

K. GOLDSTEIN

QUANTUM MECHANICS WORKBOOK



IMAGE SOURCE: [HTTPS://WWW.FLICKR.COM/PHOTOS/29233640@N07/8132455446](https://www.flickr.com/photos/29233640@N07/8132455446)

Copyright © 2025 K. Goldstein

TYPESET WITH TUFTE-LATEX

Contents

1	<i>Stern-Gerlach Experiments</i>	9
2	<i>Simplified Stern-Gerlach experiments</i>	16
3	<i>Quantum State Vectors</i>	24
4	<i>Stern-Gerlach re-examined</i>	33
5	<i>New notation & Generalisations</i>	42
6	<i>Operators and Measurement</i>	50
7	<i>New operators</i>	61
8	<i>New operators (continued...)</i>	67
9	<i>Hermitian Operators</i>	73
10	<i>Measurement, Observables and Uncertainty</i>	78
11	<i>S^2, Spin-1 and General systems</i>	88
12	<i>Schrödinger Time Evolution</i>	94
13	<i>Spin precession</i>	100
14	<i>Magnetic Field in a (more) general direction</i>	107
15	<i>Neutrino Oscillations</i>	113
16	<i>Time Dependent Hamiltonians</i>	117
17	<i>Quantum Spookiness</i>	125
18	<i>Spectroscopy & the Energy Eigenvalue equation</i>	135
19	<i>The wave function</i>	140
20	<i>Stationary states</i>	150
21	<i>Infinite Square well</i>	157
22	<i>Finite square well</i>	165
23	<i>Review & Superposition states</i>	173
24	<i>Superposition states</i>	181
25	<i>Harmonic Oscillator I</i>	190
26	<i>Harmonic Oscillator II</i>	195
27	<i>Harmonic Oscillator III</i>	203
28	<i>Unbound states – Free Particle</i>	209
29	<i>Continuous bases</i>	218
30	<i>Momentum space and uncertainty</i>	227
31	<i>Scattering and tunnelling</i>	234
32	<i>Quantum Mechanics in 3D</i>	242
33	<i>Angular momentum</i>	248
34	<i>Angular momentum and Spherical Harmonics</i>	254
35	<i>Hydrogen atom</i>	260
36	<i>Hydrogen atom II</i>	269
A	<i>Review of linear algebra I</i>	276
B	<i>Review of linear algebra II</i>	294
	<i>Bibliography</i>	304

Preface

This is a workbook based introduction to Quantum Mechanics. You will not get more out this book than you put into it. Most of the book is based on “Quantum Mechanics: A Paradigms Approach” by McIntyre¹ with a few sections, namely the harmonic oscillator, 3d quantum mechanics and part of the hydrogen atom, based on “Introduction to Quantum mechanics” by Griffiths².

Workbook Features

The workbook is broken up into a series of Readings. Before coming to class, you have to work through assigned reading material for that session – you can read more about the reading assignments in the next section.

In addition to the text, each Reading will contain various kinds of boxes.

Boxes of emphasis

Important Box

Particularly important formulae or text may appear in boxes like this one to draw your attention to them.

Discussion boxes

Each reading contains discussion boxes like the one below. When you get to a discussion box in the text, you should stop and attempt to answer it. You can discuss the boxes with your peers online and try answer them together. Don’t worry if you don’t manage to completely answer every discussion box before coming class but you **must** at least make an attempt. By the time a test or exam comes round you should have answers to each discussion box that you are happy with in your notes.

¹ D. H. McIntyre, C. A. Manogue, and J. Tate. *Quantum mechanics: a paradigms approach*, volume 192. Pearson Boston, 2012

² D. J. Griffiths. *Introduction to quantum mechanics*. Cambridge University Press, 2016

This is a workbook not a textbook.

Discussion boxes encourage active reading

 **Discussion 0.1**

What do you think the role of the discussion boxes is? How can they help you learn?

Examples

 **Example 0.1:**

Examples, which usually illustrate how to use techniques discussed in the text will often appear in boxes like this one. You should work through and understand the details of these examples.

Reading Assignments

The reading assignments help you **learn faster** by collaboratively annotating the readings and communicating with your classmates. Collaboration gets you help whenever you need it, makes learning more fun, enables you to help others (which research shows is also a great way for you to learn), and helps the instructor make class better by emphasizing information that you need.

If you have a question or information to share about a passage in the readings, highlight the text and type in a comment as an annotation. You can also respond to a classmate's annotation in threads (Facebook style) in real time or upvote questions you find helpful. Good annotations contribute to the class by stimulating discussion, explaining your thought processes, helping others, and drawing attention to good points. If a particular classmate's point is relevant, you can explicitly "mention" them and they will be immediately notified, even if not presently signed on.

Research shows that the following reading assignment behaviours predict higher end-of-course marks and long term mastery of the subject.

If you would like to understand more about the reasons for the way this course is structured, you can watch [What Research on Learning Tells Us About Teaching Physics](#)

Reading assignments help you learn.

Supplementary material

Enrichment

Supplementary material may appear in green boxes like this one. Such material is included for your interest and may introduce things you will learn about in future courses but will **not** be examined in this course.

This will not be in the exam.

Margins

You may notice that the text has particularly wide margins. The idea is that you can use these margins to annotate your own hard copies with notes and questions.

There are also marginal notes which are intended to guide you and help with locating ideas.

[This is a margin guide.](#)

Links

The pdf version of the notes has links to supplementary material. For those of you who print out the notes, I have included qrcodes³ to access these links. You can also click on the qrcode in the pdfs to go to the link.



³

Prologue

It was a dark and stormy night. Erwin huddled under his covers as he had done numerous times that summer. As the wind and rain lashed at the window, he feared having to retreat to the storm cellar once again. The residents of Erwin's apartment building sought shelter whenever there were threats of tornadoes in the area. While it was safe down there, Erwin feared the ridicule he would face once again from the other school boys. In the rush to the cellar, Erwin seemed to always end up with a random pair of socks, and the other boys teased him about it mercilessly. Not that Erwin hadn't tried hard to solve this problem. He had a very simple collection of socks—black or white, for either school or play; short or long, for either trousers or shorts. After the first few teasing episodes from the other boys, Erwin had sorted his socks into two separate drawers. He placed all the black socks in one drawer and all the white socks in another drawer.

Erwin figured he could determine an individual sock's length in the dark of night simply by feeling it, but he had to have them presorted into white and black because the apartment generally lost power before the call to the shelter.

Unfortunately, Erwin found that this presorting of the socks by color was ineffective. Whenever he reached into the white sock drawer and chose two long socks, or two short socks, there was a 50% probability of any one sock being black or white. The results from the black sock drawer were the same. The socks seemed to have "forgotten" the color that Erwin had determined previously.

Erwin also tried sorting the socks into two drawers based upon their length, without regard to color. When he chose black or white socks from these long and short drawers, the socks had also "forgotten" whether they were long or short.

After these fruitless attempts to solve his problem through experiments, Erwin decided to save himself the fashion embarrassment, and he replaced his sock collection with a set of medium length brown socks. However, he continued to ponder the mysteries of the socks throughout his childhood.

After many years of daydreaming about the mystery socks, Erwin Schrödinger proposed his theory of "Quantum Socks" and became famous. And that is the beginning of the story of the quantum socks.

The End.

Farfetched?? You bet. But Erwin's adventure with his socks is the way quantum mechanics works. Read on.

– McIntyre et al. [2012, Prologue]



"Erwin"

Original image by Hans-Michael Tappen, CC BY-NC-SA 2.0

1

Stern-Gerlach Experiments

Quantum mechanics is a new (and absolutely necessary) way of predicting the behaviour of microscopic objects. In quantum mechanics, aspects of the world essentially depend on chance and can not be predicted. Furthermore, we can not always know everything about a quantum system – measuring one quantity may lead to ignorance about another. In addition, certain quantities, that are classically continuous, may only take on discrete values. Examples of such quantities are energy, orbital angular momentum and intrinsic angular momentum (or spin).

Quantum particles behave as mysteriously as Erwin's socks – sometimes forgetting what we have already measured. Erwin's adventure with the mystery socks is farfetched because you know that everyday objects do not behave like his socks. If you observe a sock to be black, it remains black no matter what other properties of the sock you observe. However, the Stern-Gerlach experiment goes against these ideas. Microscopic or quantum particles do not behave like the classical objects of your everyday experience. The act of observing a quantum particle affects its measurable properties in a way that is foreign to our classical experience.

– McIntyre et al. [2012, Chapter 1]

Stern-Gerlach experiment – overview

The Stern-Gerlach experiment is a very important in quantum mechanics. We can learn a lot about the essential features of quantum mechanics and how it differs from the expectations we may have had from classical physics.

The simple version of the experiment has 3 essential parts shown in Figure 1.1:

1. An **oven**¹ which produces electrically neutral atoms.
2. A **magnet** producing an *inhomogeneous* magnetic field that may deflect the atoms
3. A **detector** for the atoms.

