_1\rangle|b_2\rangle + C_{2,1}|a_2\rangle|b_1\rangle\tag{698}$$

is not separable, because $|\psi\rangle$ cannot be written as a *single* Cartesian product of two states. Try it!

It is not possible to attribute one state solely to a and a different state solely to b .

States which cannot be separated in this way are called *entangled*, because a measurement on one subsystem may change the outcome of a measurement on the other.

Whenever a measurement is made on one particle, and the result known, the state collapses, and the result of a measurement on the other particle becomes certain. The results appear in pairs, even though there is some uncertainty as to which pair will be seen on any given measurement.

The use of joint probability densities leads to curious possibilities: if the position of particle a is measured, and the system collapses into the appropriate eigenstate, the possible outcomes of a measurement of the position of b might instantaneously change!

This possibility is allowed for by the existence of non-separable states.

Spooky action at a distance indeed!

Show that Eq. (698) implies that once b has been measured, the result of a measurement on a is already known.

Has it changed, or do we simply know more about the likely outcome?

Was the outcome of the second experiment predetermined, or do the particles communicate in some way?

Indeed, the position of b might also be known with certainty, but this has occurred without any dynamical interaction!

This behaviour pervades all space at an instant in time, a feature of quantum mechanics that Einstein did not like!

It gave rise to the famous Einstein Podolsky Rosen (EPR) paradox.

*Which of the above interpretations is most believable?
What is now the common belief? You may like to briefly look up Bell's Theorem.*

The state (698) where $C_{12} = C_{21}$ has a special place in quantum mechanics.

In normalised form, it is written as

$$|\psi\rangle = \frac{1}{\sqrt{2}} [|a_1\rangle|b_2\rangle + |a_2\rangle|b_1\rangle]. \quad (699)$$

The probability of finding particle a in state $|a_1\rangle$ and particle b in state $|b_2\rangle$, say at positions 1 and 2 respectively, is

$$|(\langle b_2| \langle a_1|) |\psi\rangle|^2 = \frac{1}{2}. \quad (700)$$

Likewise, the probability of finding particle a in state $|a_2\rangle$ and particle b in state $|b_1\rangle$, at positions 2 and 1

respectively, is

$$|\langle b_1 | \langle a_2 |) |\psi\rangle|^2 = \frac{1}{2}. \quad (701)$$

There is maximum uncertainty about which particle is where!

A measurement can reveal the two possible outcomes with equal probabilities.

The full significance of this form will be discussed towards the end of the course.

The same reasoning applies to multiple particles, but the number of needed basis states increases rapidly.

For N subsystems

$$|\psi\rangle = \sum_{n_1, \dots, n_N} c_{n_1, \dots, n_N} |a_{n_1}\rangle \cdots |a_{n_N}\rangle. \quad (702)$$

Determining whether a particular state is separable is of crucial importance, but challenging, giving rise to concepts such as *separability criteria*.

This topic is beyond the scope of this course.

78 Two-particle observables

The observables relating to individual particles in a multiparticle system correspond to operators acting on the relevant dependence of the wavefunction.

For example,

$$\hat{\mathbf{p}}_a = -i\hbar\nabla_a = -i\hbar \left(\frac{\partial}{\partial x_a} \mathbf{i} + \frac{\partial}{\partial y_a} \mathbf{j} + \frac{\partial}{\partial z_a} \mathbf{k} \right), \quad (703)$$

and

$$\hat{\mathbf{p}}_b = -i\hbar\nabla_b = -i\hbar \left(\frac{\partial}{\partial x_b} \mathbf{i} + \frac{\partial}{\partial y_b} \mathbf{j} + \frac{\partial}{\partial z_b} \mathbf{k} \right). \quad (704)$$

The commutators are, as usual,

$$[\hat{x}_a, \hat{p}_{xa}] = [\hat{y}_a, \hat{p}_{ya}] = [\hat{z}_a, \hat{p}_{za}] = i\hbar, \quad (705)$$

and also for b .

Operations such as $\partial/\partial x_b$ have no effect on variables such as x_a , and so

$$[\hat{\mathbf{r}}_a, \hat{\mathbf{p}}_b] = \mathbf{0} = [\hat{\mathbf{r}}_b, \hat{\mathbf{p}}_a], \quad (706)$$

and

$$[\hat{\mathbf{r}}_a, \hat{\mathbf{r}}_b] = \mathbf{0} = [\hat{\mathbf{p}}_a, \hat{\mathbf{p}}_b]. \quad (707)$$

The position of a and momentum of b can be measured simultaneously and precisely, which is not to say that a measurement on a does not change the outcome of a measurement on b .

79 Interactions

In the presence of a potential, the Hamiltonian is given by

$$\hat{H} = \frac{\hat{p}_a^2}{2m_a} + \frac{\hat{p}_b^2}{2m_b} + \hat{V}(\mathbf{r}_a, \mathbf{r}_b), \quad (708)$$

where \hat{p}^2 is understood to mean $\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}$ in the sense of it being the square-magnitude of a vector.

$\hat{V}(\mathbf{r}_a, \mathbf{r}_b)$ includes external potentials, say that due to an electric field, as well as interactions between the particles.

In many cases, the interaction potential is only a function of the distance between the particles:

$$\hat{V}(\mathbf{r}_a, \mathbf{r}_b) = \hat{V}(|\mathbf{r}_b - \mathbf{r}_a|) = \hat{V}(r), \quad (709)$$

and the problem simplifies.

80 Conservation of total momentum

For a free system, namely one without an external potential, we can show that the total momentum is conserved.

The total momentum operator is

$$\hat{\mathbf{p}} = \hat{\mathbf{p}}_a + \hat{\mathbf{p}}_b . \quad (710)$$

Consider $\hat{p}_x = \hat{p}_{xa} + \hat{p}_{xb}$ and $\hat{r}_x = \hat{x}_b - \hat{x}_a$, which is the x -component of the relative separation \mathbf{r} .

Now,

$$\begin{aligned} [\hat{r}_x, \hat{p}_x] &= [\hat{x}_b - \hat{x}_a, \hat{p}_{xa} + \hat{p}_{xb}] \\ &= (\hat{x}_b - \hat{x}_a)(\hat{p}_{xa} + \hat{p}_{xb}) - (\hat{p}_{xa} + \hat{p}_{xb})(\hat{x}_b - \hat{x}_a) \\ &= i\hbar - i\hbar \\ &= 0. \end{aligned} \quad (711)$$

Each component of the total momentum $\hat{\mathbf{p}}$ commutes with each component of the *relative* position \mathbf{r} , and hence with any potential that is only a function of \mathbf{r} .

Each of $\hat{\mathbf{p}}_a$ and $\hat{\mathbf{p}}_b$ also commutes with \hat{p}_a^2 and \hat{p}_b^2 , and therefore $\hat{\mathbf{p}}$ commutes with \hat{p}_a^2 and \hat{p}_b^2 .

It follows that $\hat{\mathbf{p}}$ commutes with the total Hamiltonian given in Eq. (708).

The total momentum $\hat{\mathbf{p}}$ is a *conserved* quantity because there are no external, as distinct from internal, forces.

81 Motion of the Centre of Mass

Return to the dynamics of the two-particle system: Fig 64.

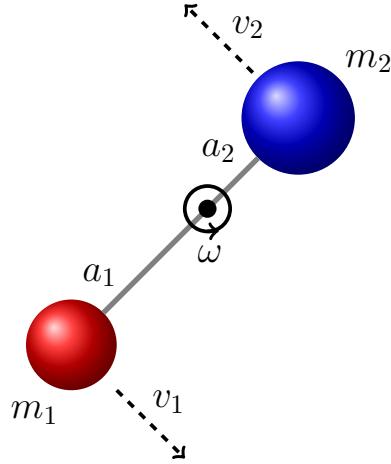


Figure 64: Diatomic molecule modelled as a rigid rotor.

The position of the *centre of mass* (CoM) is

$$\hat{\mathbf{R}} = \frac{(m_a \hat{\mathbf{r}}_a + m_b \hat{\mathbf{r}}_b)}{(m_a + m_b)}. \quad (712)$$

The commutator of $\hat{\mathbf{R}}$ with the two-particle Hamiltonian (708) becomes

$$[\hat{\mathbf{R}}, \hat{H}] = \frac{[\hat{\mathbf{r}}_a, \hat{p}_a^2] + [\hat{\mathbf{r}}_b, \hat{p}_b^2]}{2(m_a + m_b)}. \quad (713)$$

Now

$$\begin{aligned} [\hat{\mathbf{r}}_a, \hat{p}_a^2] &= [\mathbf{i}\hat{x}_a + \mathbf{j}\hat{y}_a + \mathbf{k}\hat{z}_a, \hat{p}_{xa}^2 + \hat{p}_{ya}^2 + \hat{p}_{za}^2] \quad (714) \\ &= \mathbf{i} [\hat{x}_a, \hat{p}_{xa}^2] + \mathbf{j} [\hat{y}_a, \hat{p}_{ya}^2] + \mathbf{k} [\hat{z}_a, \hat{p}_{za}^2], \end{aligned}$$

and

$$\begin{aligned} [\hat{x}_a, \hat{p}_{xa}^2] &= \hat{x}_a \hat{p}_{xa} \hat{p}_{xa} - \hat{p}_{xa} \hat{p}_{xa} \hat{x}_a \quad (715) \\ &= (i\hbar + \hat{p}_{xa} \hat{x}_a) \hat{p}_{xa} - \hat{p}_{xa} (\hat{x}_a \hat{p}_{xa} - i\hbar) \\ &= 2i\hbar \hat{p}_{xa} \\ [\hat{\mathbf{r}}_a, \hat{p}_a^2] &= 2i\hbar \hat{\mathbf{p}}_a. \end{aligned}$$

Similarly

$$[\hat{\mathbf{r}}_b, \hat{p}_b^2] = 2i\hbar \hat{\mathbf{p}}_b.$$

It follows that

$$[\hat{\mathbf{R}}, \hat{H}] = i\hbar \frac{\hat{\mathbf{p}}}{m}, \quad (716)$$

where $m = m_a + m_b$ is the total mass.

Recall that Ehrenfest's theorem states that

$$\frac{d\langle \hat{A} \rangle}{dt} = \frac{1}{\hbar} \langle i [\hat{H}, \hat{A}] \rangle, \quad (717)$$

and so the time dependence of the expectation value of $\hat{\mathbf{R}}$ is

$$\boxed{\frac{d\langle \hat{\mathbf{R}} \rangle}{dt} = \frac{\langle \hat{\mathbf{p}} \rangle}{m}.} \quad (718)$$

The velocity of the centre of mass is \mathbf{p}/m as expected.

82 Internal motion

Does a free, two-particle system, behave quantum mechanically when regarded as a single composite object?

In the absence of an external potential,

$$\hat{V}(\mathbf{r}_a, \mathbf{r}_b) = \hat{V}(|\mathbf{r}_a - \mathbf{r}_b|) = \hat{V}(r). \quad (719)$$

The Hamiltonian can be written in terms of the variables

$$\mathbf{R} = \frac{(m_a \mathbf{r}_a + m_b \mathbf{r}_b)}{(m_a + m_b)} = \mathbf{i}R_x + \mathbf{j}R_y + \mathbf{k}R_z \quad (720)$$

$$\mathbf{r} = \mathbf{r}_b - \mathbf{r}_a = \mathbf{i}r_x + \mathbf{j}r_y + \mathbf{k}r_z.$$

These can be used to describe the dynamics of the CoM, and the internal dynamics, respectively.

The x -component of the linear momentum operator for a is essentially

$$\begin{aligned} \frac{\partial}{\partial x_a} &= \frac{\partial R_x}{\partial x_a} \frac{\partial}{\partial R_x} + \frac{\partial r_x}{\partial x_a} \frac{\partial}{\partial r_x} \\ &= \frac{m_a}{m} \frac{\partial}{\partial R_x} - \frac{\partial}{\partial r_x}, \end{aligned} \quad (721)$$

where $m = m_a + m_b$ is the total mass.

It follows that

$$\hat{\mathbf{p}}_a = \frac{m_a}{m} \hat{\mathbf{p}} - \hat{\mathbf{p}}_r, \quad (722)$$

and

$$\hat{\mathbf{p}}_b = \frac{m_b}{m} \hat{\mathbf{p}} + \hat{\mathbf{p}}_r. \quad (723)$$

$\hat{\mathbf{p}}$ is the operator for the momentum associated with the movement of the centre of mass.

$\hat{\mathbf{p}}_r$ is the operator for the momentum associated with the relative motion of a and b .

Now

(724)

$$\begin{aligned} \frac{\hat{p}_a^2}{2m_a} &= \frac{m_a}{2m^2} \hat{p}^2 + \frac{1}{2m_a} \hat{p}_r^2 - \frac{1}{2m} (\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}_r + \hat{\mathbf{p}}_r \cdot \hat{\mathbf{p}}), \\ \frac{\hat{p}_b^2}{2m_b} &= \frac{m_b}{2m^2} \hat{p}^2 + \frac{1}{2m_b} \hat{p}_r^2 + \frac{1}{2m} (\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}_r + \hat{\mathbf{p}}_r \cdot \hat{\mathbf{p}}). \end{aligned}$$

The overall Hamiltonian (708) becomes

$$\hat{H} = \underbrace{\frac{\hat{p}^2}{2m}}_{\hat{H}_{CoM}} + \underbrace{\frac{\hat{p}_r^2}{2\mu} + \hat{V}(r)}_{\hat{H}_r}, \quad (725)$$

where μ is the reduced mass:

$$\frac{1}{\mu} = \frac{1}{m_a} + \frac{1}{m_b}. \quad (726)$$

The Hamiltonian separates into two parts: one associated with the movement of the CoM, and the other associated with the relative motion of the two objects, subject to the relative potential $V(r)$.

The wavefunction can be transformed into the new coordinate system: $\psi(\mathbf{r}_a, \mathbf{r}_b, t) \longrightarrow \psi(\mathbf{R}, \mathbf{r}, t)$, to give

$$\boxed{\hat{H}\psi(\mathbf{R}, \mathbf{r}) = (\hat{H}_{CoM} + \hat{H}_r)\psi(\mathbf{R}, \mathbf{r}) = E\psi(\mathbf{R}, \mathbf{r}).} \quad (727)$$

Look for a solution of the form

$$\psi(\mathbf{R}, \mathbf{r}) = U(\mathbf{R})u(\mathbf{r}). \quad (728)$$

Substituting gives

$$\frac{1}{U(\mathbf{R})}\hat{H}_{CoM}U(\mathbf{R}) + \frac{1}{u(\mathbf{r})}\hat{H}_ru(\mathbf{r}) = E. \quad (729)$$

The first term is a function of \mathbf{R} only, and the second term is a function of \mathbf{r} only.

Therefore, when $\hat{V}(\mathbf{r}_a, \mathbf{r}_b) = \hat{V}(|\mathbf{r}_b - \mathbf{r}_a|) = \hat{V}(r)$, the Schrödinger equation is separable with respect to the CoM and the relative motion:

$$-\hbar^2 \frac{\nabla_R^2}{2m} U(\mathbf{R}) = E_{CoM}U(\mathbf{R}) \quad (730)$$

and

$$\left(-\hbar^2 \frac{\nabla_r^2}{2\mu} + V(r) \right) u(r) = E_r u(r), \quad (731)$$

respectively.

The total energy is given by $E_{CoM} + E_r$.

The overall wavefunction has the form

$$\boxed{\psi(\mathbf{r}_a, \mathbf{r}_b) = \psi(\mathbf{R}, \mathbf{r}) = e^{i\hat{\mathbf{p}} \cdot \hat{\mathbf{R}}/\hbar} \psi_{n,\ell,m_\ell}(\mathbf{r})}. \quad (732)$$

The following are observed:

- According to Eq. (730), the whole structure moves, as referenced by the CoM, as a free particle having mass $m = m_a + m_b$.

- This part of the wave function takes the form of a free particle, as if the composite structure is a single particle.
- According to Eq. (731), the relative motion is equivalent to a single particle of mass μ moving in the potential $V(r)$.
- The eigenstates and eigenvalues of (731) are known for different potentials.
- The internal motion of the two-particle system has the same energy and angular momentum eigenstates as a single particle in a central potential $V(r)$, with mass μ .