We will start by looking at a Stern-Gerlach experiment with a beam of *silver* atoms. In this case, the initial beam is split into two beams by the magnet. One beam is deflected upwards and the

This reading is based on [McIntyre et al. 2012, Section 1.1]

Brief introduction to QM.

Stern-Gerlach setup

¹ The collimator in Figure 1.1 is part of the oven. It narrows the beam of atoms coming out of the oven ensuring that they are all roughly travelling in the same direction. Collimated means "made parallel".

The results of the S-G experiment contradict our classical expectation.

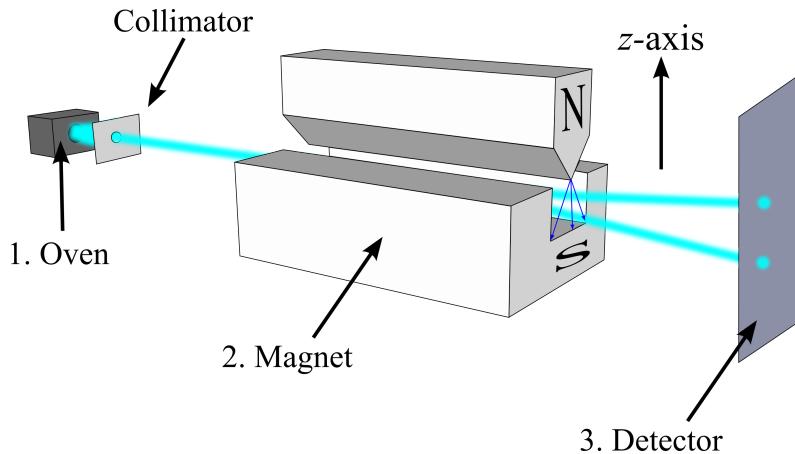


Figure 1.1: Set-up of the Stern-Gerlach experiment to measure the spin component of neutral particles along the z -axis.

Modified from an image of Tatoute, CC BY-SA 4.0, via Wikimedia Commons, [Image source](#)

other is deflected downwards. To appreciate why this is significant we first need to understand why this is **so** different from what we expect classically. This is discussed in the next section.

Classical Stern-Gerlach

We need to understand the force exerted on a *neutral* atom deflected by a magnetic field. Suppose that our atom has a magnetic moment, $\vec{\mu}$. Remember, the magnetic moment, $\vec{\mu}$, tells us about the direction and strength of a magnet. A loop of current, I , with area, A , has a magnetic moment with a magnitude $\mu = IA$. The direction of the magnetic moment of a current loop is determined by the right hand rule as shown in Figure 1.2.

Using the fact that the potential energy of a magnet in an external magnetic field is

$$V = -\vec{\mu} \cdot \vec{B}, \quad (1.1)$$

and using relationship between force and potential energy,

$$\vec{F} = -\vec{\nabla}V, \quad (1.2)$$

we conclude that the force on a magnetic moment in a magnetic field is,

$$\vec{F} = \vec{\nabla}(\vec{\mu} \cdot \vec{B}), \quad (1.3)$$

which shows that we need an *inhomogeneous*² magnetic field for an object with a constant magnetic moment to experience a force.

In the basic Stern-Gerlach experiment, as shown in Figure 1.1, the variation of the field in the z -direction results in a deflection in that direction. Calculating the force in the z -direction one finds

$$F_z = \frac{\partial}{\partial z} (\vec{\mu} \cdot \vec{B}) \approx \mu_z \frac{\partial B_z}{\partial z}, \quad (1.4)$$

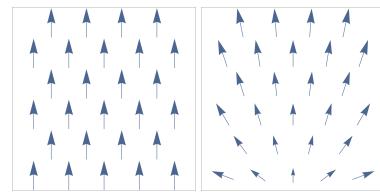
where we have assumed that the variation of the other components of \vec{B} are negligible in the z -direction³. We see that the force is proportional to the variation of the magnetic field in the z -direction. Also note that whether the force is up or down depends on sign of



Figure 1.2: Right hand rule.

Image attribution: original uploader Schorschiz at German Wikipedia, derivative work Wizard191, Public domain, via Wikimedia Commons, [Image source](#)

² Homogeneous, in this context, means the same at every point whereas an inhomogeneous field varies as a function of position. For example the first image below represents a homogeneous vector field whereas the second is inhomogeneous.



Mathematica codes for figures:

```
VectorPlot[{0, 1}, {x, -3, 3}, {y, 0.1, 3}, 
VectorColorFunction -> None, FrameTicks -> None, 
VectorPoints -> 5, VectorSizes -> {1, 2}]; 
VectorPlot[{x, 1.5 y}, {x, -3, 3}, {y, 0.1, 3}, 
VectorColorFunction -> None, VectorScaling -> Automatic, 
VectorSizes -> {1, 2}, FrameTicks -> None, 
VectorPoints -> 5]
```

³ There is a subtlety relating to the other components of the force, (1.3), that we are glossing over here. We may return to this issue once we have discussed Larmor precession in Section 13.

μ_z . Since the atoms are travelling perpendicular to the variation of the magnetic field, the deflection is perpendicular to the direction of travel.

💬 Discussion 1.1

Explain how we specialise $\vec{F} = \vec{\nabla}(\vec{\mu} \cdot \vec{B})$ to $F_z = \frac{\partial}{\partial z} (\vec{\mu} \cdot \vec{B})$.

💬 Discussion 1.2

Explain how we get the second equality in
 $F_z = \frac{\partial}{\partial z} (\vec{\mu} \cdot \vec{B}) \approx \mu_z \frac{\partial B_z}{\partial z}$.

💬 Discussion 1.3

Why is it important that we have an **inhomogeneous** magnetic field in the Stern-Gerlach experiment?

We have seen that objects with a magnetic moment are deflected by an inhomogeneous magnetic field but why do atoms have a magnetic moment at all? Classically, the magnetic moment of an atom may be due to the motion of electrons around the nucleus which constitute a current.

A loop of current, I , of area A produces a magnetic moment, $\mu = IA$. So, for a charged particle q , moving at speed v in a circle of radius r ,

$$\mu = IA = \frac{q}{2\pi r/v} (\pi r^2) = \frac{qvr}{2} = \frac{q}{2m} L, \quad (1.5)$$

where $L = mvr$ is the angular momentum of the electron.

classical perspective of atomic magnetic moments

💬 Discussion 1.4

Explain the second equality in (1.5).

Just like the earth revolves on its own axis as well as around the sun, an electron in an atom can have two kinds of angular momentum. The first kind, namely, **orbital angular momentum**, \vec{L} , due to its motion around the nucleus, has already been mentioned. The second kind is **intrinsic spin angular momentum**, \vec{S} , which is also just called the spin. In what follows, we will often refer to orbital angular momentum as just **angular momentum** and intrinsic spin angular momentum as just **spin**. Spin is due to motion of particle about its own centre of mass. Figure 1.3 shows Earth's orbital and spin angular momentum.

Orbital vs spin angular momentum

For a charged extended object, spin creates current loops, which in turn produce a magnetic moment. Finding the relationship between spin and magnetic moment for a spinning charged sphere requires a rather technical integration over the moving charge

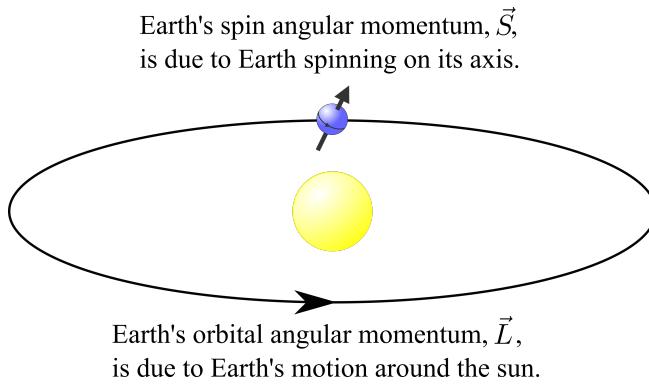


Figure 1.3: Schematic illustration of Earth's spin and orbital angular momentum. Note diagram is not to scale.

distribution which we will not look at explicitly in this course. We expect a roughly similar relationship between the magnetic moment due to a particle's spin and the relationship between orbital angular momentum and the corresponding magnetic moment, (1.5). Writing (1.5) as a vector equation gives

$$\vec{\mu}_{\text{due to } \vec{L}} = g \frac{q}{2m} \vec{L}, \quad (1.6)$$

where the **gyromagnetic ratio**, g , is some dimensionless number, which classically is related to the geometry of our current distribution. As we saw in the derivation of (1.5), for a simple current loop, $g = 1$.

The gyromagnetic ratio is a proportionality constant relating angular momentum and magnetic moment.

Discussion 1.5

The dimensions of magnetic moment are $[\mu] = [IA] = QT^{-1}L^2$ where Q denotes the units of charge. Spin has units of angular momentum, so that $[\vec{S}] = [\vec{r} \times \vec{p}] = L^2MT^{-1}$. Show that if we assume $\mu \propto q^\alpha S^\beta m^\gamma$, dimensional analysis implies $\alpha = 1$, $\beta = 1$ and $\gamma = -1$ or in other words

$$\mu \propto \frac{q}{m} S.$$

By analogy with (1.6), we write the dimensionless proportionality constant as $g/2$, with g referred to as the gyromagnetic ratio.

Using Discussion 1.5, the relationship between spin and the magnetic moment is:

$$\vec{\mu}_{\text{due to } \vec{S}} = g \frac{q}{2m} \vec{S}, \quad (1.7)$$

where g is the **gyromagnetic ratio** that can be found, as mentioned above, by integrating over all the current loops that result from the rotation of the charged object.

We will now consider a Stern-Gerlach experiment performed with silver atoms. Silver is chosen because its net magnetic moment is essentially due to the spin of just one unpaired electron.

The spin of a silver atom is essentially the same as the spin of a single electron.

Enrichment

Unfortunately to understand why the main contribution to the magnetic moment of silver atoms comes from a single electron requires a bit of knowledge about atomic quantum mechanics that we will only get to by the end of the course. Fortunately, most of the essential features of quantum mechanics we seek to learn about from the experiment do not depend on this fact. Now, with an atomic number of 47, a silver atom has 47 electrons and 47 protons. Looking back at (1.5), we see that the magnitude of the magnetic moment is inversely proportional to the mass. Since protons are about 2000 times heavier than electrons they have much less effect on an atoms magnetic moment. This means that the dominant contribution to the atoms magnetic moment comes from the electrons. You may have learnt in a chemistry course that silvers electronic configuration is: $1s^2 2s^2 3s^2 3p^6 4p^6 4d^{10} 5s^1$ (see [Silver electronic configuration](#)). The important point to notice is that there is **one unpaired** 5s electron outside the closed shells which makes main contribution to the atom's magnetic moment. This electron is in an s-state which is spherically symmetric and has no orbital angular momentum but it does have **spin** which contributes to silver's magnetic moment.

In conclusion, for silver

$$\vec{\mu} \approx -g \frac{e}{2m_e} \vec{S}, \quad (1.8)$$

where e is the charge of the electron and m_e is the mass of the electron.

Silver electronic configuration:



Discussion 1.6

Why do you think we need to use *neutral* silver atoms rather than a beam of electrons for the S-G experiment?

We are now in a position to understand why the results of the Stern-Gerlach experiment differ from what one expects from classical physics. Substituting (1.4) into (1.8) leads to

$$F_z \approx -g \frac{e}{2m_e} S_z \frac{\partial B_z}{\partial z}. \quad (1.9)$$

The equation, (1.9), tells us that the deflection of the beam is proportional to the z -component of the spin, S_z . So, assuming the magnitude of the spin, $|\vec{S}|$, of each unpaired electron of each atom is the same, the deflection of any atom is proportional to $S_z = |\vec{S}| \cos \theta$. Here θ is the angle between the z -axis and the spin, \vec{S} , of the electron in that atom. Now, since the atoms come out of a hot oven, we expect a random distribution of spins for the atoms. This means that whatever the spins are, we expect the z -component to vary

continuously from $S_z = -|\vec{S}|$ to $S_z = |\vec{S}|$. Looking at (1.9), this in turn tells us that we expect a continuous range of forces acting on the atoms in the beam so we expect a continuous range of deflections which then would give a **continuous** spread of silver atoms on the screen.

Contrary to the classical expectation outlined above, instead of a **continuous** spread of atoms on the screen, only **two** deflections are observed (see Fig. 1.4 below). This surprisingly tells us that there are only two possible values of S_z . The magnitudes of the deflections in the experiment are consistent with

$$S_z = \pm \frac{1}{2}\hbar ,$$

where $\hbar = h/2\pi$, h is Plank's constant, and

$$\begin{aligned} \hbar &= 1.0546 \times 10^{-34} \text{ J}\cdot\text{s} \\ &= 6.5821 \times 10^{-16} \text{ eV}\cdot\text{s} . \end{aligned} \quad (1.10)$$

Classically expect continuous spread spread of atoms.

Stern-Gerlach experimental results

Discussion 1.7

Show that \hbar has the same units as angular momentum,
 $\vec{L} = \vec{r} \times \vec{p}$.

The experiment is evidence that the electron's spin is **quantised** or in other word's that the electron's spin can only have certain discrete values. As previously mentioned, this is not what we expect classically. Because of the factor of $\frac{1}{2}$ in, $S_z = \pm \frac{1}{2}\hbar$, we call this a **spin-½** system. Notice that the z-axis is not special –we could have chosen any other axis (or indeed simply labelled the axes differently) along which to do the experiment and would get the same result.

After having worked through these notes, you may want to watch the videos below which also explain the experiment:

- Stern-Gerlach Experiment
- Video explaining the Stern-Gerlach Experiment

The S-G experiment shows that the electron's spin is quantised.

Videos

Discussion 1.8

The results of the original Stern-Gerlach experiment are shown in Figure 1.4. The first image on the left shows the result when there is no magnetic field – the beam of silver atoms is undeflected and we just get single line. The second image on the right shows the result when there is an inhomogeneous magnetic field.

In class we will discuss the following questions that you can think about:

- What would you expect the results to look like in the presence of a magnetic field if the spin of the electron was

not quantised?

- **Challenge:** can you explain why, rather than two distinct bands, we get a slit shape in the second picture. (**Hint:** think about the nature of the inhomogeneous magnetic field.)

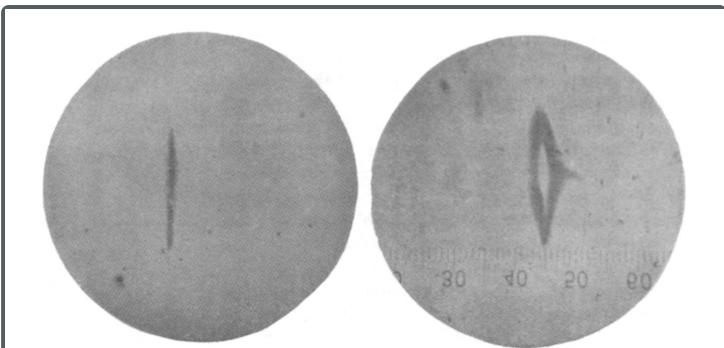


Figure 1.4: The image above shows the pattern created by a ray of silver atoms in the original Stern-Gerlach experiment. On the left is the result without a magnetic field while the right we see what happens when there is a magnetic field.

Image source: Gerlach and Stern [1922]

💬 Discussion 1.9

Write a summary of the important points in this chapter.

2

Simplified Stern-Gerlach experiments

Give me a sign
Hit me, baby, one more time

– Britney Spears¹

Continuing on from the unexpected results in the last chapter, we are going to take a look at generalisations of the Stern-Gerlach experiment. As the quote suggests, we will be subjected to a series of blows to our classical “common sense”.

In what follows we will not worry about the details of the Stern-Gerlach experiment but work with simplified schematic experiments as shown in Figure 2.1

This reading is based on [McIntyre et al. 2012, Section 1.1]

¹ actually written by Max Martin. The phrase “Hit me” is really meant to be slang for “Call me”.

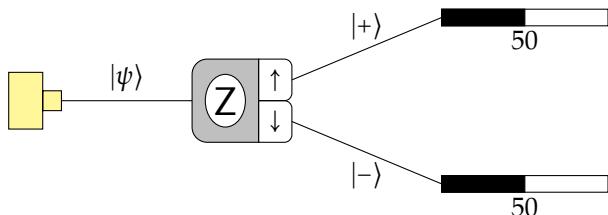


Figure 2.1: Simplified schematic of the Stern-Gerlach experiment, depicting a source of atoms, a Stern-Gerlach analyzer, and two counters.

There are 3 basic parts of Figure 2.1:

[Stern-Gerlach schematic](#)

1. An **oven** that produces the atoms:

2. A **Stern-Gerlach device** :

The device has two output ports for the two possible values of the spin component. It is labelled with the axis along which the magnetic field gradient is oriented. The up and down arrows on the output ports represent the two possible measurements for the device. For a device oriented along the z -axis the up and down arrows correspond to the results $S_z = \pm \hbar/2$ respectively whereas for a device oriented along the y -axis they would denote $S_y = \pm \hbar/2$. No matter what direction we align the analyser along, we still only get two possible outcomes in our spin $\frac{1}{2}$ system. The two possible results are called **spin up** and **spin down** (relative to a particular direction) and we call the physical quantity that is measured, in this case the component of the spin vector \vec{S} in a particular direction, an **observable**. We shall refer to

the Stern-Gerlach device as an **analyser** since it tells us whether the particle is spin up or spin down.

3. Two counters:



As the name suggests, the counters count how many particles are spin up and spin down. In Figure 2.1, 50 atoms have been detected coming out of the upper (\uparrow) port and 50 atoms have been detected coming out of the lower (\downarrow) port².