The quantum mechanical description of a diatomic molecule has the following features:

- Its translational behaviour is the same as that of a single composite object.
- Its vibrational motion is the same as that of a single mass moving in a central potential.
- Its rotational behaviour is the same as that of a single mass acting as a rigid rotor at a fixed distance from the origin.

These observations underpin the spectroscopic analysis of diatomic gases.

XI Spin

In the previous chapter, we began our study of multi-particle systems in quantum mechanics. Specifically, we considered a system composed of two particles, and how one can describe the states and their physical properties, and how to solve the two-body problem. This lead to the important concept of entangled states. In this chapter, we shall introduce another key and exquisitely quantum mechanical notion: the intrinsic angular momentum, or spin, of a particle. Using concepts developed in the previous chapter, we shall see how one can combine the spin and orbital angular momentum into the total angular momentum, and how this can be formulated in terms of eigenstates and eigenvalues, and composite wave functions. We will also study how to combine spins of different particles.

83 Stern-Gerlach Experiment

In 1922 Otto Stern and Walther Gerlach devised an experiment to investigate whether angular momentum is quantised.

The basic idea was to pass a beam of neutral atoms through the poles of a long magnet: Fig. 65.

The poles were shaped so that a non-uniform magnetic field was produced between the pole pieces in the direction normal to the beam (and uniform along the direction of the beam).



Plaque at the Frankfurt Institute

Assume that the field has only a z -component, and that it varies in the z direction, normal to the beam: $\mathbf{B} = B_z(z) \mathbf{z}$.

The potential energy of a magnetic dipole $\boldsymbol{\mu}$ in the field is given by

$$U(\mathbf{r}) = -\boldsymbol{\mu} \cdot \mathbf{B}(\mathbf{r}) = -\mu_z B_z(z). \quad (733)$$

The dipole experiences a translation force as it attempts to minimise the potential energy:

$$\begin{aligned} \mathbf{F}(\mathbf{r}) &= -\nabla U(\mathbf{r}) \\ F_z(z) &= \mu_z \frac{\partial B_z(z)}{\partial z}. \end{aligned} \quad (734)$$

If the spatial gradient is constant, the dipole experiences a lateral force that is proportional to the z -component of the dipole moment.

When an atom passes through the magnet, it will travel, because of the translational force, in a direction that depends on the z -component of its dipole moment.

By recording the point of arrival on a distant screen, one Cartesian component of the angular momentum can be measured: assuming of course that the magnetic dipole moment is a good indicator of angular momentum.

Classically, a continuous range of outcomes is possible, whereas quantum mechanically, only a limited number of discrete momenta are possible.

Also, the result should not depend on the orientation of the magnetic field.

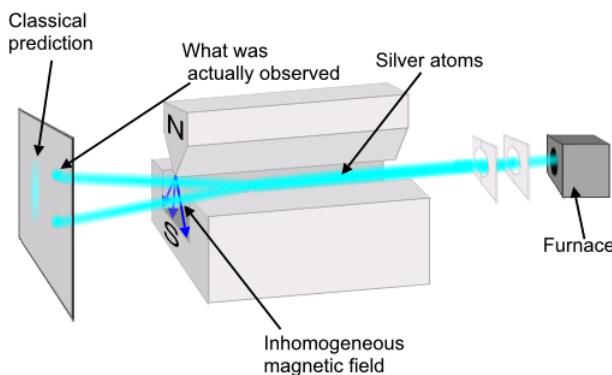
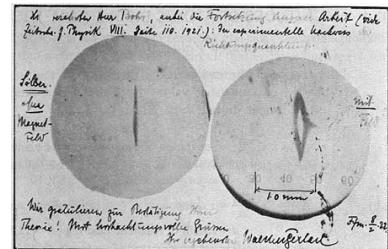


Figure 65: The Stern-Gerlach experiment of 1922.

Stern sent a card to Bohr saying: “The experimental proof of directional quantisation. Silver without magnetic field / with magnetic field. We congratulate you on the confirmation of your theory.”



Postcard from Stern to Bohr (1922).

One might speculate as to when the wavefunction collapses. When does the ‘experiment’ take place?

However, there was a problem that was not resolved until 1925 by Uhlenbeck and Goudsmit when studying the line spectrum of hydrogen.

They found that some of the lines were split into double lines, even though this was not predicted by the quantisation of orbital angular momentum.

Quantum mechanically, orbital angular momentum $\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$ is quantised, and one would expect the beam to be split into the $2\ell + 1$ possible values of m_ℓ .

ℓ is an integer and so an *odd* number of beams is always expected, but for some atoms, the experiment produced an even number of beams!

Because $2\ell + 1$ must be an even number, ℓ must be able to take on half-integer values, which as we have seen cannot be attributed to orbital angular momentum.

The non-orbital angular momentum is called *spin* and is an intrinsic property of fundamental particles, such as electrons.

Spin is not attributable to a particle's motion through space.

In any case, it would be quantised in units of \hbar , not $\hbar/2$.

The Stern-Gerlach experiment and atomic line spectra require additional angular momenta of $\pm \hbar/2$ to be present.

84 Spin operators

How does quantum mechanics describe spin?

Clearly it is measurable and so must be associated with a Hermitian operator; also it has certain qualities of angular momentum, such as a magnetic dipole moment.

Denote the spin angular momentum operator

$$\hat{\mathbf{S}} = \mathbf{i}\hat{S}_x + \mathbf{j}\hat{S}_y + \mathbf{k}\hat{S}_z. \quad (735)$$

Assume that the spin components obey the same commutation relations as the orbital angular momentum operators (see Handout IX):

$[\hat{S}_x, \hat{S}_y] = i\hbar\hat{S}_z$	(736)
$[\hat{S}_y, \hat{S}_z] = i\hbar\hat{S}_x$	
$[\hat{S}_z, \hat{S}_x] = i\hbar\hat{S}_y.$	

Following the same reasoning as for the orbital angular momentum, we can assert the following:

- \hat{S}^2 and \hat{S}_z are compatible observables.
- The eigenvalues of \hat{S}_z are $m_s\hbar$.
- The eigenvalues of \hat{S}^2 are $s(s+1)\hbar^2$.
- Two sequences are possible, one based on integer values of m_s , and one based on half-integer values of m_s .
- m_s is called the *spin magnetic quantum number*.
- $m_s = -s, -s+1, \dots, s-1, s$.
- The $(2s+1)$ possible values can be odd or even.

For hydrogen, with a single electron, and $\ell = 0$, the Stern-Gerlach experiment produces two beams.

The proton also has spin, but its magnetic dipole moment is small compared with that of the electron.

So for an electron $2s+1=2$, and $s=1/2$.

The electron is a *spin-half particle*, meaning that the only allowed values of any one component of its spin angular momentum are $\pm\hbar/2$.

The magnitude of the electron's total spin angular momentum is

$$\boxed{\sqrt{s(s+1)}\hbar = \sqrt{\frac{1}{2}\left(\frac{1}{2}+1\right)}\hbar = \frac{\sqrt{3}}{2}\hbar.} \quad (737)$$

An electron with $m_s = +1/2$ is called *spin up*, whereas an electron with $m_s = -1/2$ is called *spin down*.

In each case, the direction of the total spin lies at a direction of 55° to the z -axis: Fig. 66.

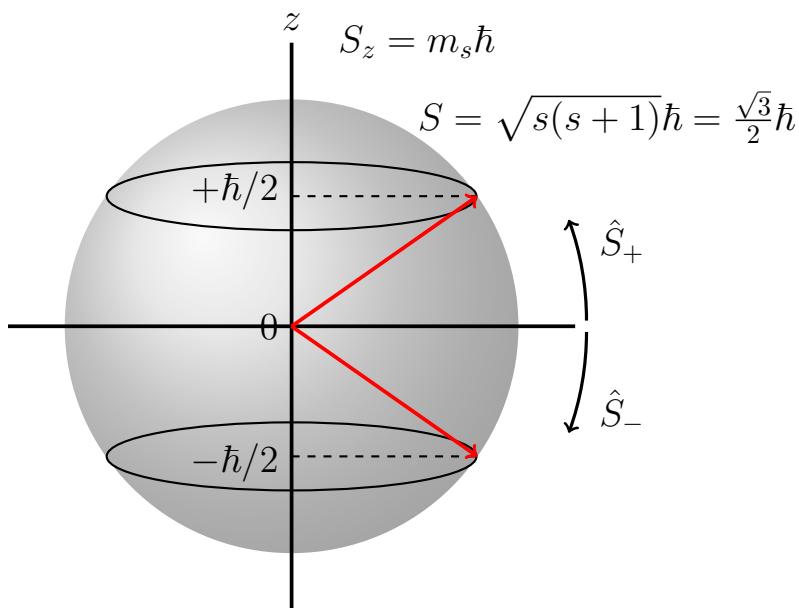


Figure 66: Spin of a single electron.

The x and y components cannot be determined precisely once the z component has been measured.

Where does spin come from? After all, there can be no spatial wave function describing spin, because spatial components would be needed, and the results are not compatible with the quantisation of orbital angular momentum.

Spin arises from a relativistic wave equation, which we shall not consider here.

The *Dirac equation*, which can be regarded as a relativistic version of the Schrödinger equation, shows that spin can be included in the quantum mechanical description of a particle by simply combining the additional degrees of freedom with the usual degrees of freedom associated with space:

$$|\psi\rangle = |\mathbf{r}, t\rangle |S\rangle. \quad (738)$$

$|S\rangle$ does not depend on spatial coordinates. It is representative of other (intrinsic) degrees of freedom present in the particles themselves.

The spin operators \hat{S}^2 and \hat{S}_z do not involve spatial coordinates, and so only act on $|S\rangle$, not on $|\mathbf{r}, t\rangle$.

Conversely, \hat{L}^2 and \hat{L}_z only act on $|\mathbf{r}, t\rangle$, not on $|S\rangle$.

Although there are many different notations for the *spin up* and *spin down* states, we shall use ‘arrows’:

$$\boxed{\begin{aligned}\hat{S}_z |\uparrow\rangle &= +\frac{\hbar}{2} |\uparrow\rangle \\ \hat{S}_z |\downarrow\rangle &= -\frac{\hbar}{2} |\downarrow\rangle.\end{aligned}} \quad (739)$$

Any spin state of the system can be written as a linear superposition of the basis states:

$$|\chi\rangle = c_1 |\uparrow\rangle + c_2 |\downarrow\rangle, \quad (740)$$

where $|c_1|^2 + |c_2|^2 = 1$.

In general, the outcome of a measurement is uncertain, and it depends on the relative values of c_1 and c_2 .

85 Anticommutation

There is a neat way of ensuring that, while the commutation relations in Eq. (736) are satisfied, \hat{S}_z is limited to having eigenvalues $\pm\hbar/2$.

In order to do this, let us define the *anticommutator* of any two operators

$$\boxed{\{\hat{A}\hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}}, \quad (741)$$

and then for an electron let us postulate that

$$\boxed{\{\hat{S}_i, \hat{S}_j\} = \frac{\hbar^2}{2} \delta_{ij}}, \quad (742)$$

where i, j can each stand for any of the Cartesian components.

From Eq. (742) we see that

$$\hat{S}_z^2 = \frac{\hbar^2}{4}. \quad (743)$$

The solution to the eigenvalue problem is clearly

$$\begin{aligned} \hat{S}_z |\uparrow\rangle &= +\frac{\hbar}{2} |\uparrow\rangle \\ \hat{S}_z |\downarrow\rangle &= -\frac{\hbar}{2} |\downarrow\rangle, \end{aligned} \quad (744)$$

as required.

For an electron, we therefore have

$$\begin{aligned} s &= \frac{1}{2} \\ m_s &= \pm \frac{1}{2}. \end{aligned} \quad (745)$$

86 Spin raising and lowering

There are also spin raising and lowering operators,

$$\hat{S}_{\pm} = \hat{S}_x \pm i\hat{S}_y, \quad (746)$$

which have the effects of

$$\boxed{\begin{aligned} \hat{S}_+ |s, m_s\rangle &= \hbar \sqrt{s(s+1) - m_s(m_s+1)} |s, m_s+1\rangle \\ \hat{S}_- |s, m_s\rangle &= \hbar \sqrt{s(s+1) - m_s(m_s-1)} |s, m_s-1\rangle. \end{aligned}}$$

They have the following effects on the basis states:

$$\boxed{\begin{aligned} \hat{S}_+ |\uparrow\rangle &= |0\rangle \\ \hat{S}_+ |\downarrow\rangle &= \hbar |\uparrow\rangle \\ \hat{S}_- |\uparrow\rangle &= \hbar |\downarrow\rangle \\ \hat{S}_- |\downarrow\rangle &= |0\rangle. \end{aligned}} \quad (747)$$

Note that the spin angular momentum changes in increments of \hbar : Fig. 67

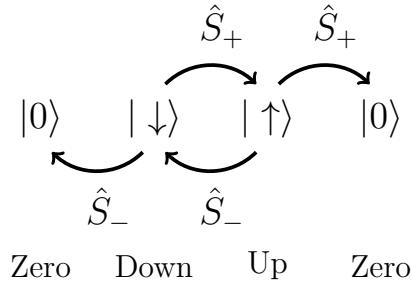


Figure 67: The raising operator \hat{S}_+ increases the spin state by \hbar . The lowering operator \hat{S}_- decreases the spin state by \hbar . The series terminates at the bottom and top of the ladder by generating the zero vector $|0\rangle$.

What are the eigenstates of \hat{S}_x and \hat{S}_y ?

Using

$$\hat{S}_x = \frac{\hat{S}_+ + \hat{S}_-}{2}, \quad (748)$$

and Eq. (747) gives

$$\hat{S}_x |\uparrow\rangle = \frac{\hbar}{2} |\downarrow\rangle \quad \text{and} \quad \hat{S}_x |\downarrow\rangle = \frac{\hbar}{2} |\uparrow\rangle. \quad (749)$$

It follows that

$$\begin{aligned} \hat{S}_x (|\uparrow\rangle + |\downarrow\rangle) &= \frac{\hbar}{2} (|\downarrow\rangle + |\uparrow\rangle) = +\frac{\hbar}{2} (|\uparrow\rangle + |\downarrow\rangle) \\ \hat{S}_x (|\uparrow\rangle - |\downarrow\rangle) &= \frac{\hbar}{2} (|\downarrow\rangle - |\uparrow\rangle) = -\frac{\hbar}{2} (|\uparrow\rangle - |\downarrow\rangle). \end{aligned} \quad (750)$$

The eigenstates of \hat{S}_x , say $|\chi_+\rangle$ and $|\chi_-\rangle$, can be written in terms of the eigenstates of \hat{S}_z , which are $|\uparrow\rangle$ and $|\downarrow\rangle$:

$$\begin{aligned} |\chi_+\rangle &= \frac{1}{\sqrt{2}} (|\uparrow\rangle + |\downarrow\rangle) \\ |\chi_-\rangle &= \frac{1}{\sqrt{2}} (|\uparrow\rangle - |\downarrow\rangle), \end{aligned} \quad (751)$$

and

$$\begin{aligned} \hat{S}_x |\chi_+\rangle &= +\frac{\hbar}{2} |\chi_+\rangle \\ \hat{S}_x |\chi_-\rangle &= -\frac{\hbar}{2} |\chi_-\rangle. \end{aligned} \quad (752)$$

Two possible values of the spin angular momentum can result from a measurement of \hat{S}_x : $\pm\hbar/2$, as expected.

Similarly,

$$\hat{S}_y = \frac{\hat{S}_+ - \hat{S}_-}{2i}, \quad (753)$$

where the eigenstates, say $|\phi_+\rangle$ and $|\phi_-\rangle$, are given by

$$\begin{aligned} |\phi_+\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\rangle + i|\downarrow\rangle) \\ |\phi_-\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\rangle - i|\downarrow\rangle), \end{aligned} \quad (754)$$

and

$$\begin{aligned} \hat{S}_y|\phi_+\rangle &= +\frac{\hbar}{2}|\phi_+\rangle \\ \hat{S}_y|\phi_-\rangle &= -\frac{\hbar}{2}|\phi_-\rangle. \end{aligned} \quad (755)$$

87 Succession of Stern-Gerlach experiments

Notice the peculiar behaviour of quantum mechanics and non-commuting variables highlighted by carrying out a succession of Stern-Gerlach experiments.