In Figure 2.1, the states of atoms are represented by new objects called **kets**. The $|+\rangle$ ket represents the quantum states of atoms with $S_z = \hbar/2$ while the $|-\rangle$ ket represents the quantum states of atoms with $S_z = -\hbar/2$. Since we don't have any information about it, the input beam is represented with the generic ket $|\psi\rangle$.

Ket notation was invented by Paul Dirac and is consequently called **Dirac notation**. Kets are central to the notation used in this course. The symbol(s) inside the ket label different states and there are many different ways of writing the same ket. For example, $|+\rangle$, $|+\hbar/2\rangle$, $|S_z = +\hbar/2\rangle$, $|+\hat{z}\rangle$ and $|\uparrow\rangle$ are all different ways of writing the same state while $|\pm\rangle$ refers to both the $|+\rangle$ and $|-\rangle$ kets.

Our first postulate about quantum mechanics relates to kets. Essentially it is the assumption that all the information we have about the state of some quantum system is represented by a ket.

² This is meant to schematically indicate the probabilities. In a real experiment, we would expect statistical fluctuations as you can see in the [animated simulation here](#). The animation was made using the SPINS program .

Dirac Notation

Quantum Mechanics Postulate 1

Quantum Mechanics Postulate 1

The state of a quantum mechanical system, including all the information you can know about it, is represented mathematically by a normalised ket $|\psi\rangle$.

A “normalised” ket has a “magnitude” of 1. We'll see in next chapter how to calculate the magnitude of a ket. While initially this postulate may seem vacuous and tautological, you will hopefully see its relevance later.

Now we will look at various simple Stern-Gerlach experiments which will teach us some features of Quantum Mechanics. We'll then see how the mathematics of kets can be used to predict the results of each experiment.

Experiment 1

Experiment 1, shown in Figure 2.2, has an additional Stern-Gerlach analyser. Both analysers in this experiment are aligned along the z -axis. Rather than sending the output of the 1st analyser's upper port to a counter, it is sent to a second analyser. We use counters to see how many atoms leave each of the remaining ports.

The most significant thing to notice from Figure 2.2 is that no atoms are detected in the lower (\downarrow) output of the 2nd analyser. We

Experimental set-up

Experimental result

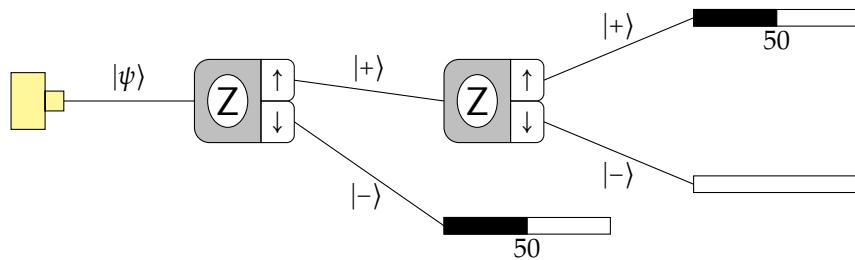


Figure 2.2: Experiment 1 measures the spin component along the z -axis twice in succession.

see that all atoms from the upper (\uparrow) port of the 1st analyser pass to upper (\uparrow) output port of the 2nd analyser.

This experiment has a rather simple interpretation which should not be that surprising. When the 1st analyser measures an atom to have $S_z = +\hbar/2$, the 2nd analyser also measures $S_z = +\hbar/2$ for that atom. In other words, having measured $S_z = +\hbar/2$, we find the same thing when we measure S_z again.

Interpretation of the experiment

Discussion 2.1

Starting with the set-up in fig. 2.2, what do you expect to find if we remove the upper output of the first analyser and feed its lower output into a second analyser aligned in the z -direction? What do you expect to happen if we measure any component of \vec{S} twice in a row?

The basic set up for this experiment will be similar to many of the experiments that follow. The analysers are positioned to play different roles in the experiment. The 1st analyser *prepares* the beam in a particular state – in this case the $|+\rangle$ state. We call the 1st analyser a **state preparation device**. The 2nd analyser then *measures* the resultant beam we have prepared.

State preparation

Comparing Experiment 1 with the later experiments will help you start to understand some of the surprising features of quantum mechanics.

Experiment 2

The set-up for this experiment, as shown in Figure 2.3, is similar to experiment 1 **except** that the 2nd analyser has been aligned along the x -axis. Since the first analyser is unchanged, atoms entering the 2nd analyser are still represented by the $|+\rangle$ ket, or in other words they have $S_z = +\hbar/2$. The first analyser prepares the beam to be in the state $|+\rangle$ and then the second analyser measures the x -component of the spin.

Experiment 2 results

We see that in contrast to the Experiment 1, atoms appear at both possible outputs of 2nd analyser. Atoms leaving upper port of 2nd analyser have been measured to have $S_x = \hbar/2$. We will label this state with the ket $|+\rangle_x$. On the other hand, atoms leaving lower port of 2nd analyser have been measured to have $S_x = -\hbar/2$ and we

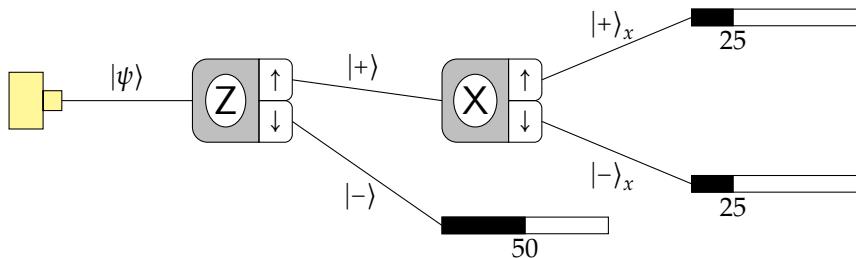


Figure 2.3: Experiment 2 measures the spin component along the z -axis and then along the x -axis.

indicate this state with the ket $|-\rangle_x$. From Figure 1.4 we conclude that atoms are evenly split, 50/50, between two states. Note that, in this course, we will follow the convention that, if there is no subscript, the spin component is along the z -axis or in other words, $|\pm\rangle = |\pm\rangle_z$.

The results would have been the same if we had used the lower output of the 1st analyser. In this case the first device, prepares atoms in the state $|-\rangle$. Then atoms entering the 2nd analyser will be measured to be split 50/50 between the $|\pm\rangle_x$ states.

We can not predict which of the 2nd analyser ports a particular atom will come out – it appears to be completely random. This can be demonstrated in actual experiments by recording the counts of each port of the second analyser one by one. One finds that the sequence of arrivals at each port is random *but* on average there is a 50% probability of a particular atom exiting either port. Looking back there is also randomness at the first analyser. We can not predict whether a particular atom coming out of the hot oven is spin up or spin down in the z -direction. The randomness at the first analyser is due to our ignorance. We don't know the history of a particular atom inside the hot oven so we have no information with which we can make a definitive prediction about its spin. The randomness at the second analyser is on the other hand a fundamental feature of quantum mechanics. Unlike atoms entering the first analyser, we are not ignorant about the atoms entering the second analyser – we know it is in the $|+\rangle$ state. Despite knowing that the particle is in the $|+\rangle$ state, we can only say with certainty that there is a 50/50 chance that the particle is spin up or spin down in the x -direction.

There is still a loophole in the claim that the randomness we observe at the second analyser is not just due to our ignorance. You could say that, just because we have measured the z -component of the spin why should that affect the x -component? The two directions are after all orthogonal to each other. Just like atoms came out of the oven with S_z randomised due to our ignorance of what happened in the oven, perhaps the S_x components remain randomised, after passing through the first analyser for the same reason. We will see why this argument doesn't work when we look at Experiment 3.

One might think that, perhaps we are just ignorant about some

Comments on Experiment 2

Randomness is a fundamental feature of QM

other aspects of the system and that there are *other* variables that we do not know about which would allow us to completely predict the properties of a particular atom. This idea goes by the name **local hidden variable theory**. John Bell proved that such theories are not compatible with the experimental results of quantum mechanics.

Quantum mechanics is a theory which makes **predictions about probabilities**. For example given that we know our atom is in the $|+\rangle$ state, we can make a prediction about the probability that it will be measured to be in the $|+\rangle_x$ or $|-\rangle_x$ states. Under some circumstances we *can* make predictions with 100% probability or in other words with certainty. For example, in **Experiment 1** (see Figure 2.2), we can be certain that the atoms coming out of the 2nd analyser are in the $|+\rangle$ state.

Experiment 3

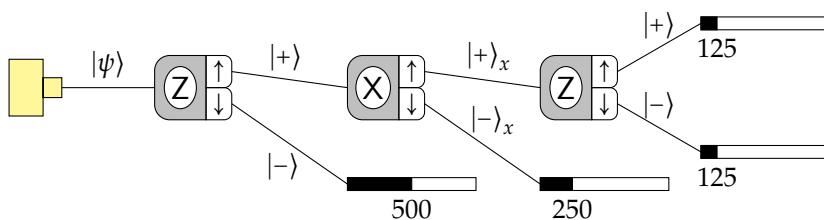


Figure 2.4: Experiment 3 measures the spin component three times in succession.