Let \hat{S}_x refer to an experiment with field gradient along the x direction, such that it measures the x -component of the spin.

Let \hat{S}_z refer to an experiment with field gradient along the z direction, such that it measures the z -component of the spin.

If the incident beam is not in any particular spin state, two beams emerge having eigenvalues $\pm\hbar/2$.

Suppose that these two outcomes correspond to the eigenvectors of an x -directed spin measurement: $|\chi_+\rangle$ and $|\chi_-\rangle$.

The $|\chi_+\rangle$ beam is passed into a second experiment with a z -directed field gradient.

The possible measurements are $\pm\hbar/2$, and they occur with equal probability.

The two emerging beams are spin up and spin down.

The spin down beam is passed into a third experiment with an x -directed field gradient.

Because the spin down state is a superposition of the $|\chi_+\rangle$ and $|\chi_-\rangle$ states, both of these beams emerge from the last experiment.

Thus a perfect $|\chi_+\rangle$ state ends up generating both $|\chi_+\rangle$ and $|\chi_-\rangle$ states after being subjected to a sequence of measurements.

This behaviour is seen experimentally.

88 Uncertainty

Suppose that a system is in the spin up eigenstate. Then

$$\hat{S}_z |\uparrow\rangle = \frac{\hbar}{2} |\uparrow\rangle \text{ and } \hat{S}_z^2 |\uparrow\rangle = \frac{\hbar^2}{4} |\uparrow\rangle, \quad (756)$$

and therefore

$$\langle \hat{S}_z \rangle = \frac{\hbar}{2} \text{ and } \langle \hat{S}_z^2 \rangle = \frac{\hbar^2}{4}, \quad (757)$$

giving

$$\Delta S_z = \sqrt{\langle \hat{S}_z^2 \rangle - \langle \hat{S}_z \rangle^2} = 0. \quad (758)$$

There is no uncertainty in the measurement of the z component of the spin angular momentum when the system is in either the spin-up or spin-down state.

Now measure \hat{S}_x when the system is in the spin up state.

$$\hat{S}_x |\uparrow\rangle = \frac{\hbar}{2} |\downarrow\rangle \text{ and } \hat{S}_x^2 |\uparrow\rangle = \frac{\hbar^2}{4} |\uparrow\rangle, \quad (759)$$

giving

$$\langle \hat{S}_x \rangle = 0 \text{ and } \langle \hat{S}_x^2 \rangle = \frac{\hbar^2}{4}, \quad (760)$$

and therefore

$$\Delta S_x = \sqrt{\langle \hat{S}_x^2 \rangle - \langle \hat{S}_x \rangle^2} = \frac{\hbar}{2}. \quad (761)$$

It can also be shown that

$$\Delta S_y = \frac{\hbar}{2}, \quad (762)$$

and therefore the x and y components of the spin angular momentum are completely undetermined. There is no information about its direction in the $x - y$ plane.

89 Combining spin and orbital angular momentum

How do we combine spin and orbital angular momentum?

Both spin, \mathbf{S} , and orbital angular momentum, \mathbf{L} , contribute to the total angular momentum, \mathbf{J} .

‘Classically’ we have $\mathbf{J} = \mathbf{L} + \mathbf{S}$, and so we should write

$$\boxed{\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}.} \quad (763)$$

The associated state vectors have the form

$$|\psi\rangle = |\mathbf{r}\rangle |S\rangle. \quad (764)$$

$\hat{L}_x, \hat{L}_y, \hat{L}_z$ act only on $|\mathbf{r}\rangle$, and $\hat{S}_x, \hat{S}_y, \hat{S}_z$ act only on $|S\rangle$.

All components of $\hat{\mathbf{L}}$ commute with all components of $\hat{\mathbf{S}}$.

It follows that

$$\begin{aligned}
 [\hat{J}_x, \hat{J}_y] &= [\hat{L}_x + \hat{S}_x, \hat{L}_y + \hat{S}_y] \\
 &= (\hat{L}_x + \hat{S}_x)(\hat{L}_y + \hat{S}_y) - (\hat{L}_y + \hat{S}_y)(\hat{L}_x + \hat{S}_x) \\
 &= [\hat{L}_x, \hat{L}_y] + [\hat{S}_x, \hat{S}_y] \\
 &= i\hbar(\hat{L}_z + \hat{S}_z) \\
 &= i\hbar\hat{J}_z.
 \end{aligned} \tag{765}$$

Similar results can be derived for the other components giving

$[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z$	
$[\hat{J}_y, \hat{J}_z] = i\hbar\hat{J}_x$	
$[\hat{J}_z, \hat{J}_x] = i\hbar\hat{J}_y.$	

(766)

Regardless of the constituents of $\hat{\mathbf{J}}$, we would, on the basis of these commutators, come to the following conclusions:

- \hat{J}^2 and \hat{J}_z are compatible observables.
 - The eigenvalues of \hat{J}_z are $m_j\hbar$.
 - The eigenvalues of \hat{J}^2 are $j(j+1)\hbar^2$.
 - m_j takes one of the $2j+1$ values $-j, -j+1, \dots, j-1, j$.
 - j can take integer or half integer values.
-

But how does this indexing relate to the indexing associated with the contributory orbital and spin parts (Fig.68)?

ℓ, m_ℓ are generated through ladder operators, s, m_s are generated through ladder operators, and j, m_j are generated through ladder operators.

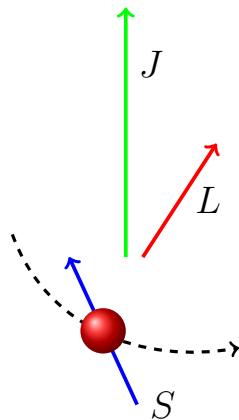


Figure 68: A particle has orbital angular momentum (red) and spin angular momentum (blue), how do we calculate the quantised behaviour of the total angular momentum (green)?

How do ℓ, m_ℓ, s, m_s relate to j, m_j ?

The quantisation conditions for \hat{J}_z are constrained by the quantisation conditions for \hat{L}_z and \hat{S}_z : Fig. 69.

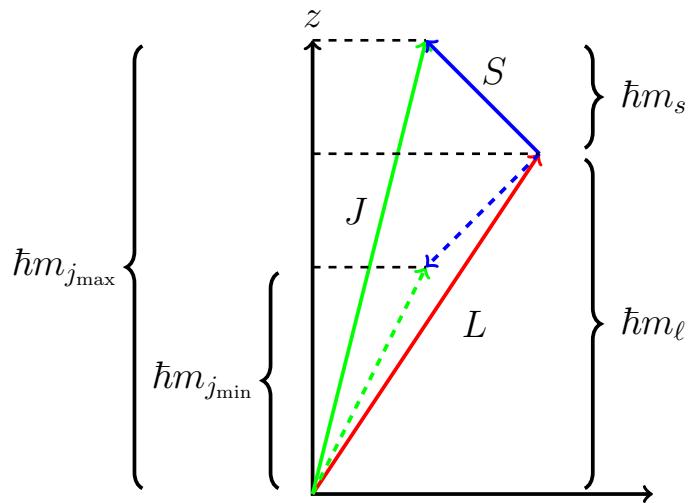


Figure 69: Vector addition of orbital and spin angular momentum.

The procedure for combining angular momentum quantum numbers can be found by algebraic means, but the best insight comes from a simple diagram: Fig. 70.

Since $\hat{J}_z = \hat{L}_z + \hat{S}_z$, then $m_j = m_\ell + m_s$ covers all possibilities.

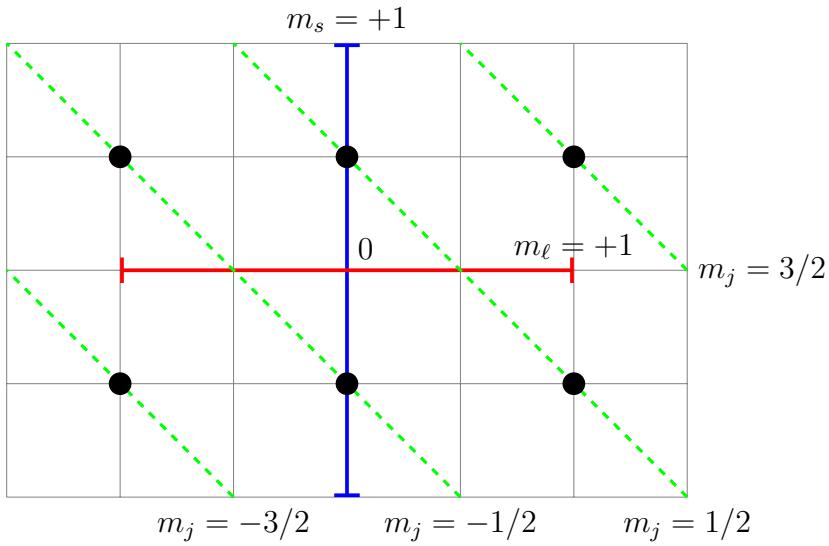


Figure 70: Combining the orbital (red) and spin (blue) ladder operators to create the total (green) ladder operator. The $\ell = 1$ and $s = 1/2$ combinations are shown.

The maximum possible value for m_j is $m_{\ell_{\max}} + m_{s_{\max}}$, and therefore the maximum possible value for j is $\ell + s$.

The maximum possible value of J^2 occurs when **L** and **S** are ‘maximally parallel’: Fig. 69 (green arrow).

For $j = j_{\max} = \ell + s$, m_j can take the following values

$$m_j = -(\ell+s), -(\ell+s)+1, \dots, (\ell+s)-1, (\ell+s). \quad (767)$$

The minimum possible value of J^2 occurs when **L** and **S** are ‘maximally antiparallel’: Fig. 69 (dashed green arrow), and so $j_{\min} = |\ell - s|$.

For $j = j_{\min} = |\ell - s|$, m_j can take the following values:

$$m_j = -|\ell - s|, -|\ell - s| + 1, \dots, |\ell - s| - 1, |\ell - s|. \quad (768)$$

The allowed values for the total angular momentum operator j are

$$j = (\ell + s), (\ell + s) - 1, \dots, |\ell - s| + 1, |\ell - s|. \quad (769)$$

There are $2s + 1$ or $2\ell + 1$ such values of j , whichever is fewer.

For example:

$$\ell = 3, s = 1 \left\{ \begin{array}{ll} j = 4 & m_j = -4, -3, -2, -1, 0, 1, 2, 3, 4 \\ j = 3 & m_j = -3, -2, -1, 0, 1, 2, 3 \\ j = 2 & m_j = -2, -1, 0, 1, 2. \end{array} \right. \quad (770)$$

Taking $s \leq \ell$, the total number of states is

$$\sum_{\ell-s}^{\ell+s} (2j+1) = (2\ell+1)(2s+1). \quad (771)$$

For a spin 1/2 system, j takes only two values,

$$j = \ell + \frac{1}{2}, \left| \ell - \frac{1}{2} \right|, \quad (772)$$

and the total number of states is

$$2(2\ell+1). \quad (773)$$

90 Composite wavefunction

The total state describing orbital angular momentum and spin must have the form

$$|\Psi\rangle = |\mathbf{r}\rangle |S\rangle. \quad (774)$$

The radial dependence does not enter into the orbital angular momentum wavefunction, and so

$$|\Psi\rangle = Y_{\ell,m_\ell}(\theta, \phi) |S\rangle. \quad (775)$$

Each possible spin state has its associated wave function based on azimuthal coordinates.

For states corresponding to a specific j and m_j , denoted Ψ_{j,m_j} , the overall angular momentum wavefunctions are linear combinations of the states having the allowable values of m_ℓ and m_s , such that $m_j = m_\ell + m_s$:

$$\boxed{\Psi_{j,m_j} = \sum_{\ell,m_\ell,s,m_s} C_{\ell,m_\ell,s,m_s} Y_{\ell,m_\ell}(\theta, \phi) |s, m_s\rangle. \quad (776)}$$

The C_{ℓ,m_ℓ,s,m_s} are called the *Clebsch-Gordan coefficients*.

Say $\ell = 1, s = \frac{1}{2}$, then $j = 1 \pm \frac{1}{2} = \frac{3}{2}, \frac{1}{2}$.

The following values of m_j are then possible (see Fig. 70):

$j = \frac{3}{2}$		$j = \frac{1}{2}$	
m_j	(m_ℓ, m_s)	m_j	(m_ℓ, m_s)
$+\frac{3}{2}$	$(+1, +\frac{1}{2})$	$+\frac{1}{2}$	$(+1, -\frac{1}{2})$ or $(0, +\frac{1}{2})$
$+\frac{1}{2}$	$(+1, -\frac{1}{2})$ or $(0, +\frac{1}{2})$	$-\frac{1}{2}$	$(-1, +\frac{1}{2})$ or $(0, -\frac{1}{2})$
$-\frac{1}{2}$	$(-1, +\frac{1}{2})$ or $(0, -\frac{1}{2})$	$-\frac{3}{2}$	$(-1, -\frac{1}{2})$

These states are generated by the total ladder operators:

- $\hat{J}_\pm = \hat{L}_\pm + \hat{S}_\pm$.
- \hat{L}_\pm acts only on Y_{ℓ,m_ℓ} .
- \hat{S}_\pm acts only on $|s, m_s\rangle$.
- $\hat{J}_\pm |j, m_j\rangle = \hbar \sqrt{j(j+1) - m_j(m_j \pm 1)} |j, m_j \pm 1\rangle$.

For example consider the highest total angular momentum state

$$|\frac{3}{2}, \frac{3}{2}\rangle = Y_{1,1} |\uparrow\rangle, \quad (777)$$

where we have used the notation $|j, m_j\rangle$. Then

$$\hat{J}_- |\frac{3}{2}, \frac{3}{2}\rangle = \sqrt{3} \hbar |\frac{3}{2}, \frac{1}{2}\rangle \quad (778)$$

$$(\hat{L}_- + \hat{S}_-) Y_{1,1} |\uparrow\rangle = \sqrt{2} \hbar Y_{1,0} |\uparrow\rangle + \hbar Y_{1,1} |\downarrow\rangle \quad (779)$$

$$\Rightarrow |\frac{3}{2}, \frac{1}{2}\rangle = \sqrt{\frac{2}{3}} Y_{1,0} |\uparrow\rangle + \sqrt{\frac{1}{3}} Y_{1,1} |\downarrow\rangle. \quad (780)$$

The complete set is as follows.

For the $j = 3/2$ ladder

$$\left| \frac{3}{2}, \frac{3}{2} \right\rangle = Y_{1,1} | \uparrow \rangle \quad (781)$$

$$\left| \frac{3}{2}, \frac{1}{2} \right\rangle = \sqrt{\frac{2}{3}} Y_{1,0} | \uparrow \rangle + \sqrt{\frac{1}{3}} Y_{1,1} | \downarrow \rangle \quad (782)$$

$$\left| \frac{3}{2}, -\frac{1}{2} \right\rangle = \sqrt{\frac{1}{3}} Y_{1,-1} | \uparrow \rangle + \sqrt{\frac{2}{3}} Y_{1,0} | \downarrow \rangle \quad (783)$$

$$\left| \frac{3}{2}, -\frac{3}{2} \right\rangle = Y_{1,-1} | \downarrow \rangle, \quad (784)$$

and for the $j = 1/2$ ladder

$$\left| \frac{1}{2}, \frac{1}{2} \right\rangle = \sqrt{\frac{1}{3}} Y_{1,0} | \uparrow \rangle - \sqrt{\frac{2}{3}} Y_{1,1} | \downarrow \rangle \quad (785)$$

$$\left| \frac{1}{2}, -\frac{1}{2} \right\rangle = \sqrt{\frac{2}{3}} Y_{1,-1} | \uparrow \rangle - \sqrt{\frac{1}{3}} Y_{1,0} | \downarrow \rangle. \quad (786)$$

Note that (782) and (785) are orthogonal, and also (783) and (786) are orthogonal.

91 Combining spins

This same structure applies when combining the spins of two particles: Fig. 71.

Consider two particles having individual spin operators $\hat{\mathbf{S}}_1$ and $\hat{\mathbf{S}}_2$, such that their individual spin quantum numbers are s_1 and s_2 .

How do we find the eigenvalues and eigenstates of the total spin operator

$$\hat{\mathbf{S}} = \hat{\mathbf{S}}_1 + \hat{\mathbf{S}}_2 ? \quad (787)$$

Consider the case of a system of two spin-1/2 particles, say the two electrons of a He atom, or the two electrons

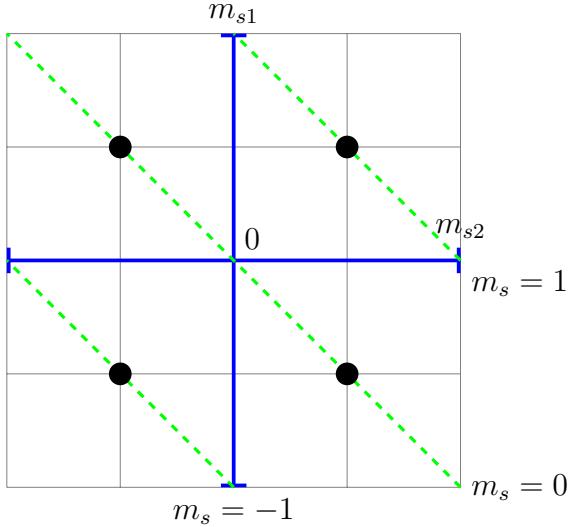


Figure 71: Combining two spin ladder operators (blue) to create a total (green) spin ladder operator.

in a H_2 bond.

There are 4 states in total:

$$|X_{s,m_s}\rangle \quad (788)$$

where s is the total spin quantum number, and m_s is the quantum number associated with the z component.

The maximum value of the total spin is $s = 1/2 + 1/2 = 1$, and therefore the highest state is

$$|X_{1,1}\rangle = |\uparrow\rangle_1|\uparrow\rangle_2. \quad (789)$$

Now create the other states in the series through the lowering operator $\hat{\mathbf{S}}_- = \hat{\mathbf{S}}_{1-} + \hat{\mathbf{S}}_{2-}$, where

$$\hat{\mathbf{S}}_-|X_{s,m_s}\rangle = \hbar\sqrt{s(s+1) - m_s(m_s-1)}|X_{s,m_s-1}\rangle, \quad (790)$$

and so

$$\hat{\mathbf{S}}_-|X_{1,1}\rangle = \hbar\sqrt{2}|X_{1,0}\rangle. \quad (791)$$

Also

$$\begin{aligned} \hat{\mathbf{S}}_-|X_{1,1}\rangle &= (\hat{\mathbf{S}}_{1-} + \hat{\mathbf{S}}_{2-})|\uparrow\rangle_1|\uparrow\rangle_2 \\ &= \hbar(|\downarrow\rangle_1|\uparrow\rangle_2 + |\uparrow\rangle_1|\downarrow\rangle_2), \end{aligned} \quad (792)$$

and so

$$|X_{1,0}\rangle = \frac{1}{\sqrt{2}} (|\downarrow\rangle_1 |\uparrow\rangle_2 + |\uparrow\rangle_1 |\downarrow\rangle_2). \quad (793)$$

Using the lowering operator again gives

$$\hat{\mathbf{S}}_- |X_{1,0}\rangle = \hbar\sqrt{2} |X_{1,-1}\rangle, \quad (794)$$

but also

$$\begin{aligned} \hat{\mathbf{S}}_- |X_{1,0}\rangle &= (\hat{\mathbf{S}}_{1-} + \hat{\mathbf{S}}_{2-}) \frac{1}{\sqrt{2}} (|\downarrow\rangle_1 |\uparrow\rangle_2 + |\uparrow\rangle_1 |\downarrow\rangle_2) \\ &= \hbar\sqrt{2} |\downarrow\rangle_1 |\downarrow\rangle_2 \\ |X_{1,-1}\rangle &= |\downarrow\rangle_1 |\downarrow\rangle_2. \end{aligned} \quad (795)$$

This sequence, which keeps the total spin constant whilst stepping down the z component, is called a *triplet state*.

The minimum value of the total spin is $s = 1/2 - 1/2 = 0$, and therefore

$$|X_{0,0}\rangle = \frac{1}{\sqrt{2}} (|\downarrow\rangle_1 |\uparrow\rangle_2 - |\uparrow\rangle_1 |\downarrow\rangle_2), \quad (796)$$

which follows because $|X_{0,0}\rangle$ must be orthogonal to $|X_{1,0}\rangle$.

There is only a single state in this ‘sequence’, and it is called a *singlet state*.

In conclusion, for the $s = 1$ ladder (Fig. 72)

$$|X_{1,1}\rangle = |\uparrow\rangle_1 |\uparrow\rangle_2 \quad (797)$$

$$|X_{1,0}\rangle = \frac{1}{\sqrt{2}} (|\downarrow\rangle_1 |\uparrow\rangle_2 + |\uparrow\rangle_1 |\downarrow\rangle_2)$$

$$|X_{1,-1}\rangle = |\downarrow\rangle_1 |\downarrow\rangle_2,$$

and for the $s=0$ ladder (Fig. 73)

$$|X_{0,0}\rangle = \frac{1}{\sqrt{2}} (|\downarrow\rangle_1 |\uparrow\rangle_2 - |\uparrow\rangle_1 |\downarrow\rangle_2). \quad (798)$$

Note that the 3 triplet states are symmetric with respect to an interchange of the two particles, whereas the singlet state is antisymmetric with respect to an interchange.

The spins of some well-known particles are as follows:

Spin	1/2	0	1
Electron	•		
Proton	•		
Neutron	•		
Quark	•		
Photon			•
Phonon		•	
Higgs boson		•	

An electron can undergo a magnetic dipole transition from the spin down to the spin up state. Why, in terms of angular momentum, can a photon induce this change?

Rules for combining the spins of multiple particles can also be deduced.

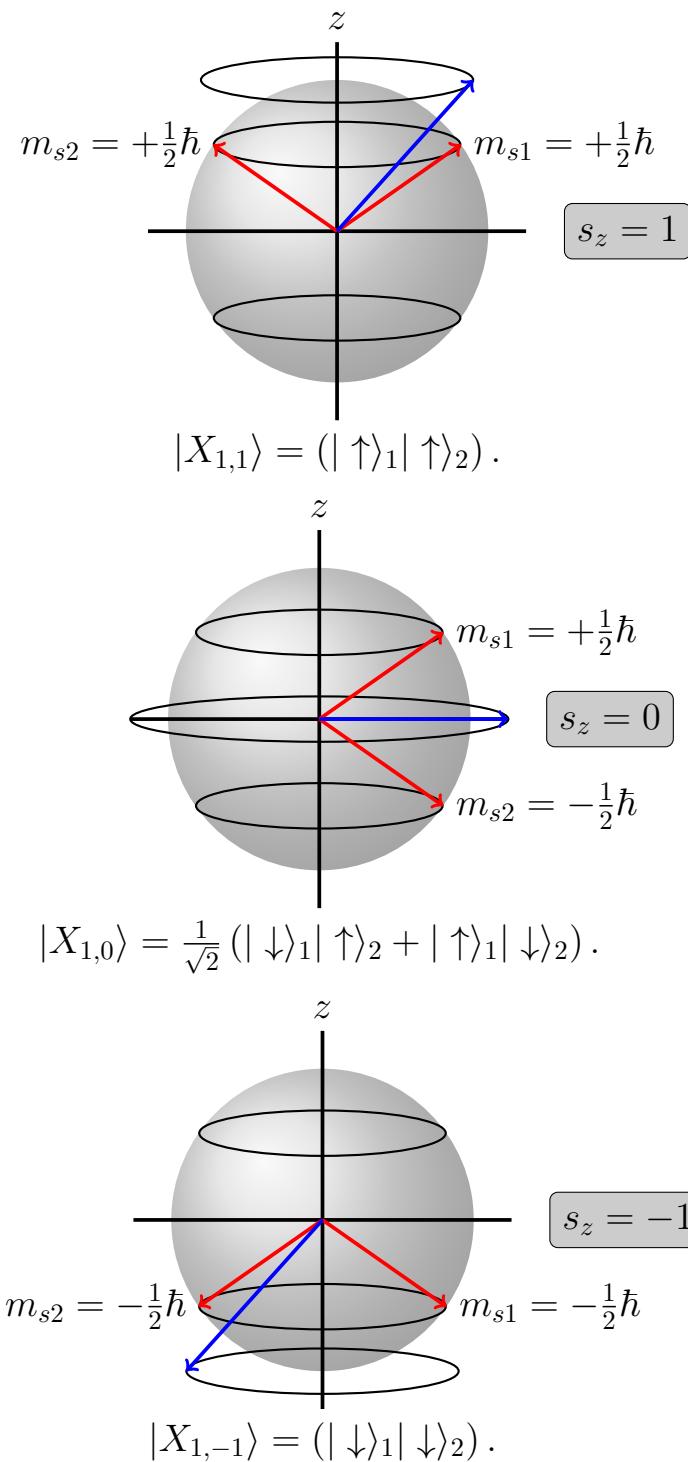


Figure 72: The three triplet states of two spin 1/2 particles.

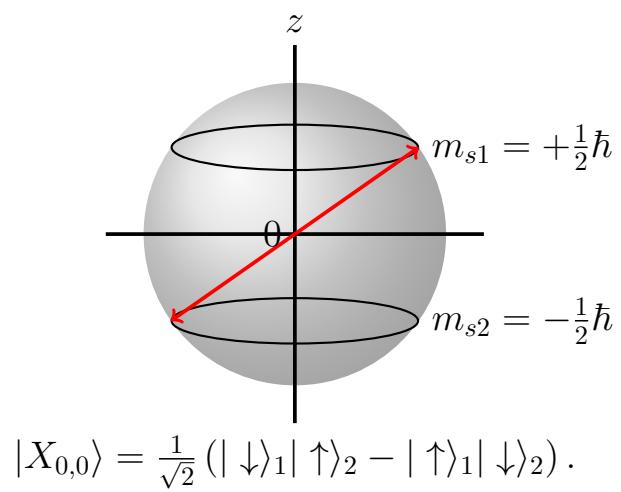


Figure 73: The singlet state of two spin 1/2 particles.

XII Matrix Mechanics

In the previous chapter we introduced the notion of intrinsic angular momentum, or spin, of a particle. We saw how one can combine the spin and orbital angular momentum into the total angular momentum, and how this can be formulated in terms of eigenstates and eigenvalues, and composite wave functions. We also studied how to combine spins of different particles. In this chapter, we shall encounter another iconic example of a quantum mechanical system that requires the combination of different states: the two level system. We shall describe it using the vector and operator-matrix formalism introduced in chapter VI, which we will review briefly. We shall introduce the notions of overlap integrals and level splitting / repulsion, in the context of a perturbative study of the problem. This chapter also includes two asides that illustrate simple applications of the matrix representation to orbital and spin angular momentum, and two further non-examinable asides on spin matrices and density matrices.

92 Matrix elements

In chapter VI, we saw the close relationship between discrete column vectors and wave functions.

These ‘vectors’ can be handled in similar ways and seem to share the same mathematical structures as wave functions.

The matrix formulation of quantum mechanics was devised by Werner Heisenberg, and it is particularly effective at describing the behaviour of discrete quantum systems where only a few states are involved.

Quantum states are represented by column vectors, and sequences of operators, or equivalently measurements, are represented by matrix multiplication.

In the following, we shall discuss this at a more formal and detailed level.

Any state can be expanded as a weighted linear combination of complete basis states:

$$|\psi\rangle = \sum_n c_n |\phi_n\rangle \text{ where } c_n = \langle \phi_n | \psi \rangle. \quad (799)$$

Suppose that

$$|\xi\rangle = \hat{A}|\psi\rangle = \sum_n c_n \hat{A}|\phi_n\rangle \quad (800)$$

and, in general,

$$|\xi\rangle = \sum_n d_n |\phi_n\rangle \text{ where } d_n = \langle \phi_n | \xi \rangle. \quad (801)$$

Forming inner products on both sides of Eq. (800),

$$\langle \phi_m | \xi \rangle = \sum_n c_n \langle \phi_m | \hat{A} | \phi_n \rangle, \quad (802)$$

and comparing with Eq. (801), gives

$$d_m = \sum_n A_{mn} c_n, \quad \forall m, \quad (803)$$

where the *matrix elements* are given by

$$A_{mn} = \langle \phi_m | \hat{A} | \phi_n \rangle. \quad (804)$$

The operator can then be reconstructed through

$$\hat{A} = \sum_{mn} A_{mn} |\phi_m\rangle \langle \phi_n|. \quad (805)$$

Eq. (803) is the matrix equation

$$\boxed{\mathbf{d} = \mathbf{A}\mathbf{c}.} \quad (806)$$

From (804), if the $|\phi_n\rangle$ are the eigenvectors of \hat{A} , then \mathbf{A} is diagonal, $A_{mn} = \delta_{mn}\lambda_n$.

\mathbf{A} can therefore be diagonalised using numerical techniques to give the eigenvectors.

If the system has a discrete spectrum, and a finite number of states, there is an *equivalent* or *isomorphic* representation of states and operators in terms of vectors and matrices.

Matrix mechanics is particularly effective at describing few-level systems, orbital angular momentum and spin.

Even when the basis set comprises an infinite number of states, matrix mechanics can be used to generate good approximations by truncating the basis set at a sufficiently high order.

This approximation applies when the probability of high-order states being occupied is low.

93 Two level systems

Two-level systems appear extensively in nature, and they provide a good example of the use of matrix mechanics.

They have beneficial applications, such as qubits in quantum computing, but are also associated with undesirable behaviour, such as excessive low-temperature heat capacity and dielectric losses in amorphous dielectrics.

As a specific case, consider a double potential well; such as an electron in the vicinity of two close atoms.

Fig. 74 shows the potential of a diatomic molecule as a function of position.

V_1 is the potential due to one nucleus, and V_2 is the potential due to the other. The total potential is shown in black. The two lowest energy states of the isolated potential wells are indicated.

The electron can transition from being in the left well to the right well through tunneling.

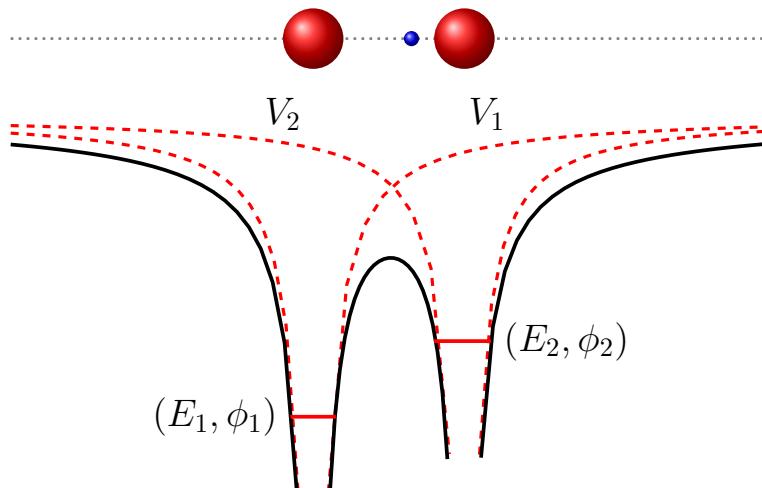


Figure 74: Electron potential as a function of position for a diatomic molecule. Dashed red: the potentials of two isolated nuclei. Solid black: the total potential. The two lowest energy states of the isolated nuclei are also shown.

Schrödinger's equation

$$\hat{H}|\psi\rangle = E|\psi\rangle \quad (807)$$

must apply.

The two spatially separated potential wells $V_1(x)$ and $V_2(x)$, when in isolation, would have the Hamiltonians \hat{H}_1 and \hat{H}_2 , respectively.

They are associated with the energy eigenstates $\{|\phi_{1i}\rangle\}$

and $\{|\phi_{2j}\rangle\}$ having eigenvalues $\{E_{1i}\}$ and $\{E_{2j}\}$.