In this experiment we extend **Experiment 2** by adding a 3rd Stern-Gerlach analyser aligned along the z-axis. It shows that, surprisingly, we can not have information about different spin components at the same time. We say that they are **incompatible observables**.

Atoms entering the 3rd analyser were measured to be have spin up along the z-axis ($|+\rangle$) at the first analyser, then the second analyser measures them to have spin up along the x-axis ($|+\rangle_x$) and finally the 3rd analyser measures whether these atoms have spin up or spin down along the z-axis

The final analyser measures 50% of the atoms to be spin down and 50 % of the atoms to be spin up. This is *certainly* not what one would expect classically. We expect the final measurement to spin-up since the first two measurements tell us that the atoms have $S_z = \hbar/2$ and $S_x = \hbar/2$. This means that the third measurement should just confirm that $S_z = \hbar/2$. Comparing Figure 2.3 and Figure 2.4 you will notice that the last two analyses behave just like the two analysers in **Experiment 2**. The fact that the initial measurement in **Experiment 3** was spin up seems to have been “forgotten”.

An important feature of quantum mechanics is demonstrated by this experiment, namely that, under certain circumstances, the act of measurement itself *necessarily* disturbs the system. We see that the 2nd analyser disturbed the atoms so that S_z no longer has a

Spin components are incompatible observables

Experiment 3 set-up

Experiment 3 results

unique value even though it was just measured by the first analyser.

💡 You might wonder whether there is another way to design the experiment so that we do not disturb the atoms – the short answer is **no!** 🤔 There is a **fundamental problem** with trying to measure the spin components of atoms along different directions. We say that S_x and S_z are **incompatible observables** which means we can not have complete information about both of them at the same time. The system can be represented by the ket $|+\rangle = |S_z = +\hbar/2\rangle$ or $|+\rangle_x = |S_x = \hbar/2\rangle$ **but** not by a ket like $|S_z = +\hbar/2, S_y = +\hbar/2\rangle$ which specifies both components. It should be noted that not all pairs of quantum mechanics observables are **incompatible**. It is possible to make some measurements of a system without disturbing certain other features of the system. Classically, we would expect to be able to independently determine all three components of a spin vector which would tell us the direction the spin is pointing. In quantum mechanics, the incompatibility of spin components means we can not know in which direction spin is pointing.

Experiment 4

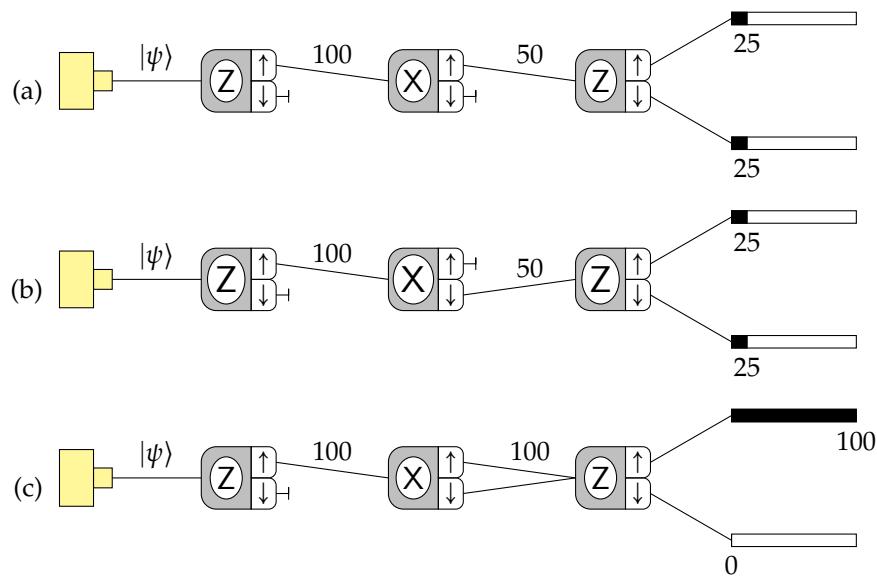


Figure 2.5: Experiment 4 measures the spin component three times in succession and uses (a and b) one or (c) two beams from the second analyser.

Experiment 4 involves slight variations on **Experiment 3** which demonstrate *another* surprising feature of quantum mechanics namely that quantum mechanics involves *interference effects*.

There are some changes in notation in Figure 2.5. Firstly, the diagrams only show counters on last analyser. We just block off the beams we are not counting as shown. Secondly, we just show average number of atoms, out of 100, going from one analyser to the next.

In **Experiment 4c** there is a new feature. The two output beams from the 2nd analyser are recombined – this must be done without “disturbing” the beams. This is analogous to the way light is recombined after passing through diffraction gratings.

Experiment 4 setup

Experiment 4 results

The results of **Experiment 4a** and **Experiment 4b** are essentially the same. From **4a** we would conclude that there is a 25% chance for an atom leaving upper port of the 1st analyser to take the upper path through the 2nd analyser and then leave through the upper port of the 3rd analyser. Similarly from **4b** we see that there is a 25% chance for an atom leaving upper port of the 1st analyser to take the lower path through the 2nd analyser and then leave through the upper port of the 3rd analyser. Now comes the surprising part, combining the results from **4a** and **4b** one might conclude that there is a $25\% + 25\% = 50\%$ possibility for an atom to exit from the upper port of the 3rd analyser when both paths are available. Similarly, we expect there to be a 50% possibility for an atom to exit from the lower port of the 3rd analyser when both paths are available. **Contrary** to what one might expect from **4a** and **4b**, all the atoms in **4c** exit from the upper port and atoms now seem to "remember" that they left the upper port of the 1st analyser. By combining the beams, we seem to have gotten rid of the quantum disturbance evident in Experiments **3**, **4a** and **4b** so that the result is the same as **Experiment 1** as if the 2nd analyser. isn't there

Things get more curious when we look closely at the various paths atoms may be following in **Experiment 4**. In experiments **4a** & **4b** 50% of the atoms are blocked after the 2nd counter and 25% exit the lower port of the 3rd analyser. On the other hand in **4c**, no atoms are blocked at the 2nd counter and yet no atoms exit the lower port of the 3rd analyser. So, in experiment **4c**, by allowing the atoms more ways to reach the lower counter of the 3rd analyser we get fewer counts. Its as if you opened a second window and the room went dark!

There is a way to explain this kind of effect that you may already know. Can you think of a example where combining two effects leads to cancellation rather than enhancement? One such example is the **wave interference** in a double slit experiment (see Figure 2.6).

In the double slit experiment, light passes through two narrow slits which leads to an interference pattern. If light passes through just one slit, we do not get interference³, but two slits combined lead to alternating bright and dark bands. This effect is explained by the electromagnetic waves from each slit cancelling in certain places, leading to dark bands, and adding in other places, leading to bright bands. One can determine the pattern by adding together the electric field vectors of the light & squaring the resulting vector to find the **total intensity**. There will be a similar recipe in quantum mechanics. Will will have to add together amplitudes & take the square to find the probability for some possible outcomes of our experiment. The fact that we will add amplitudes means there can be **interference** as is demonstrated by experiment **4c**. Part of what you will be learning in this course is the "mechanics" of doing these calculations.

³ Although there is diffraction through a single slit.

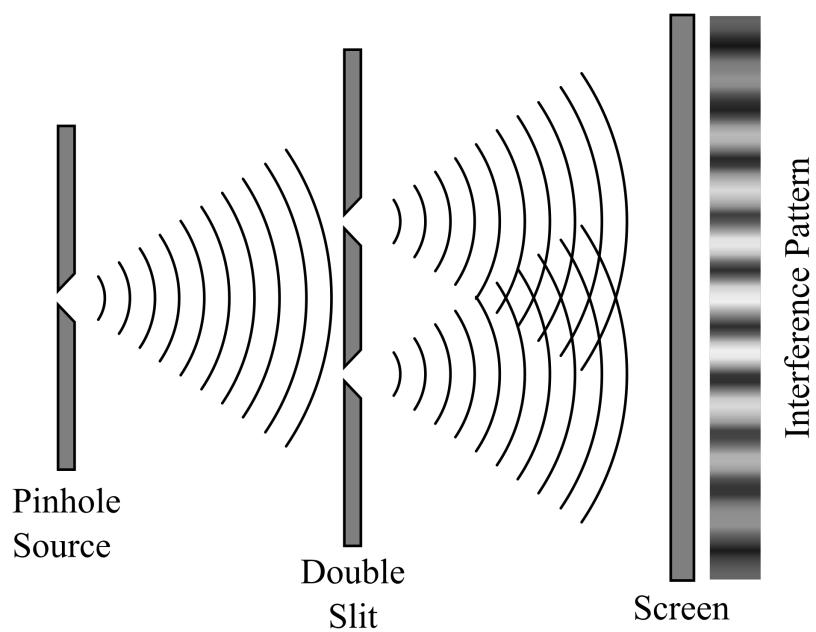


Figure 2.6: Young's double-slit interference experiment.

Modification of this image

Attribution: Lacatosias, Stannered, CC BY-SA 3.0 <http://creativecommons.org/licenses/by-sa/3.0/>, via Wikimedia Commons

留言板 Discussion 2.2

Write a summary of the important points in this chapter

3

Quantum State Vectors

Review

This reading is based on [McIntyre et al. 2012, Section 1.2]

QM Postulate 1

The state of a quantum mechanical system, including all the information you can know about it, is represented mathematically by a normalised ket, $|\psi\rangle$.

$|Kets\rangle$

Kets give us a mathematical description of quantum mechanical systems. They obey some of the same algebraic rules as ordinary vectors¹. Since they do not describe physical vectors, but rather the state of our quantum mechanical system, we call them *abstract* vectors. Kets live in a vector space that we call **Hilbert space**. The number of dimensions of our Hilbert space depends on the system we are looking at. For example, in the Stern-Gerlach experiment with silver atoms, there are two possibilities for a measurement of the spin component – spin up or spin down. This means that our Hilbert space is two dimensional.

In this section we will discuss some of the mathematical properties of kets and how we can use them to make predictions about physical quantum systems. Just like \hat{i} , \hat{j} and \hat{k} constitute a basis for 3d vectors, $|+\rangle$ and $|-\rangle$ are a basis for a spin $\frac{1}{2}$ system. The $|\pm\rangle$ basis has properties analogous to some of the familiar properties of the $\{\hat{i}, \hat{j}, \hat{k}\}$ basis as summarised in Table² 3.1.

Name	Expression
Normalisation	$\hat{i} \cdot \hat{i} = \hat{j} \cdot \hat{j} = \hat{k} \cdot \hat{k} = 1$
Orthogonality	$\hat{i} \cdot \hat{j} = \hat{j} \cdot \hat{k} = \hat{k} \cdot \hat{i} = 0$
Completeness	$\forall \vec{A}, \vec{A} = A_x \hat{i} + A_y \hat{j} + A_z \hat{k}$, where, for eg, $A_x = \vec{A} \cdot \hat{i}$

Completeness, which may be less familiar than the other two, is the property that any general vector, \vec{A} , can be written as a linear combination of the basis vectors. Notice that all of these properties involve the **dot product**.

¹ Mathematicians would say that kets are elements of a *vector space*. Vector spaces are mathematical constructions whose elements share many of the properties you are familiar with from ordinary spacial vectors.

² The symbol \forall is short hand for “for all”.

Table 3.1: Properties of a Cartesian orthonormal basis

As *abstract* vectors, a basis for kets should also have analogous properties. Starting with completeness, note that when we measure S_z there are only two possible results which are represented by the states $|+\rangle$ and $|-\rangle$. Since these are the only two possibilities, they must constitute a complete basis that we call the S_z basis, and in this case, **completeness** means that a general state can be written:

$$|\psi\rangle = a|+\rangle + b|-\rangle , \quad (3.1)$$

where a and b are *complex* numbers multiplying each ket. When writing a ket, the number can appear before or after the ket or in other words,

$$a|+\rangle = |+\rangle a . \quad (3.2)$$

Whereas the components of ordinary spacial vectors are real, it turns out that we need to use *complex numbers* in quantum mechanics. You may have seen the *optional* use of complex numbers in other areas of physics, for example, sometimes it is easier to describe oscillations using complex exponentials like $e^{i\omega t}$ rather than say $\cos(\omega t)$. In quantum mechanics, complex numbers are *essential*.

Now, before we can talk about orthogonality and normalisation for kets, we need to define a dot product. To understand what features this dot product should have, let's review some of the properties of the dot product shown in Table 3.2

Property	Formula
Distributive over vector addition	$\vec{a} \cdot (\vec{b} + \vec{c}) = \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c}$
Positive definite	$\vec{a} \cdot \vec{a} > 0$ if $\vec{a} \neq 0$
Symmetric	$\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}$

To ensure that it is positive definite, the dot product of vectors with complex components have to be slightly generalised. In particular, we need to take the complex conjugate of one of the vectors. One way to define the dot product between two vectors, \vec{a} and \vec{b} , with complex components is, $\vec{a}^* \cdot \vec{b}$, where “ \cdot ” is the usual dot product for 3d vectors. For example, say we have the $\vec{a} = (1)\hat{i} + (i)\hat{j}$ and $\vec{b} = (2)\hat{i} + (3)\hat{j}$ then the dot product of \vec{a} and \vec{b} is

$$\vec{a}^* \cdot \vec{b} = ((1)\hat{i} + (-i)\hat{j}) \cdot ((2)\hat{i} + (3)\hat{j}) = 2 - 3i$$

As mentioned, the reason we take the complex conjugate of one of the vectors is so that the dot product of a complex vector with itself is always positive. This can clearly be seen by looking at the dot product of a 2d vector, \vec{a} , with itself:

$$(\vec{a})^* \cdot \vec{a} = ((a_x)^*\hat{i} + (a_y)^*\hat{j}) \cdot ((a_x)\hat{i} + (a_y)\hat{j}) = |a_x|^2 + |a_y|^2 \geq 0 .$$

Discussion 3.1

If $\vec{A} = (1+i)\hat{i} + 3\hat{j} + 2i\hat{k}$ and $\vec{B} = 3\hat{i} + 7\hat{k}$, calculate the dot product $(\vec{A})^* \cdot \vec{B}$. How does this differ from $(\vec{B})^* \cdot \vec{A}$?

Table 3.2: Some properties of the dot product for (real) vectors

Discussion 3.2

Show that for a general vector, $\vec{C} = C_x \hat{i} + C_y \hat{j} + C_z \hat{k}$, with complex components, the dot product with itself is a greater than or equal to zero, ie. $(\vec{C})^* \cdot \vec{C} \geq 0$.

$\langle \text{Bra} |$

In quantum mechanics, with abstract vectors, the analogue of the complex conjugate vector is called a **bra**³. So, corresponding to a general ket $|\psi\rangle$, there is a bra which we write as $\langle\psi|$. If we have a general spin- $\frac{1}{2}$ ket, $|\psi\rangle = a|+\rangle + b|-\rangle$, then the corresponding **bra**, $\langle\psi|$, is defined to be

$$\langle\psi| = a^* \langle+| + b^* \langle-|$$

You can see that, in the expression above, the basis bras, $\langle+|$ and $\langle-|$, correspond to the basis kets, $|+\rangle$ and $|-\rangle$. You should also note that the complex coefficients a and b have been complex conjugated.

Inner Product

In quantum mechanics, we define the scalar product as the product of a bra and a ket :

$$(\langle \text{bra} |) (\text{ket} \rangle) .$$

when the bra and ket are combined like this we get a *braket* which is some complex number⁴. The short hand for the *braket* is

$$\langle \text{bra} | \text{ket} \rangle .$$

For example, the inner product of $|+\rangle$ and $|-\rangle$ is written

$$(\langle+|) (\text{ket} \rangle) = \langle+|-\rangle .$$

To make things look a little neater, we leave out an extra “|” in the middle (in other words we don’t write $\langle+||-\rangle$). The scalar product in quantum mechanics is often called the **inner product** or a **projection**.

You might now wonder how to actually calculate scalar products like $\langle+|+\rangle$. Well by analogy with, for example, $\hat{i} \cdot \hat{i} = 1$, we define $\langle+|+\rangle = 1$. Using unit basis kets make working with them much easier⁵. In summary, in our spin- $\frac{1}{2}$ basis we take

$$\langle+|+\rangle = \langle-|-\rangle = 1$$

Finally, the remaining property is **orthogonality**. The 3d unit vectors \hat{i} , \hat{j} and \hat{k} are orthogonal because they are at 90° to each other. Mathematically we write $\hat{i} \cdot \hat{j} = \hat{j} \cdot \hat{k} = \hat{k} \cdot \hat{i} = 0$. For the basis kets $|+\rangle$ and $|-\rangle$ there is no spacial geometry so we can not visualise them as being at 90° to each other. We will **assume** that $|+\rangle$ and $|-\rangle$ are mathematically orthogonal. This means we will assume that they satisfy the equation $\langle+|-\rangle = \langle-|+\rangle = 0$.

³ Mathematicians refer to “bras” as co-vectors or dual vectors. One way of thinking of co-vectors is as the set of linear maps which map vectors to real numbers. In special relativity, we indicate vectors with an upper index, as in v^α , and co-vectors with a lower index, as in v_α .

⁴ this was Dirac’s little joke....

⁵ It turns out that using **normalised** basis vectors in quantum mechanics has a physical meaning that we will look at soon.