Can we find the eigenstates and eigenvalues of the system when the two wells are allowed to interact weakly?

A starting approximation is to form a basis for the coupled system consisting of $\{|\phi_{1i}\rangle\}$ and $\{|\phi_{2j}\rangle\}$.

This process is called *hybridisation*.

For simplicity, assume that only the lowest two states, one in each well, are of importance.

For the isolated systems,

$$\begin{aligned}\hat{H}_1|\phi_1\rangle &= E_1|\phi_1\rangle \\ \hat{H}_2|\phi_2\rangle &= E_2|\phi_2\rangle.\end{aligned}\quad (808)$$

For any *two-level system*, the matrix-equivalent of Schrödinger's equation is

$$\begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = E \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}. \quad (809)$$

The total potential is the sum of the two potential wells:

$$\hat{V} = \hat{V}_1 + \hat{V}_2. \quad (810)$$

For the composite system, having only one particle,

$$\hat{H} = \hat{T} + \hat{V}_1 + \hat{V}_2, \quad (811)$$

and

$$\hat{H}|\psi\rangle = E|\psi\rangle. \quad (812)$$

Then

$$\begin{aligned}\hat{H} &= \hat{H}_1 + \hat{V}_2 \\ \hat{H} &= \hat{H}_2 + \hat{V}_1.\end{aligned}\quad (813)$$

To solve the problem use

$$|\psi\rangle = c_1|\phi_1\rangle + c_2|\phi_2\rangle \quad (814)$$

as the expansion.

The matrix elements of \hat{H} are

$$\begin{aligned} H_{11} &= \langle\phi_1|\hat{H}_1 + \hat{V}_2|\phi_1\rangle \\ &= E_1 + \langle\phi_1|\hat{V}_2|\phi_1\rangle. \end{aligned} \quad (815)$$

The wavefunction of well 1 is well-separated from the potential function of well 2, and so to a very good approximation, the matrix element $\langle\phi_1|\hat{V}_2|\phi_1\rangle$ is small compared with E_1 .

We conclude that

$$\begin{aligned} H_{11} &\approx E_1 \\ H_{22} &\approx E_2. \end{aligned} \quad (816)$$

The off-diagonal terms are given by

$$\begin{aligned} H_{12} &= \langle\phi_1|\hat{H}|\phi_2\rangle = t \\ H_{21} &= \langle\phi_2|\hat{H}|\phi_1\rangle = t^*, \end{aligned} \quad (817)$$

where the second line follows from the Hermiticity of \hat{H} , namely $\hat{H}^\dagger = \hat{H}$.

Terms of the form $\langle\phi_1|\dots|\phi_2\rangle$ are called *overlap integrals*, because they describe the ‘coupling’ between the wavefunctions of the two wells.

The matrix H is Hermitian, as is required for an observable.

Interaction occurs because we have taken the ‘isolated wavefunctions’ to be the basis set, not because we have explicitly included a coupling energy term in the Hamiltonian.

Nevertheless, t has the dimensions of energy, takes the form of a *coupling potential*, and appears as an off-diagonal element in the Hamiltonian.

To find the eigenvalues and eigenvectors of the coupled system, diagonalise H :

$$\begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \lambda \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}. \quad (818)$$

The secular equation is

$$\begin{vmatrix} H_{11} - \lambda & H_{12} \\ H_{21} & H_{22} - \lambda \end{vmatrix} = \begin{vmatrix} E_1 - \lambda & t \\ t^* & E_2 - \lambda \end{vmatrix} = 0. \quad (819)$$

Therefore

$$\boxed{\lambda = \frac{1}{2} \left(E_1 + E_2 \pm \sqrt{(E_2 - E_1)^2 + 4|t|^2} \right)}. \quad (820)$$

Using each value of λ , the eigenvectors are found by solving

$$\begin{pmatrix} E_1 - \lambda & t \\ t^* & E_2 - \lambda \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = 0. \quad (821)$$

As a check, in the case of $t = 0$, $\lambda^+ = E_2$ and $\lambda^- = E_1$.

Consider the case of two identical wells: $E_1 = E_2 = E$, and then

$$\boxed{\lambda_{\pm} = E \pm |t|}. \quad (822)$$

The energy level splitting is strongly dependent on the well separation a (recall from early on in the notes the fact that a quantum wave function decays exponentially with distance as it penetrates into a barrier; and $|t|$ depends on the overlap between the two wave functions ϕ_1 and ϕ_2).

Eq. (821) becomes

$$\begin{pmatrix} \mp|t| & t \\ t^* & \mp|t| \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = 0. \quad (823)$$

Assuming that t is real,

$$c_1 = \pm c_2. \quad (824)$$

Also

$$c_1^2 + c_2^2 = 1, \quad (825)$$

and therefore

$$|\psi\rangle_{\mp} = \frac{1}{\sqrt{2}} [|\phi_1\rangle \pm |\phi_2\rangle].$$

(826)

Note the relevant similarity to the \hat{S}_x eigenstates in the previous chapter, $(|\uparrow\rangle \pm |\downarrow\rangle)/\sqrt{2}$. Indeed the overlap matrix elements t in a two level system with $E_1 = E_2$ play the same role as the matrix elements of \hat{S}_x for a spin-1/2 degree of freedom, and a rigorous mapping between the two can be established.

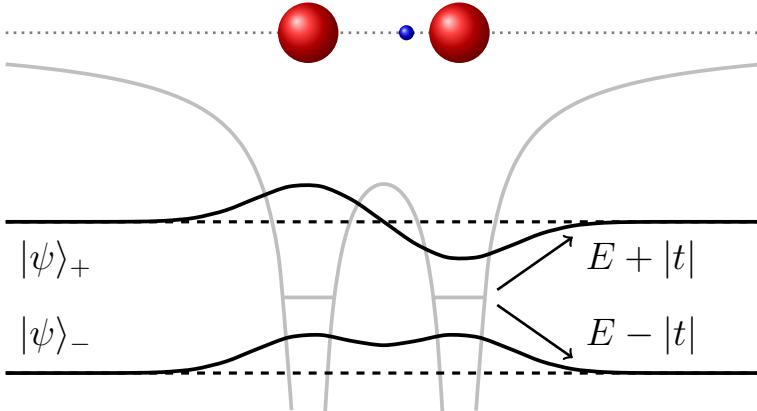


Figure 75: The original states, having energy E , split into two states having energies $E + |t|$ and $E - |t|$.

For a pair of identical potential wells, the original states interact and form symmetric and an antisymmetric states: Fig. 75.

The original energy levels split, the degeneracy is lost, and two new energy levels are formed. This is a characteristic feature of many interacting physical systems.

The symmetric wavefunction is associated with an energy that is lower than the original, whereas the antisymmetric wavefunction is associated with an energy that is higher than the original.

The lower state is associated with ‘bonding’, whereas the higher state is associated with ‘repulsion’ (also known as ‘antibonding’).

If the original levels each contained one electron, both can occupy the lower energy state of the double well, reducing the overall energy by $2|t|$ (in the approximation where we neglect electron-electron interactions).

The sharing of electrons to lower the total energy of a system is in essence the formation of a covalent bond between the two bound systems, as in the case of H_2 : Fig. 76.

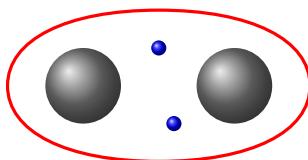


Figure 76: Two electrons being shared in a covalent bond.

Solving the problem for multiple interacting potential wells requires numerical techniques, but ultimately gives rise to band structures, and calculations of the equilibrium arrangements of atoms in crystalline materials.

Consider the situation when the two wells are different: $E_1 \neq E_2$, and without loss of generality assume that

$E_1 < E_2$.

Assume also that $t = \langle \phi_1 | \hat{H} | \phi_2 \rangle$ is much smaller than $E_2 - E_1$ (i.e., weak coupling):

$$\frac{|t|}{E_2 - E_1} = \delta \ll 1. \quad (827)$$

Then

$$\lambda = \frac{1}{2} \left(E_1 + E_2 \pm (E_2 - E_1) \sqrt{1 + 4\delta^2} \right). \quad (828)$$

The two energy levels become

$$\begin{aligned} \lambda_- &= E_1 - (E_2 - E_1)\delta^2 \\ &= E_1 - \Delta, \end{aligned} \quad (829)$$

and

$$\begin{aligned} \lambda_+ &= E_2 + (E_2 - E_1)\delta^2 \\ &= E_2 + \Delta, \end{aligned} \quad (830)$$

where Δ is a positive quantity, but E_1 and E_2 are negative.

λ_- is the lowest energy and λ_+ is the highest energy.

Remembering that $\delta \ll 1$, one can obtain the approximate eigenstates

$|\psi\rangle_- = \frac{1}{\sqrt{1 + \delta^2}} [|\phi_1\rangle + \delta|\phi_2\rangle]$

(831)

and

$|\psi\rangle_+ = \frac{1}{\sqrt{1 + \delta^2}} [-\delta|\phi_1\rangle + |\phi_2\rangle].$

(832)

In this case, the upper energy level shifts up and the lower energy level shifts down: Fig. 77

This is called *level repulsion* and it is a generic outcome

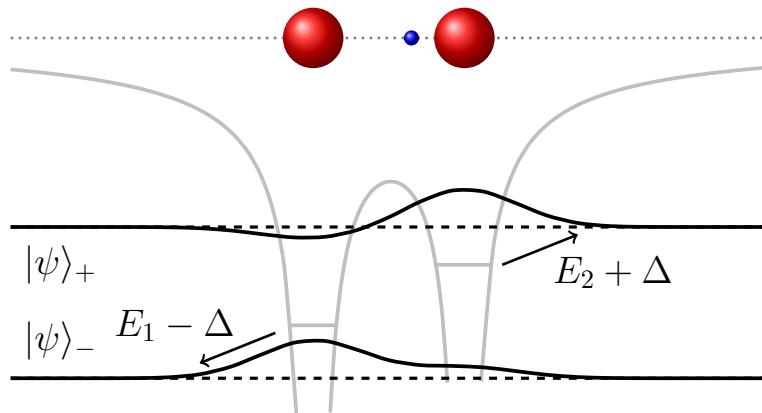


Figure 77: The original states, having energy E_1 and E_2 , change into two states having energies $E_1 - \Delta$ and $E_2 + \Delta$.

when distinct energy eigenstates are perturbatively coupled to one another.

$|\psi\rangle_-$ is dominated by $|\phi_1\rangle$ for the level $E_1 - \Delta$, whereas $|\psi\rangle_+$ is dominated by $|\phi_2\rangle$ for the level $E_2 + \Delta$ (remember $\delta \ll 1$).

If the original levels each hold 1 electron, they can both occupy the lower level of the double well, reducing the overall energy by $(E_2 + E_1) - 2(E_1 - \Delta) = E_2 - E_1 + 2\Delta$.

This behaviour is characteristic of ionic bonding, as in the case of NaCl, because it is associated with a net transfer of charge between the two wells: Fig. 78.

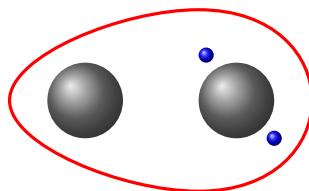


Figure 78: An electron being transferred in an ionic bond.

If the system is shown in Fig. 77 is not in one of the energy eigenstates, how does the likely location of the electron vary with time?

94 Aside 1: Orbital angular momentum in matrix representation

Matrix mechanics can be used to represent angular momentum.

As an example, use the $\ell = 1$ states as the basis:
 $\{|+1\rangle, |0\rangle, |-1\rangle\}$.

Use the following representations for these states:

$$|+1\rangle \equiv \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |0\rangle \equiv \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad |-1\rangle \equiv \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (833)$$

The matrix elements of the \hat{L}_z operator are then

$$\begin{pmatrix} \langle +1|\hat{L}_z|+1\rangle & \langle 0|\hat{L}_z|+1\rangle & \langle -1|\hat{L}_z|+1\rangle \\ \langle +1|\hat{L}_z|0\rangle & \langle 0|\hat{L}_z|0\rangle & \langle -1|\hat{L}_z|0\rangle \\ \langle +1|\hat{L}_z|-1\rangle & \langle 0|\hat{L}_z|-1\rangle & \langle -1|\hat{L}_z|-1\rangle \end{pmatrix}, \quad (834)$$

giving

$$\mathbf{L}_z = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \quad (835)$$

Recall from Handout IX that

$$\hat{L}_{\pm}|\ell, m_{\ell}\rangle = \hbar\sqrt{\ell(\ell+1) - m_{\ell}(m_{\ell} \pm 1)}|\ell, m_{\ell} \pm 1\rangle. \quad (836)$$

Therefore

$$\hat{L}_+|+1\rangle = 0, \quad \hat{L}_+|0\rangle = \sqrt{2}\hbar|+1\rangle, \quad \hat{L}_+|-1\rangle = \sqrt{2}\hbar|0\rangle, \quad (837)$$

and

$$\hat{L}_-|+1\rangle = \sqrt{2}\hbar|0\rangle, \quad \hat{L}_-|0\rangle = \sqrt{2}\hbar|-1\rangle, \quad \hat{L}_-|-1\rangle = 0. \quad (838)$$

The matrix elements become

$$\mathbf{L}_+ = \hbar \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix} \quad (839)$$

and

$$\mathbf{L}_- = \hbar \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix}. \quad (840)$$

\mathbf{L}_+ is the conjugate transpose of \mathbf{L}_- .

Since $\hat{\mathbf{L}}_+ = \hat{\mathbf{L}}_x + i\hat{\mathbf{L}}_y$ and $\hat{\mathbf{L}}_- = \hat{\mathbf{L}}_x - i\hat{\mathbf{L}}_y$, we find

$$\mathbf{L}_x = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad (841)$$

and

$$\mathbf{L}_y = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}. \quad (842)$$

It is straightforward to show that

$$\mathbf{L}_x \mathbf{L}_y - \mathbf{L}_y \mathbf{L}_x = i\hbar \mathbf{L}_z, \quad (843)$$

which is the usual commutation relationship.

It is possible to extend this scheme to higher-order eigenstates.

The matrix formulation encompasses much of the structure of orbital angular momentum without introducing wavefunctions.

95 Aside 2: Spin angular momentum in matrix representation

Matrix mechanics can be used to represent spin in an elegant way.

Use the state designations

$$|\uparrow\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\downarrow\rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (844)$$

Consider the matrix elements of \hat{S}_z :

$$\begin{pmatrix} \langle \uparrow | \hat{S}_z | \uparrow \rangle & \langle \uparrow | \hat{S}_z | \downarrow \rangle \\ \langle \downarrow | \hat{S}_z | \uparrow \rangle & \langle \downarrow | \hat{S}_z | \downarrow \rangle \end{pmatrix}, \quad (845)$$

or

$$\mathbf{S}_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (846)$$

The matrix elements of \hat{S}_+ are

$$\begin{pmatrix} \langle \uparrow | \hat{S}_+ | \uparrow \rangle & \langle \uparrow | \hat{S}_+ | \downarrow \rangle \\ \langle \downarrow | \hat{S}_+ | \uparrow \rangle & \langle \downarrow | \hat{S}_+ | \downarrow \rangle \end{pmatrix}, \quad (847)$$

giving

$$\mathbf{S}_+ = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}. \quad (848)$$

Similarly,

$$\mathbf{S}_- = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \quad (849)$$

Also,

$$\hat{S}_x = \frac{1}{2} (\hat{S}_+ + \hat{S}_-) \quad (850)$$

and

$$\hat{S}_y = \frac{1}{2i} (\hat{S}_+ - \hat{S}_-), \quad (851)$$

giving

$$\mathbf{S}_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (852)$$

and

$$\mathbf{S}_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (853)$$

Why in quantum mechanics is it necessary to introduce the notion of ‘spin’ when referring to photons?

96 Aside 3: Spin matrices – non-examinable

The three matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (854)$$

are of special importance, and they are called *Pauli spin matrices* or *Pauli spin operators*.

When combined with the identity matrix they form a complete set in terms of which any 2×2 complex matrix can be expanded.