Summary	Formula
Normalisation	$\langle+ +\rangle = 1$ $\langle- -\rangle = 1$
Orthogonality	$\langle+ -\rangle = 0$ $\langle- +\rangle = 0$
Completeness	$ \psi\rangle = a +\rangle + b -\rangle$

Table 3.3: Properties of the $|\pm\rangle$ basis

Three properties of the $|\pm\rangle$ basis, analogous to those in Table 3.1, are summarised in Table 3.3. Using the ket notation, we can write some of the properties of the inner product analogous to the ones we write down for the 3d dot product in Table 3.2.

Property	Formula
Distributive over vector addition	$\langle \alpha (\beta\rangle + \gamma\rangle) = \langle \alpha \beta\rangle + \langle \alpha \gamma\rangle$
Positive definite	$\langle \alpha \alpha \rangle > 0$ if $ \alpha\rangle \neq 0$

Notice that something analogous to the symmetry of the dot product is missing from Table 3.4 – we will look at the analogous symmetry of the inner product presently.

You may be familiar with the way one can extract the component of a 3d vector using the dot product. For example, we can find the x -component of a vector using, $A_x = \hat{i} \cdot \vec{A}$. We also call, $\hat{i} \cdot \vec{A}$, the **projection** of \vec{A} along the x -axis. More generally, we can write $\vec{A} = (\hat{i} \cdot \vec{A})\hat{i} + (\hat{j} \cdot \vec{A})\hat{j} + (\hat{k} \cdot \vec{A})\hat{k}$.

Doing the analogous thing for kets, taking the inner product of the bra, $\langle +|$, with a general state, $|\psi\rangle = a|+ \rangle + b|-\rangle$ we get

$$\begin{aligned}
 \langle + | \psi \rangle &= \langle + | (a|+ \rangle + b|-\rangle) \\
 &= \langle + | a |+ \rangle + \langle + | b |-\rangle \quad (\text{distributive property}) \\
 &= a \langle + |+ \rangle + b \langle + |-\rangle \quad (\text{moving } a \text{ and } b \text{ through kets}) \\
 &= a \quad (\text{normalisation \& orthogonality})
 \end{aligned} \tag{3.3}$$

Discussion 3.3

Show that $\langle - | \psi \rangle = b$

We see that the coefficients in our expression for a general state, $|\psi\rangle = a|+ \rangle + b|-\rangle$, are just the inner products or **projections** of a general state $|\psi\rangle$ along each basis ket⁶. In summary, given $|\psi\rangle = a|+ \rangle + b|-\rangle$, we can extract a and b using $\langle + | \psi \rangle = a$ and $\langle - | \psi \rangle = b$. This means we can write the general state as

$$\begin{aligned}
 |\psi\rangle &= a|+ \rangle + b|-\rangle \\
 &= \{\langle + | \psi \rangle\}|+ \rangle + \{\langle - | \psi \rangle\}|-\rangle .
 \end{aligned} \tag{3.4}$$

Can you see how the result above is analogous to the expression $\vec{A} = (\hat{i} \cdot \vec{A})\hat{i} + (\hat{j} \cdot \vec{A})\hat{j} + (\hat{k} \cdot \vec{A})\hat{k}$?

Finally, we can get to the symmetry property of the inner product. For a general state, $|\psi\rangle = a|+ \rangle + b|-\rangle$, we have the corresponding bra, $\langle \psi| = a^* \langle +| + b^* \langle -|$, so

$$\begin{aligned}
 \langle \psi | + \rangle &= \langle + | a^* |+ \rangle + \langle - | b^* |+ \rangle \\
 &= a^* \langle + |+ \rangle + b^* \langle - |+ \rangle \\
 &= a^*
 \end{aligned} \tag{3.5}$$

Table 3.4: Some properties of the inner product

⁶ In the language of 3d vectors, a is the “+”-component and b is the “-”-component.

We see that reversing the order of states in the inner product, leads to complex conjugation:

$$\langle + | \psi \rangle = \langle \psi | + \rangle^*$$

This property holds for any inner product, so that, for two general states, $|\psi\rangle$ and $|\varphi\rangle$,

Conjugation of the inner product

$$\langle \varphi | \psi \rangle = \langle \psi | \varphi \rangle^* \quad (3.6)$$

Discussion 3.4

Show that for the general states, $|\psi\rangle = a|+\rangle + b|-\rangle$, and, $|\varphi\rangle = c|+\rangle + d|-\rangle$, that $\langle \varphi | \psi \rangle = \langle \psi | \varphi \rangle^*$

The symmetry property of our quantum mechanical inner product, $\langle \varphi | \psi \rangle = \langle \psi | \varphi \rangle^*$, is analogous to the symmetry of the 3d vector dot product, $\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}$.

So far we have seen similarities between quantum vectors, or in other words kets, and the spacial vectors we are used to from classical mechanics but there is one very important difference. The rules of quantum mechanics (postulate 1 and postulate 4 which we will come to shortly) **require** that **all** state vectors, not just the basis kets, describing a quantum system are normalised. This is obviously **very** different from ordinary spacial vectors which can be any length (except for unit vectors).

Requiring that a general ket, $|\psi\rangle$, is normalised leads to

$$\begin{aligned} 1 &= \langle \psi | \psi \rangle = \{a^* \langle + | + b^* \langle - | \} \{a | + \rangle + b | - \rangle\} \\ &\implies 1 = a^* a \langle + | + b^* b \langle - | + b^* a \langle - | + \rangle + b^* b \langle - | - \rangle \\ &\implies 1 = a^* a + b^* b \\ &\implies 1 = |a|^2 + |b|^2. \end{aligned} \quad (3.7)$$

Discussion 3.5

Show how we go from the 2nd to the 3rd line of (3.7).

Example 3.1: Normalisation

Problem Normalise the vector $|\psi\rangle = C(1|+\rangle + 2i|-\rangle)$. We call the constant C , the **normalisation constant**.

Solution To normalise $|\psi\rangle$, we set the inner product of the state with itself to be 1 and solve for C :

$$\begin{aligned}
 1 &= \langle\psi|\psi\rangle \\
 &= C^* \{1(+| + (-2i)(-|)\} C \{1|+ + 2i|-)\} \\
 &= CC^* \{1(+|+) + 2i(+|-) - 2i(-|+) + 4(-|-)\} \\
 &= 5|C|^2 \\
 \implies |C| &= \frac{1}{\sqrt{5}} \\
 \implies C &= \frac{1}{\sqrt{5}}e^{i\theta} \tag{3.8}
 \end{aligned}$$

for some phase θ .

Discussion 3.6

Explain how we go from the 2nd last to the last line of (3.8).

You will show in a problem set that the overall phase of a normalisation constant does not affect the physics and, following convention, we take C to be real and positive leading to $C = 1/\sqrt{5}$. This means the normalised quantum state vector is $|\psi\rangle = \frac{1}{\sqrt{5}}(1|+ + 2i|-) = \frac{1}{\sqrt{5}}|+ + \frac{2i}{\sqrt{5}}|-$.

QM postulate 4

We can also rewrite (3.7) as,

$$|\langle +|\psi\rangle|^2 + |\langle -|\psi\rangle|^2 = 1, \tag{3.9}$$

which leads us to a very important feature of quantum mechanics. We **postulate** that each term in the sum (3.9), is equal to the probability that the quantum state described by the ket $|\psi\rangle$, is measured to be in the corresponding basis state. This means that when a measurement of S_z is made, $|\langle +|\psi\rangle|^2$ is the probability that state $|\psi\rangle$ will be found to be in the state $|+\rangle$, and $|\langle -|\psi\rangle|^2$ is the probability that state $|\psi\rangle$ will be found to be in the state $|-\rangle$:

$$\underbrace{|\langle +|\psi\rangle|^2}_{\text{Probability of measuring } S_z = +\hbar/2} + \underbrace{|\langle -|\psi\rangle|^2}_{\text{Probability of measuring } S_z = -\hbar/2} = 1. \tag{3.10}$$

Notation

We will use a curly \mathcal{P} as short hand for “probability” and $\mathcal{P}_{\text{Observation } X}$ is short hand for “probability that we measure Observation X” for a given state.

So for example⁷, $\mathcal{P}_{S_z=+\hbar/2} = |\langle +|\psi\rangle|^2$, is probability that we mea-

⁷ If you use L^AT_EX, you can use `\mathscr{P}` to produce “ \mathcal{P} ”.

sure $S_z = +\hbar/2$ for a given state, $|\psi\rangle$. In other words, $\mathcal{P}_{S_z=+\hbar/2}$, is the probability, our particle, which starts off in the state $|\psi\rangle$, is measured to be in the state $|+\rangle$. On the other hand, $\mathcal{P}_{S_z=-\hbar/2} = |\langle -|\psi\rangle|^2$, is probability that we measure $S_z = -\hbar/2$ for a given state, $|\psi\rangle$. In other words, $\mathcal{P}_{S_z=-\hbar/2}$, is the probability, our particle, which starts off in the state $|\psi\rangle$, is measured to be in the state $|-\rangle$.