The appropriate inner product *for matrices* is

$$\text{Tr} (\mathbf{AB}^\dagger), \quad (855)$$

which when applied to Eq. (854), and the identity matrix, shows that they form an orthogonal, but not orthonormal, set.

We can write any 2×2 complex matrix in the form

$$\mathbf{A} = a_1 \mathbf{I} + a_2 \sigma_x + a_3 \sigma_y + a_4 \sigma_z \quad (856)$$

and use Eq. (855) to determine the coefficients.

Also note that

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \mathbf{I}, \quad (857)$$

and that they anti-commute

$$\begin{aligned} \sigma_x \sigma_y + \sigma_y \sigma_x &= 0 \\ \sigma_z \sigma_x + \sigma_x \sigma_z &= 0 \\ \sigma_y \sigma_z + \sigma_z \sigma_y &= 0. \end{aligned} \quad (858)$$

The eigenvectors of the various operators are called *spinors*.

It is straightforward to show the spin matrices obey all of the behaviour of the various spin operators $\hat{S}_{x,y,z}$, \hat{S}_\pm , \hat{S}^2 .

spin matrix	eigenvalue	spinor	eigenvalue	spinor
S_x	$\frac{1}{2}\hbar$	$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$	$-\frac{1}{2}\hbar$	$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$
S_y	$\frac{1}{2}\hbar$	$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}$	$-\frac{1}{2}\hbar$	$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}$
S_z	$\frac{1}{2}\hbar$	$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$	$-\frac{1}{2}\hbar$	$\frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

The Pauli spin matrices are similar to *Stokes parameters* in classical electromagnetism, giving linear polarisation and left and right circular polarisation.

97 Aside 4: Density matrices – non-examinable

Now return to the notion of classically incoherent combinations of quantum systems.

Consider an ensemble of two-state quantum systems.

Moreover assume that the states are represented by the column vectors

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (859)$$

If the probability that a system is prepared in the first state is P , and the probability that it is prepared in the second state is $1 - P$, then the density matrix becomes

$$\mathbf{d} = P \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} + (1 - P) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix}, \quad (860)$$

or

$$\boxed{\mathbf{d} = \begin{pmatrix} P & 0 \\ 0 & 1 - P \end{pmatrix}}. \quad (861)$$

$\text{Tr}[\mathbf{d}] = 1$ as expected, and $\text{Tr}[\mathbf{d}^2] < 1$, indicating that the system is in a mixed state (unless $p = 1$ or $p = 0$).

For coherent superpositions, where the relative phase is important, we add column vectors; for incoherent superpositions, where the relative phase is randomised, we add density matrices.

Consider coupling an ensemble of two-level systems, having energy states E_1 and E_2 , to a thermal bath.

In this case the normalised density matrix is given by

$$\frac{1}{e^{-E_1/kT} + e^{-E_2/kT}} \begin{pmatrix} e^{-E_1/kT} & 0 \\ 0 & e^{-E_2/kT} \end{pmatrix}. \quad (862)$$

The expectation value of the total energy is $\langle \hat{H} \rangle = \text{Tr} [\mathbf{d}\mathbf{H}]$, which becomes

$$\text{Tr} \left[\frac{1}{e^{-E_1/kT} + e^{-E_2/kT}} \begin{pmatrix} e^{-E_1/kT} & 0 \\ 0 & e^{-E_2/kT} \end{pmatrix} \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} \right] \quad (863)$$

giving

$$\boxed{\langle \hat{H} \rangle = \frac{E_1 e^{-E_1/kT} + E_2 e^{-E_2/kT}}{e^{-E_1/kT} + e^{-E_2/kT}}.} \quad (864)$$

For small T , and assuming $E_2 > E_1$, one obtains $\langle \hat{H} \rangle \approx E_1$, whereas for large T one obtains $\langle \hat{H} \rangle \approx (E_1 + E_2)/2$, indicating that the states are equally occupied at high temperatures.

There are many situations where systems interact through the exchange of energy, but which are not quantum mechanically entangled. Consider a two level system having energy difference $\hbar\omega = E_2 - E_1$ resonantly coupled to a simple harmonic oscillator. The Hamiltonian takes the form

$$\begin{aligned} \hat{H} = & E_1 |\phi_1\rangle\langle\phi_1| + E_2 |\phi_2\rangle\langle\phi_2| + \hbar\omega(\hat{a}^\dagger\hat{a} + 1/2) \\ & + V (\hat{a}^\dagger|\phi_1\rangle\langle\phi_2| + \hat{a}|\phi_2\rangle\langle\phi_1|). \end{aligned}$$

What is the physical significance of the last two terms when applied to a separable composite state describing the two systems?

XIII Indistinguishable Particles

In the previous chapter we presented a perturbative study of a two level system, using the vector and operator-matrix formalism introduced in chapter VI. This lead us to introduce the notions of overlap integrals and level splitting / repulsion, typical of covalent / ionic bonds between atoms. In this chapter, we shall return to the description of systems comprising more than one particle, paying particular attention to the situation where the particles are *identical*. We will encounter the concept of exchange symmetry and the spin statistics theorem, whereby particles at the quantum mechanical level can either be symmetric upon exchange (bosons) or antisymmetric (fermions). We will close the lecture notes with a brief discussion of how this affects multi-particle states.

98 Identical particles

Consider an N -particle system where every particle has the same mass, charge, etc.

The composite wave function is written as

$$\Psi_N \equiv \Psi(\mathbf{r}_1, \dots, \mathbf{r}_N), \quad (865)$$

and the Hamiltonian has the form

$$\hat{H}_N \equiv \hat{H}(\mathbf{r}_1, \dots, \mathbf{r}_N; \mathbf{p}_1, \dots, \mathbf{p}_N). \quad (866)$$

Eq. (865) encodes information about the likely outcomes of experiments on N particles, such as measuring their

positions.

Eqs. (865) and (866) are related by the *many-particle* Schrödinger equation

$$\hat{H}_N \Psi_N = E_N \Psi_N, \quad (867)$$

where Ψ_N are the eigenfunctions of the complete system, and E_N are the corresponding eigenvalues, giving the possible values of the total energy.

There is some energy associated with finding particles simultaneously at particular locations and having particular momenta.

Different arrangements are generally associated with different energies.

If two identical particles are swapped, the energy of the configuration does not change.

For identical particles, the N -particle Hamiltonian must be invariant to particles being swapped.

The possible outcomes of a measurement, the eigenvalues, must be unchanged.

Exchange degeneracies appear because different arrangements have the same energy.

In fact, every operator corresponding to every two-particle measurement that might be performed, such as energy and angular momentum, must be invariant with respect to ‘swapping the labels’ on identical particles.

Otherwise, certain measurements would be available that would give different results when identical particles are swapped!

For example, the Hamiltonian of a He atom is approximately

$$\hat{H}_{He} = \frac{\hat{p}_n^2}{2m_n} + \frac{\hat{p}_1^2}{2m_e} + \frac{\hat{p}_2^2}{2m_e} - \frac{2e^2}{4\pi\epsilon_0|\mathbf{r}_1 - \mathbf{r}_n|} - \frac{2e^2}{4\pi\epsilon_0|\mathbf{r}_2 - \mathbf{r}_n|} + \frac{e^2}{4\pi\epsilon_0|\mathbf{r}_1 - \mathbf{r}_2|}, \quad (868)$$

where the subscript n refers to the nucleus, and the subscript e to the electrons.

\hat{H}_{He} is unaffected by relabelling the two electrons: $1 \leftrightarrow 2$.

For an N -particle system, interchanging pairs makes no difference if the particles are identical:

$$\begin{aligned} \hat{H}_N &\equiv \hat{H}(\mathbf{r}_1, \dots, \mathbf{r}_i, \dots, \mathbf{r}_j, \dots, \mathbf{r}_N; \mathbf{p}_1, \dots, \mathbf{p}_i, \dots, \mathbf{p}_j, \dots, \mathbf{p}_N) \\ &= \hat{H}(\mathbf{r}_1, \dots, \mathbf{r}_j, \dots, \mathbf{r}_i, \dots, \mathbf{r}_N; \mathbf{p}_1, \dots, \mathbf{p}_j, \dots, \mathbf{p}_i, \dots, \mathbf{p}_N). \end{aligned} \quad (869)$$

The exchange symmetry of *operators* (as distinct from states) of all possible observables defines what is meant by *identical particles*.

On a notational point, rather than writing out all of the labels, as in Eq. (869), it can be beneficial to only refer explicitly to pairs of particles that are identical.

For example, rather than writing

$$\Psi_N \equiv \Psi(\mathbf{r}_1, \dots, \mathbf{r}_i, \dots, \mathbf{r}_j, \dots, \mathbf{r}_N), \quad (870)$$

and

$$\hat{H}(\mathbf{r}_1, \dots, \mathbf{r}_i, \dots, \mathbf{r}_j, \dots, \mathbf{r}_N; \mathbf{p}_1, \dots, \mathbf{p}_i, \dots, \mathbf{p}_j, \dots, \mathbf{p}_N), \quad (871)$$

it is more convenient to write

$$\Psi_N \equiv \Psi(\xi_i, \xi_j), \quad (872)$$

and

$$\hat{H}(\xi_i, \xi_j), \quad (873)$$

with all unreferenced positions and momenta being assumed.

ξ_i corresponds to the dynamical variables of particle i . Thus, when we write $\hat{H}(\xi_j, \xi_i)$, it is the labelling of the particles that has changed, not simply the positions.

Identical particles have the property that

$$\hat{A}(\xi_i, \xi_j) = \hat{A}(\xi_j, \xi_i), \quad (874)$$

where \hat{A} is the operator corresponding to any physical quantity that can be measured.

Particles cannot be identified beyond any of the properties that might be measured.

99 Particle exchange operator

Define the *particle exchange operator* \hat{P}_{ij} , which has the effect of interchanging the indices:

$$\hat{P}_{ij}|\xi_i, \xi_j\rangle = |\xi_j, \xi_i\rangle. \quad (875)$$

Suppose that we have completely symmetrised / anti-symmetrised states:

$$\hat{P}_{ij}|\xi_i, \xi_j\rangle^{a/s} = \pm|\xi_i, \xi_j\rangle^{a/s}. \quad (876)$$

Then $|\xi_i, \xi_j\rangle^{a/s}$ are eigenstates of \hat{P}_{ij} with eigenvalues ± 1 .

It is straightforward to show that \hat{P}_{ij} is Hermitian. Indeed,

$$\begin{aligned} \hat{P}_{ij}|\xi_i, \xi_j\rangle &= |\xi_j, \xi_i\rangle \\ \langle \xi_i, \xi_j | \hat{P}_{ij}^\dagger &= \langle \xi_j, \xi_i |, \end{aligned} \quad (877)$$

and therefore

$$\begin{aligned} \langle \xi_i, \xi_j | \hat{P}_{ij}^\dagger \hat{P}_{ij} | \xi_i, \xi_j \rangle &= \langle \xi_j, \xi_i | \xi_j, \xi_i \rangle \\ &= 1. \end{aligned} \quad (878)$$

Eq. (878) implies that

$$\hat{P}_{ij}^\dagger \hat{P}_{ij} = \hat{I}, \quad (879)$$

but since applying \hat{P}_{ij} twice has no effect (it exchanges the same two particles twice), we also have that

$$\hat{P}_{ij}^2 = \hat{I}, \quad (880)$$

and so

$$\boxed{\hat{P}_{ij}^\dagger = \hat{P}_{ij}.} \quad (881)$$

The exchange operator \hat{P}_{ij} is Hermitian, and exchange symmetry is an observable.

It is possible to determine from a measurement whether particles are exchange symmetric or exchange antisymmetric.

100 Time evolution of the exchange operator

Suppose that \hat{A} represents some observable.

Mathematically, it is clear that

$$\langle \xi_i, \xi_j | \hat{A} | \xi_i, \xi_j \rangle = \langle \xi_j, \xi_i | \hat{A} | \xi_j, \xi_i \rangle, \quad (882)$$

which indicates that the expectation values of the states are the same, and that no measurement can distinguish between the particles.

Then

$$\begin{aligned} \langle \xi_i, \xi_j | \hat{A} | \xi_i, \xi_j \rangle &= \langle \xi_i, \xi_j | \hat{P}_{ij}^\dagger \hat{A} \hat{P}_{ij} | \xi_i, \xi_j \rangle \\ \hat{A} &= \hat{P}_{ij}^\dagger \hat{A} \hat{P}_{ij} \\ \hat{P}_{ij} \hat{A} &= \hat{A} \hat{P}_{ij}. \end{aligned} \quad (883)$$

Then

$$\boxed{[\hat{A}, \hat{P}_{ij}] = 0.} \quad (884)$$

The exchange operator and any observable are compatible.

No experiment can possibly distinguish one particle from another, because the particles can be swapped prior to a measurement without changing the result.

For example,

$$\boxed{[\hat{H}, \hat{P}_{ij}] = 0,} \quad (885)$$

and so exchange symmetry is a constant of motion.

If a wavefunction is exchange symmetric, or antisymmetric, it remains so as time evolves.

Fermions do not change into bosons as time progresses.

The time-independence of exchange symmetry is also preserved when measurements are made.

Suppose that a system is in state $|\psi\rangle$; it must be exchange symmetric or antisymmetric with respect to i and j .

It is therefore an eigenstate of \hat{P}_{ij} with eigenvalue $\eta = \pm 1$.

Consider some general measurement \hat{A} .

$|\psi\rangle$ can be expanded in terms of the eigenstates of \hat{A} :

$$|\psi\rangle = \sum_n a_n |a_n\rangle. \quad (886)$$

Because $[\hat{A}, \hat{P}_{ij}] = 0$, the $|a_n\rangle$ are also eigenstates of \hat{P}_{ij} with eigenvalues $\eta_n = \pm 1$.

Therefore

$$\begin{aligned}\hat{P}_{ij}|\psi\rangle &= \sum_n a_n \hat{P}_{ij}|a_n\rangle \\ &= \sum_n a_n \eta_n |a_n\rangle\end{aligned}\quad (887)$$

but also

$$\eta|\psi\rangle = \eta \sum_n a_n |a_n\rangle. \quad (888)$$

Taking the inner product of (887) and (888) with respect to each of the $\langle a_m |$ shows that $\eta_m = \eta, \forall m$.

Each of the eigenvectors of \hat{A} must have the same exchange symmetry.

If $|\psi\rangle$ is an eigenstate of \hat{P}_{ij} with eigenvalue η then its expansion in terms of the eigenstates of \hat{A} can only include those terms which are also eigenstates of \hat{P}_{ij} with the same eigenvalue as $|\psi\rangle$.

An exchange symmetric state can only be expressed as a weighted linear combination of exchange symmetric states.

An exchange antisymmetric state can only be expressed as a weighted linear combination of exchange antisymmetric states.

When a measurement is made, and the wavefunction collapses, the exchange symmetry remains unaltered.

Exchange symmetry is a conserved quantity unaffected by the measurement of a particle's characteristics: position, momentum, etc.

Fermions do not change into bosons after a measurement.

101 Indistinguishable particles

The notion that particles may be *indistinguishable*, as distinct from *identical*, is more fundamental than one might at first imagine.

What does it mean to swap the arguments of a wavefunction: $\Psi(\mathbf{r}_1, \mathbf{r}_2) \rightarrow \Psi(\mathbf{r}_2, \mathbf{r}_1)$?

Suppose that particle *A* is at position \mathbf{r}_1 , and particle *B* is at position \mathbf{r}_2 . It is also possible to find particle *B* at position \mathbf{r}_1 and particle *A* at position \mathbf{r}_2 . Swapping the arguments refers to these two arrangements.

Classically they can be distinguished because we can imagine tracking the particles as they move around.

We can watch *A* being moved to position \mathbf{r}_1 , and *B* being moved to position \mathbf{r}_2 .

In quantum mechanics, the idea of a classical trajectory does not make sense because every time a position measurement is made the particle can 'jump' between possible outcomes (albeit with some weighting), giving a non-continuous distribution of sampled positions.

If the wavefunctions of two particles overlap significantly, it is not possible to say which particle has been detected when an experiment is performed, as illustrated pictorially in Fig. 79.

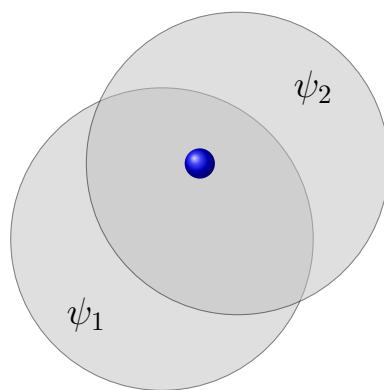


Figure 79: Particle A is most likely found in region 1, whereas particle B is most likely found in region 2. If an experiment is performed and a particle is found at the position shown, is it particle A or particle B?

Equivalently, when two particles elastically scatter, it is often not possible to relate the outgoing particles to the incoming particles in one-to-one correspondence, because of the uncertainty in their positions during the scattering event.

Or consider a collection of particles. If two experiments are carried out to measure the positions of all particles on two systems in the same quantum state, the distribution of positions will in general be different, and so which particle is which?

The state corresponding to ‘A’ being at position \mathbf{r}_1 and ‘B’ being at \mathbf{r}_2 is *exactly the same state* as ‘B’ being at position \mathbf{r}_1 and ‘A’ being at \mathbf{r}_2 .

The state gives the probability of a particle being found at position \mathbf{r}_1 and a particle being found at position \mathbf{r}_2 without any reference to which particle is which.

The two *apparently different states* are in fact only a *single state*, and this has profound consequences for understanding the physical world at the microscopic level.

Quantum mechanically, when a number of identical particles are collected together, the notion of individuality is lost.

We cannot attach some indelible, or even imagined, label beyond the qualities that are characterised by a complete set of quantum numbers.

The argument can be extended to large collections of identical particles.

Each available element in the argument labels a specific particle, whereas each position vector labels a particular point in space; thus, $\Psi(\mathbf{r}_1, \mathbf{r}_2)$ corresponds to particle A being at position \mathbf{r}_1 and particle B at \mathbf{r}_2 , whereas $\Psi(\mathbf{r}_2, \mathbf{r}_1)$ corresponds to particle A being at position \mathbf{r}_2 and particle B at \mathbf{r}_1 .

By swapping the position labels, we swap the positions of the particles.

By having an ordered list in the argument we are already falling into the bad habit of identifying particles!

More problematically we seem to have too many degrees of freedom in the original notation.

We seem to have two different state vectors describing the same experimental outcome.

Should we simply throw away the unused degrees of freedom?

Should we change the notation?

102 Exchange symmetry

For two particles, the two apparently different arrangements refer, profoundly, to the same physical situation. Perhaps identical particles do not have their own ‘identities’ at any level. This is what is meant by ‘indistinguishable’.

The probability density

$$|\Psi(\mathbf{r}_1, \mathbf{r}_2)|^2 \quad (889)$$

is the joint probability of finding particle A in a differential region around \mathbf{r}_1 and particle B in a differential region around \mathbf{r}_2 .

Eq. (889) is a joint probability, not the product of two single probabilities; this joint information constitutes a single state in its own right.

All that can be known is that two particles, or excitations, are found simultaneously at the two locations.

Then

$$|\Psi(\mathbf{r}_2, \mathbf{r}_1)|^2 = |\Psi(\mathbf{r}_1, \mathbf{r}_2)|^2. \quad (890)$$

And therefore,

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = \pm \Psi(\mathbf{r}_2, \mathbf{r}_1). \quad (891)$$

The probability of finding a particle at \mathbf{r}_1 and a particle at \mathbf{r}_2 does not depend on which particle is which.

In conclusion,

$\begin{aligned} \Psi(\mathbf{r}_1, \mathbf{r}_2) &= +\Psi(\mathbf{r}_2, \mathbf{r}_1) \\ &\text{or} \\ \Psi(\mathbf{r}_1, \mathbf{r}_2) &= -\Psi(\mathbf{r}_2, \mathbf{r}_1). \end{aligned}$	(892)
---	---------

Eq. (892) suggests that two different particle-exchange symmetries are possible.

103 Exchange from a state-vector perspective

The above reasoning is based on wavefunction considerations, but how does exchange symmetry appear in a general state-vector description?

Consider again an N -particle Hamiltonian.

We need a basis to describe the action of general operators.

As seen in earlier chapters, one can use Cartesian products of the states that characterise the behaviour of each particle separately.

A typical basis ket is

$$|\chi_1\rangle|\chi_2\rangle\cdots|\chi_i\rangle\cdots|\chi_N\rangle. \quad (893)$$

$|\chi_1\rangle$ is a ket describing the state of particle 1. χ_1 is shorthand for a complete set of quantum numbers.

For orbital electrons, $\chi \equiv n, \ell, m_\ell, s, m_s$ as appropriate. It characterises the outcomes of all possible experiments on particle 1.

The complete collection of states having the form of Eq. (893) is a basis for the system because it includes all possible, in principle at least, experimental outcomes.

It is called the *Hartree product*.

General states must be formed from linear combinations of these basis states.

The likely outcome of an experiment will then be given by the relevant projection.

However, it is not clear what combinations of basis states are physically meaningful.

The following form of operator appears frequently when describing the behaviour of interacting particles, say electrons:

$$\hat{V} = \sum_{ijkl} V_{lkji} |\chi_l\rangle |\chi_k\rangle \langle \chi_j| \langle \chi_i|. \quad (894)$$

From a mapping perspective describe its action. Without worrying about the details, speculate on how it might be used.

Consider a two-particle system.

If particle 1 has two possible energy states, $|\alpha_1\rangle$ and $|\beta_1\rangle$, and particle two has two possible energy states, $|\alpha_2\rangle$ and $|\beta_2\rangle$, then a suitable basis to describe all possible outcomes is

$$|\alpha_1\rangle|\alpha_2\rangle, \quad |\alpha_1\rangle|\beta_2\rangle, \quad |\beta_1\rangle|\alpha_2\rangle, \quad |\beta_1\rangle|\beta_2\rangle \quad (895)$$

and these must be combined with suitable weightings to create a general state.

The problem starts when the two particles are identical because then they can be in the same energy eigenstates. Consider two electrons in the energy eigenstates of a potential well; say two electrons in orbit around a nucleus.

The two-particle basis becomes

$$|\alpha\rangle|\alpha\rangle, \quad |\alpha\rangle|\beta\rangle, \quad |\beta\rangle|\alpha\rangle, \quad |\beta\rangle|\beta\rangle. \quad (896)$$

What is the difference between particle 1 being in state $|\alpha\rangle$ and particle 2 being in state $|\beta\rangle$, and particle 1 been in state $|\beta\rangle$ and particle 2 being in state $|\alpha\rangle$?

Surely these are the same quantum mechanically. They represent the same experimental outcome. Why are they represented by different states?

The first and fourth states are clearly exchange symmetric, but the second and third states do not exhibit any particular particle exchange symmetry.

The symmetric and antisymmetric parts of the second state are

$$\begin{aligned} |s\rangle &= \frac{1}{\sqrt{2}}[|\alpha\rangle|\beta\rangle + |\beta\rangle|\alpha\rangle] \\ |a\rangle &= \frac{1}{\sqrt{2}}[|\alpha\rangle|\beta\rangle - |\beta\rangle|\alpha\rangle]. \end{aligned} \quad (897)$$

Then

$$\begin{aligned} |\alpha\rangle|\beta\rangle &= \frac{1}{\sqrt{2}}[|s\rangle + |a\rangle] \\ |\beta\rangle|\alpha\rangle &= \frac{1}{\sqrt{2}}[|s\rangle - |a\rangle]. \end{aligned} \quad (898)$$

A new basis containing the same number of degrees of freedom has been formed, but now each state has a definite particle exchange symmetry:

$$|\alpha\rangle|\alpha\rangle, \quad |s\rangle, \quad |a\rangle, \quad |\beta\rangle|\beta\rangle. \quad (899)$$

This basis comprises three symmetric states and one antisymmetric state.

Particles having symmetric state vectors under particle exchange are called *bosons*, and particles having anti-symmetric state vectors are called *fermions*.

Examples of bosons are photons, phonons, and magnons; examples of fermions are electrons, neutrons, and protons, and more fundamentally, quarks.

The *symmetrisation postulate* states that N identical particles are either exchange symmetric or exchange antisymmetric.

States with mixed symmetry do not exist. Symmetric and antisymmetric states refer to completely different kinds of object.

At the fundamental particle level, fermions relate to matter and bosons to forces.

It is now seen that the experimental interpretation of our two-particle two-state system, Eq. (899), is as follows:

- $|\alpha\rangle|\alpha\rangle$ — Two bosons in the low energy state.
- $|s\rangle$ — A boson in each of the low and high energy states, with no information about which one is which.
- $|a\rangle$ — A fermion in each of the low and high energy states, with no information about which one is which.
- $|\beta\rangle|\beta\rangle$ — Two bosons in the high energy state.

All information about which particle is which has been eliminated from the notation.

Suppose that the system is in the symmetric state $|s\rangle$, then the probability of finding boson 1 in state $|\alpha\rangle$ and boson 2 in state $|\beta\rangle$ is

$$\begin{aligned}
 P_{\alpha\beta} &= \left| [\langle\alpha|\langle\beta|] |s\rangle \right|^2 & (900) \\
 &= \left| \langle\alpha|\langle\beta| \frac{1}{\sqrt{2}} [|\alpha\rangle|\beta\rangle + |\beta\rangle|\alpha\rangle] \right|^2 \\
 &= \frac{1}{2}.
 \end{aligned}$$

The probability of finding boson 1 in state $|\beta\rangle$ and boson 2 in state $|\alpha\rangle$ is

$$\begin{aligned} P_{\beta\alpha} &= \left| [\langle\beta|\langle\alpha|] |s\rangle \right|^2 & (901) \\ &= \left| \langle\beta|\langle\alpha| \frac{1}{\sqrt{2}} [|\alpha\rangle|\beta\rangle + |\beta\rangle|\alpha\rangle] \right|^2 \\ &= \frac{1}{2}. \end{aligned}$$

There is maximum uncertainty about which particle is which, and the same applies to two fermions.

Each pair of identical particles has taken the form of an entangled pair, where it is not possible to associate a particle with a particular state.

All we know is what can be measured!

There is no antisymmetric wavefunction corresponding to two fermions being in the same state.

Bosons can share states, but fermions cannot.

The *Pauli Exclusion Principle* states that ‘no two identical fermions can be in the same single particle state’.

Two identical fermions cannot have the same set of quantum numbers.

Pauli’s exclusion principle has profound consequences, and in 1945 he was awarded the Nobel prize for his work.



Wolfgang Ernst Pauli
(1900-1958)

Remember that protons and neutrons are fermions also, and so two of them cannot share the same state. The theorem does not just apply to electrons.

As we add more and more electrons to an atom, it is only possible to place one electron in each state.

For an electron, with spin 1/2, it is possible to have two electrons in every state corresponding to some particular orbital and azimuthal angular momentum.

As more and more electrons are added, they can only build up in shells.

The rows of the Periodic Table correspond to different values of the principal quantum number n . The columns, moving left to right, correspond to filling up the angular momentum and spin states with electrons: one electron in each state.

The last column, corresponding to all of the available p states ($\ell = 1$) of the outer shell being filled, are particularly stable.

The chemical properties of the elements are largely determined by the degree to which the various shells are filled.

Collecting thoughts, for two particles in a two-state system:

- For dissimilar particles, $|\alpha_1\rangle|\alpha_2\rangle$, $|\alpha_1\rangle|\beta_2\rangle$, $|\beta_1\rangle|\alpha_2\rangle$, $|\beta_1\rangle|\beta_2\rangle$ is a suitable basis.
- For identical particles $|\alpha\rangle|\alpha\rangle$, $|s\rangle$, $|a\rangle$, $|\beta\rangle|\beta\rangle$ is a suitable basis.
- The new basis has the same number of degrees of freedom as the original.

- In making the particles identical, we eliminated the need for one degree of freedom, but this was then used to infer the existence of two fundamentally different kinds of particle.
- The new description maximises the uncertainty about which particle is which.
- One type of particles, fermions, can only exist as a superposition of states having different quantum numbers, but of course single fermions can exist.
- It seems as though there are two very specific forms of identical particle!

104 Spin statistics theorem

The *spin statistics theorem* states the following:

If identical particles have integer spin quantum numbers, $s = 0, 1, 2, 3, \dots$, the overall state must be exchange symmetric: bosons.

If identical particles have half-odd-integer spin quantum numbers, $s = 1/2, 3/2, 5/2, \dots$, the overall state must be exchange antisymmetric: fermions.

The reason the word ‘statistics’ enters the name is because the statistical behaviour of a collection of particles depends on whether the particles are distinguishable or indistinguishable.

Bosons obey Bose-Einstein statistics whereas fermions obey Fermi-Dirac statistics.



Enrico Fermi
(1901-1954)

For a collection of identical particles, the average number of bosons in a certain quantum state, j , is given by the Bose-Einstein statistics:

$$\langle n_j^b \rangle = \frac{1}{\exp[(E_j - \mu)/kT] - 1}, \quad (902)$$

which can be greater than unity, and is in fact unbounded.

The average number of fermions in a certain quantum state, j , is given by Fermi-Dirac statistics:

$$\langle n_j^f \rangle = \frac{1}{\exp[(E_j - \mu)/kT] + 1}, \quad (903)$$

which is always less than unity.

What happens when a ‘particle’ is in reality a collection of more elementary particles?

For example, a hydrogen atom can be regarded as a particle, whereas in reality it comprises an electron and a proton.

Composite particles have spin, where the overall spin is determined by adding the spins of the constituents.

If the total spin of the composite particle is an integer, the composite particle is exchange symmetric, and so it is a boson; whereas if the total spin of the composite particle is a half-odd-integer, the composite particle is exchange antisymmetric, and so it is a fermion.

It is remarkable that the statistics of a collection of identical, composite particles depends profoundly on whether the spins of the individual elementary particles add up to give an integer or half-integer spin quantum number.

The spin-statistics theorem is not just a feature of elementary particles themselves, but a more profound statement about the way matter behaves.

Consider the two isotopes, ^3He and ^4He , of the helium atom. Are they bosons or fermions? What about the hydrogen atom?

At this stage you may be wondering whether all of this manipulation can be possibly true, but remember that quantum mechanics was created over several decades to describe what is actually seen when experiments are performed.

It is a mathematical model to describe as best as possible the physical world as we witness it.

If you are at all sceptical about the symmetrisation postulate, note that the Pauli exclusion principle underpins a wealth of physical phenomena:

- The electronic structure of atoms, and therefore the form of the Periodic Table.
- The electronic properties of solids, and their electrical and thermal characteristics: insulators and conductors.
- The behaviour of ferromagnetic materials.
- The reason why White Dwarf stars do not gravitationally collapse: electrons are not able to occupy the same states, and so there is a limit to the degree to which they can occupy the same volume of space.



105 Multiple-particle exchange

To this point we have discussed particle exchange symmetry in the context of two particles in a two-level system, but we must extend our description to more complicated arrangements.

Consider N particles in a multi-state system, such as the electrons in an atom.

Each particle requires a list of quantum numbers to describe each of its possible states.

Use ξ_i as shorthand for the list of quantum numbers that identifies the state of particle i , say $(n, \ell, m_\ell, s, m_s)$.

Use $|\xi_i^r\rangle$ to denote the r 'th state of the i 'th particle.

$i \in \{1, \dots, N\}$ where N is the number of particles, and $r \in \{1, \dots, M\}$ where M is the number of states.

A canonical basis to describe the outcome of any experiment is then

$$|\xi_1^a\rangle \cdots |\xi_i^r\rangle \cdots |\xi_N^z\rangle \quad (904)$$

for all values of $a, \dots, z \in \{1, \dots, M\}$.

There are M^N such basis vectors.

If the particles are dissimilar, this basis is suitable for representing the outcomes of all possible measurements.

If the N particles are identical, then

$$|\xi_i^r\rangle = |\xi_j^r\rangle \quad \forall i, j, r, \quad (905)$$

and the basis would be

$$|\xi^a\rangle \cdots |\xi^r\rangle \cdots |\xi^z\rangle. \quad (906)$$

This can be appreciated by imagining the individual basis sets to be the eigenvectors of some measurement operator.

In Eq. (906), the particle is indicated by its position in the row.

Terms for which $r = s$ have a definite exchange symmetry when particles are swapped between states:

$$|\xi^a\rangle \cdots |\xi^r\rangle \cdots |\xi^r\rangle \cdots |\xi^z\rangle = |\xi^a\rangle \cdots |\xi^r\rangle \cdots |\xi^r\rangle \cdots |\xi^z\rangle, \quad (907)$$

but terms for which $r \neq s$ do not:

$$|\xi^a\rangle \cdots |\xi^r\rangle \cdots |\xi^s\rangle \cdots |\xi^z\rangle \neq |\xi^a\rangle \cdots |\xi^s\rangle \cdots |\xi^r\rangle \cdots |\xi^z\rangle. \quad (908)$$

Because the two sides of Eq. (908) must represent the same state, i.e., the same outcome when an experiment is performed, it is necessary to symmetrise or antisymmetrise the states, depending on the type of particles involved.

But if all of the particles are identical, how can we do this?

Let us consider first the symmetric case, for bosons. It is possible to create exchange symmetric states by combining all possible particle permutations and renormalising.

For each combination of a, \dots, r, \dots, z , of which there are M^N , one has to form the exchange symmetric vector

$$|a, \dots, r, \dots, z\rangle^s = \sqrt{\frac{\Pi_n(m_n!)}{N!}} \sum_P |\xi^a\rangle \cdots |\xi^r\rangle \cdots |\xi^z\rangle, \quad (909)$$

where \sum_P corresponds to a sum over all available unique pair-wise particle permutations acting on the N particles.

There are $N!$ of them.

m_n in the normalisation factor corresponds to the number of times the n 'th state appears in the Cartesian product.

Eq. (909) generates a unique set of normalised vectors, which are symmetric with respect to particle exchange.

For example, consider three particles in a three-state system.

A canonical basis vector is

$$|\xi^a\rangle|\xi^b\rangle|\xi^c\rangle \quad a, b, c \in \{1, 2, 3\}. \quad (910)$$

The lowest order state,

$$|1, 1, 1\rangle, \quad (911)$$

is exchange symmetric, and corresponds to finding all three bosons in state 1.

The next state,

$$|1, 1, 2\rangle, \quad (912)$$

does not have any obvious exchange symmetry $|1, 1, 2\rangle \neq \pm|1, 2, 1\rangle$, etc., and so we should symmetrise it using (909).

We find that

$$|1, 1, 2\rangle^s = \frac{1}{\sqrt{3}} [|1, 1, 2\rangle + |1, 2, 1\rangle + |2, 1, 1\rangle], \quad (913)$$

which is an entangled state involving all three particle permutations where there are two particles in state 1 and one particle in state 2.

Eq. (913) is normalised, and moreover the probability of finding each possible arrangement is the same:

$$\begin{aligned} |\langle 1, 1, 2 | 1, 1, 2 \rangle^s|^2 &= \frac{1}{3} \\ |\langle 1, 2, 1 | 1, 1, 2 \rangle^s|^2 &= \frac{1}{3} \\ |\langle 2, 1, 1 | 1, 1, 2 \rangle^s|^2 &= \frac{1}{3}. \end{aligned} \quad (914)$$

The situation is shown in Fig. 80. Notice that the labelling of any two particles can be swapped without changing the state (dashed lines):

$$|1, 1, 2\rangle^s = |1, 2, 1\rangle^s = |2, 1, 1\rangle^s. \quad (915)$$

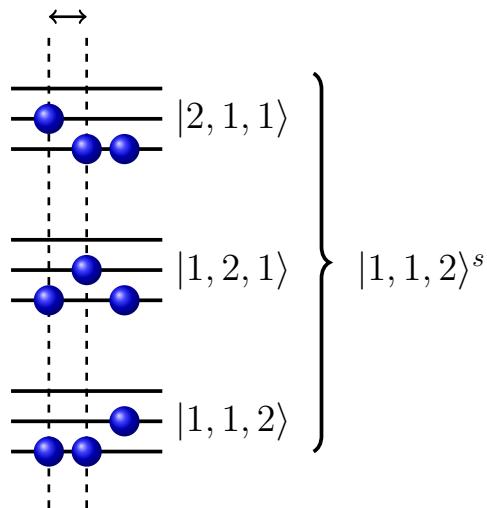


Figure 80: Three identical bosons in a three-state system. Symmetrising $|1, 1, 2\rangle$ leads to entangling the three states shown.

It is then necessary to move to the next unsymmetrised state, say $|1, 2, 2\rangle$ and carry out the procedure again.

Each unique combination gets symmetrised until we have created a complete system of symmetric states.

We can also form antisymmetric states, which correspond to possible arrangements of fermions.

For each combination of a, \dots, r, \dots, z , let us form the exchange antisymmetric vector

$$|a, \dots, r, \dots, z\rangle^a = \frac{1}{\sqrt{N!}} \sum_P (-1)^P |\xi^a\rangle \cdots |\xi^r\rangle \cdots |\xi^z\rangle, \quad (916)$$

where again \sum_P corresponds to the sum over all available unique pair-wise particle permutations acting on the N particles.

$(-1)^P = 1$ for an even number of permutations, and $(-1)^P = -1$ for an odd number of permutations.

Now, however, we can only form antisymmetrised states from arrangements for which a, \dots, r, \dots, z are all different.

Make sure that you find all $N!$ permutations.

For three particles in a three-state system, only one antisymmetric state is available:

$$|1, 2, 3\rangle^a = \frac{1}{\sqrt{3!}} [|1, 2, 3\rangle - |1, 3, 2\rangle + |2, 3, 1\rangle - |2, 1, 3\rangle + |3, 1, 2\rangle - |3, 2, 1\rangle], \quad (917)$$

which is more complicated.

The various states are shown in Fig. 81. Notice that when the labelling of two particles is swapped, illustrated by the dashed lines, the state changes sign:

$$|1, 2, 3\rangle^a = -|1, 3, 2\rangle^a = |2, 3, 1\rangle^a, \text{ etc.} \quad (918)$$

It is clear that there are fewer ways of arranging fermions than bosons, and this accounts for the different statistics.

Once the symmetrisation and antisymmetrisation have been completed, a new basis has been formed.

It allows for two distinct classes of particles, bosons and fermions, and it prevents to have more than one fermion in any quantum state, thus satisfying Pauli's exclusion principle.

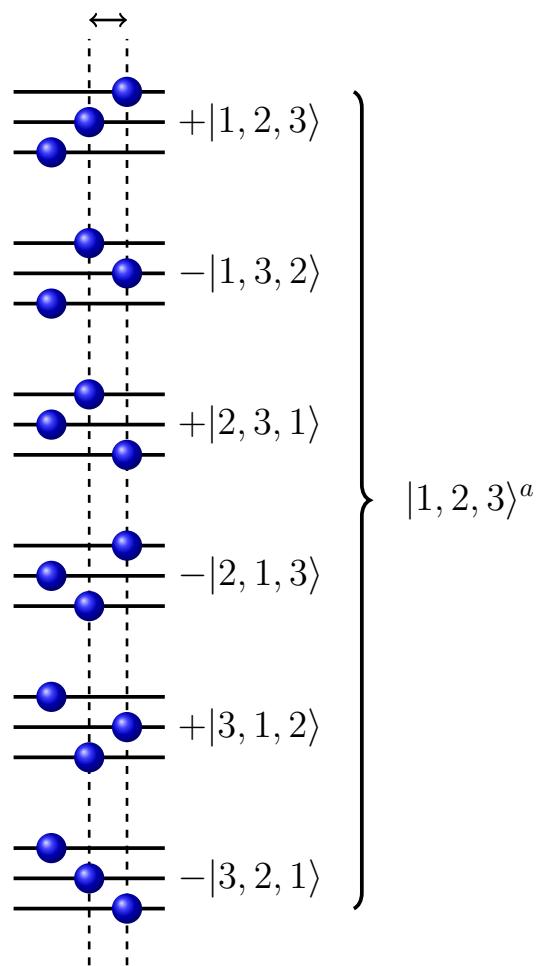


Figure 81: Three identical fermions in a three-state system. Antisymmetrising $|1, 2, 3\rangle$ leads to entangling the six states shown.

Consider two particles in a three-state system. If the single-particle states are labelled a , b and c , what bosonic states are available and what fermionic are available? Draw a three-level diagram to illustrate these arrangements. How many different ways are there of arranging two bosons in the system, and how many different ways are there of arranging two fermions?

Why cannot we fully symmetrise / antisymmetrise the state vectors corresponding to three particles in a two-state system? Try it!

The fermionic state corresponding to N identical parti-

cles in N -states can be represented in the following way

$$|N\rangle^a = \frac{1}{\sqrt{N!}} \begin{vmatrix} |\xi^1\rangle & |\xi^1\rangle & \cdots & |\xi^1\rangle \\ |\xi^2\rangle & |\xi^2\rangle & \cdots & |\xi^2\rangle \\ \vdots & \vdots & & \vdots \\ |\xi^N\rangle & |\xi^N\rangle & \cdots & |\xi^N\rangle \end{vmatrix}, \quad (919)$$

which is known as the *Slater determinant*.

When wavefunctions are used, it becomes a convenient numerical tool.

Each column corresponds to a particle, and each row to a single-particle state that a particle can be in.

By swapping columns we swap particles, and this results in the sign of the determinant changing, as expected.

By swapping rows we swap states, and this also results in the sign of the determinant changing, as expected.

All N fermions must exist in different states, whereas all of the bosons can, at sufficiently low temperatures, exist in the same state, which is called a *Bose-Einstein condensate*.



Satyendra Nath Bose
(1894-1974)

For example, if N bosons all condense into the ground state, the overall state vector would be

$$|\xi^1\rangle \cdots |\xi^1\rangle \cdots |\xi^1\rangle. \quad (920)$$

Consider two spin-half particles, such as electrons. A generic basis would be

$$|\uparrow\rangle|\uparrow\rangle, |\uparrow\rangle|\downarrow\rangle, |\downarrow\rangle|\uparrow\rangle, |\downarrow\rangle|\downarrow\rangle, \quad (921)$$

but this does not reveal spin degeneracies.

Create a new basis that explicitly exhibits exchange symmetry.

Why can't two electrons having the same spin exist in a single spatial quantum state, whereas two electrons having different spin can?

XIV Summary

In these notes, we have seen how mounting experimental evidence at the beginning of the 20th century forced scientists to re-think their understanding of the world in terms of distinct wave and particle concepts, and to look for a unified description that encompasses both. This lead to the formulation of wave-particle duality and quantum mechanics.

We discussed, at a phenomenological level, some of the key concepts that one must introduce in order to formulate a new description of reality that encompasses the wave-particle duality required by quantum mechanics. In particular, we focused in the first instance on the notion of wave function and its meaning in regards to physical properties of a system and the outcome of measurements.

We introduced a wave equation that describes the time evolution of a wave function, and we discussed its understanding in terms of energy and momentum operators. Moreover, we introduced the concept of probability current and the corresponding continuity equation.

We began our investigation of how one solves the Schrödinger equation, focusing in the first instance on the case of 1D unbound particles in (piece-wise) constant potentials. We discussed the physical significance of the results and remarked on the difference between classical and quantum behaviour, and we encountered the exquisitely quantum mechanical concept of tunnelling.

We continued the study by turning our attention to the behaviour of bound particles. We discussed the 1D simple harmonic oscillator, and the notions of zero point

energy and of parity of the wavefunction. A correspondence principle between classical and quantum mechanics was also formulated.

In this first part of the notes, we developed an initial understanding of quantum mechanics, starting from the fundamental physics questions that it aimed to address and continuing to propose a formalism based on wave functions and the Schrödinger equation. We then ventured into solving a range of problems concerning the motion of a particle in 1D, encompassing piece-wise constant potentials and the simple harmonic oscillator. In the process, we encountered and discussed several new and important concepts brought about by quantum mechanics.

We then moved beyond the wave function formulation and introduced a new, more powerful description based on vector spaces and operators – a description that has come to take centre stage and is widely used in modern physics. In this language, we revisited the notion of position, momentum and energy operators, and discussed the concepts of eigenvalues and eigenvectors, of expectation values and the uncertainty principle. The full force of quantum mechanics only emerges when the tools of linear operators are brought to bear. Operator algebra is central to quantum mechanics. Its advanced forms pervade many concepts that are now regarded as ‘physical reality’.

We demonstrated the power and importance of the operator formulation of quantum mechanics by considering again the simple harmonic oscillator, to be contrasted with the Schrödinger description in earlier chapters, leading to the discussion of ladder operators. The quantisation of the harmonic oscillator is central to quantum mechanics because it highlights the relationship between particle-like and wave-like behaviour. Ultimately, it defines what one means by elementary excitations

such as photons, phonons, magnons, etc., and it provides an elegant scheme for describing their behaviour and interactions. We closed this part of the notes by introducing the notion of density operator and by discussing how to form functions of operators.

We introduced two formulations of quantum mechanics, in terms of Schrödinger's equation and wave functions, and in terms of vector spaces and operators. In both cases, time is a parameter. For example, the operator algebra we have discussed applies at an instant in time, and at every time. We can only calculate the probability of observing some outcome at a chosen time t .

Next, we turned our attention to time evolution in quantum mechanics. We started from the so-called *Schrödinger picture*, where it is quantum states and not observables (operators) that change with time. We introduced the concept of time translation operator and Ehrenfest's theorem describing the evolution of the expectation value of operators. We also discussed the connection to classical mechanics and the time-energy uncertainty principle, and we gave an example involving a particle in an infinite well. We also introduced the so-called *Heisenberg picture*, where quantum states do not change with time, but operators do; in many ways, this is closer to the classical viewpoint, where position and momentum are dynamical variables that vary with time in familiar ways. We finally gave an example involving the density operator.

Up to this point, we considered 1D systems, for simplicity. While this allowed us to introduce some of the main concepts and formalisms used in quantum mechanics, it also had substantial limitations – for example, there is no notion of angular momentum in 1D. In order to progress in our understanding of quantum mechanics, we had to expand our discussion to higher dimensions, and specifically to 3D systems.

Through the investigation of quantum mechanical systems in 3D, we introduced the notion of angular momentum and we saw how it plays a fundamental part in quantum mechanics. We defined the angular momentum operator and derived its properties, commutation and uncertainty relations, eigenstates and eigenvalues. We encountered again ladder operators. And we further discussed the connection between angular momentum and magnetic moment, and gave detailed examples such as diatomic molecules and the Hydrogen-like atom.

This concluded the part of the lecture notes dedicated to single particle mechanics. The rest of the course was dedicated to multi-particle / multi-state systems. This set the stage to discuss the notion of spin and total angular momentum, as well as the role of interactions and identical particles.

We started going beyond single particle mechanics by considering a system composed of two particles, and how one can describe the states and their physical properties, and how to solve the quantum two-body problem. This lead to the important concept of entangled states and spooky action at a distance.

We introduced the notion of intrinsic angular momentum, or spin, of a particle. We saw how one can combine the spin and orbital angular momentum into the total angular momentum, and how this can be formulated in terms of eigenstates and eigenvalues, and composite wave functions. We also studied how to combine spins of different particles.

We then encountered another iconic example of a quantum mechanical system that requires the combination of different states: the two level system. We described it using the vector and operator-matrix formalism and presented a complete perturbative study. This lead us to introduce the notions of overlap integrals and level splitting / repulsion, typical of covalent / ionic bonds

between atoms. The notes also included two asides that illustrate simple applications of the matrix representation to orbital and spin angular momentum, and two further non-examinable asides on spin matrices and density matrices.

Finally, the lecture notes closed by returning to the description of systems comprising more than one particle, paying particular attention to the situation where the particles are *identical*. We encountered the concept of exchange symmetry and the spin statistics theorem, whereby particles at the quantum mechanical level can either be symmetric upon exchange (bosons) or antisymmetric (fermions). We finished with a brief discussion of how this affects multi-particle states.