Since we will often refer to spin measurements, we adopt the abbreviations in Table 3.5. Using this notation, the 4th postulate for the z-component of a spin-½ system is:

QM Postulate 4 (spin-½ system)

The probability of obtaining the value, $\pm\hbar/2$, in a measurement of the observable, S_z , on a system in the state $|\psi\rangle$ is, $\mathcal{P}_\pm = |\langle \pm|\psi\rangle|^2$, where, $|\pm\rangle$, are the basis kets of S_z corresponding to the results $\pm\hbar/2$.

Spin-up	Spin-down
$\mathcal{P}_{S_x=\hbar/2} = \mathcal{P}_{+x}$	$\mathcal{P}_{S_x=-\hbar/2} = \mathcal{P}_{-x}$
$\mathcal{P}_{S_y=\hbar/2} = \mathcal{P}_{+y}$	$\mathcal{P}_{S_y=-\hbar/2} = \mathcal{P}_{-y}$
$\mathcal{P}_{S_z=+\hbar/2} = \mathcal{P}_+$	$\mathcal{P}_{S_z=-\hbar/2} = \mathcal{P}_-$

Table 3.5: Probability notation

Discussion 3.7

How should we generalise QM Postulate 4 above for measurements of S_x and S_y ?

We'll come to postulates 2 & 3 as well as an even more general form of postulate 4 later.

Why all states are normalised

Looking at the formula

$$\langle \psi | \psi \rangle = |\langle +|\psi\rangle|^2 + |\langle -|\psi\rangle|^2 = \mathcal{P}_+ + \mathcal{P}_- = 1 ,$$

we are now in a position to understand why, according to our postulates, states must be normalised. In our spin-½ system, there are only 2 possible measurement results (for a spin-component). This implies that the two probabilities \mathcal{P}_+ and \mathcal{P}_- must sum to 1 since there is a 100% chance to getting some outcome. In other words, requiring that all the possible probabilities should add up to 1 leads to the requirement that all state vectors must be normalised.

Example 3.2:

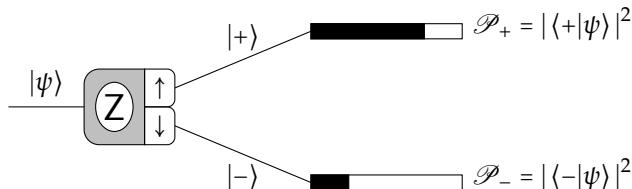
With our new notation, we'll look at a simple Stern-Gerlach experiment (as shown in Figure 3.1 (a)) with a given initial state:

$$|\psi\rangle = \textcolor{blue}{a}|+\rangle + \textcolor{green}{b}|-\rangle = \{\langle +|\psi\rangle\}|+\rangle + \{\langle -|\psi\rangle\}|-\rangle .$$

As shown in the figure, the probability of measuring $S_z = +\hbar/2$ is $\mathcal{P}_+ = |\langle +|\psi\rangle|^2$ and the probability of measuring $S_z = -\hbar/2$ is $\mathcal{P}_- = |\langle -|\psi\rangle|^2$. A histogram summarising the results, is shown in Figure 3.1 (b). We will often use such

histograms to summarise results so you should be able to interpret it.

(a)



(b)

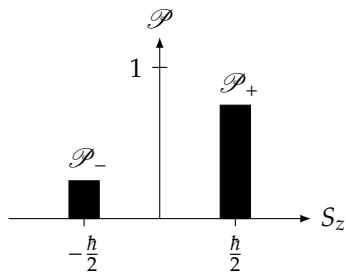
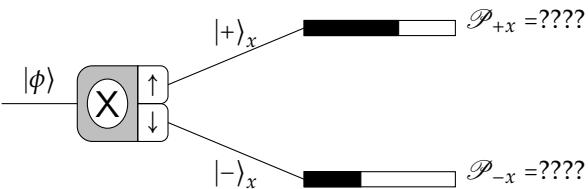


Figure 3.1: (a) Spin component measurement for a general input state and (b) histogram of measurement results.

Discussion 3.8



Consider the set up above. Using the result of Discussion 3.7, what are P_{+x} and P_{-x} in terms of $|\phi\rangle$, $|+\rangle_x \langle +|$ and $|-\rangle_x \langle -|$?

Amplitudes

We can find the intensity of a classical wave by squaring the wave amplitude. In quantum mechanics, one finds the probability of a measurement by squaring an inner product. Because of this similarity we call an inner product like $\langle + |\psi \rangle$ a **probability amplitude** or just an **amplitude**. Later we will see that there is a quantum wave function whose intensity is related to probability.

We will follow a particular convention when writing amplitudes. Suppose we have some initial state $|in\rangle$ and we are interested in the amplitude associated with some final state $|out\rangle$. We will put the input or initial state on the right and the output or final state on the left. In other words we write the amplitude as $\langle out|in\rangle$. Since probability involves the complex square of the amplitude and due to the symmetry of the inner product, $\langle out|in\rangle = \langle in|out\rangle^*$, this convention does not affect probability calculations. This is because, $P_{out} = |\langle out|in\rangle|^2 = |\langle in|out\rangle^*|^2 = |\langle in|out\rangle|^2$.

Amplitude convention

 Discussion 3.9

Write a summary of the important points and formulae in this chapter.

4

Stern-Gerlach re-examined

We will now revisit the Stern-Gerlach experiments we looked at before. This time we will be using the language of kets and the 4th postulate of Quantum mechanics.

This reading is based on [McIntyre et al. 2012, Section 1.2]

QM Postulate 4 (spin- $\frac{1}{2}$ system)

The probability of obtaining the value, $\pm\hbar/2$, in a measurement of the observable, S_z , on a system in the state $|\psi\rangle$ is,
 $\mathcal{P}_{\pm} = |\langle \pm | \psi \rangle|^2$, where, $|\pm\rangle$, are the basis kets of S_z corresponding to the results $\pm\hbar/2$.

SPINS Stern-Gerlach simulator

You can simulate the experiments discussed in this chapter using SPINS, a Stern-Gerlach simulator [here](#). 

Challenge: see if you can reproduce the experiments mentioned in the chapter. You can download a java version of the program to run on your own computer [here](#). 

Experiment 1

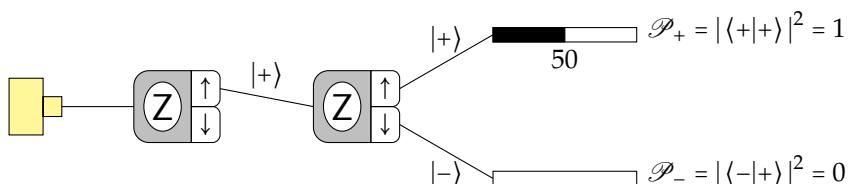


Figure 4.1: Probabilities of spin component measurements in Experiment 1.

As shown in Figure 4.1, the 1st analyser prepared atoms in the state $|+\rangle$, and then the 2nd analyser finds all those atoms to be in the state $|+\rangle$. To use the 4th postulate of QM to analyse this experiment, we start off with the fact that atoms enter the 2nd analyser in the state, $|\psi_{in}\rangle = |+\rangle$. Looking at the probability to measure spin up

in the z -direction, \mathcal{P}_+ , means that our “out” state, $|\text{out}\rangle$, is $|+\rangle$. So, since $|+\rangle$ is normalised we get:

$$\mathcal{P}_+ = |\langle \text{out} | \text{in} \rangle|^2 = |(+|\psi_{\text{in}}\rangle)|^2 = |(+|+)\rangle|^2 = 1$$

If we want the chance of measuring spin down, \mathcal{P}_- , our “out” state is, $|-\rangle$. Then using the fact that $|-\rangle$ and $|+\rangle$ are orthogonal leads to:

$$\mathcal{P}_- = |\langle \text{out} | \text{in} \rangle|^2 = |(-|\psi_{\text{in}}\rangle)|^2 = |(-|+)\rangle|^2 = 0$$

Since, $\mathcal{P}_+ = 1$ and $\mathcal{P}_- = 0$, there is a 100% chance that the 2nd analyser measures $S_z = \frac{1}{2}\hbar$. The results of the experiment are summarised in Figure 4.2. We see that if the atoms are prepared in the $|+\rangle$ state they remain in the $|+\rangle$ state.

Experiment 2

Experiment 2 is similar to the previous one. As before the 1st analyser prepares the system in the state $|+\rangle$ but then 2nd analyser measures the spin along the x -axis. One finds that there is 50/50 chance for atoms to be found in the $|+\rangle_x$ or $|-\rangle_x$ states. An alternative version of Experiment two is shown in Figure 4.4. In this version we look at the atoms coming out of the lower port of the first analyser, which is aligned along the z -axis. These atoms must have $S_z = -\hbar/2$ or in other words, they are in the $|-\rangle$ state. In both cases there is a 50/50 split between $S_x = +\hbar/2$ and $S_x = -\hbar/2$ at the 2nd analyser. This tells us that after we measure S_z , we can only say that there is a 50/50 chance that the atom is spin up or spin down in the x -direction.

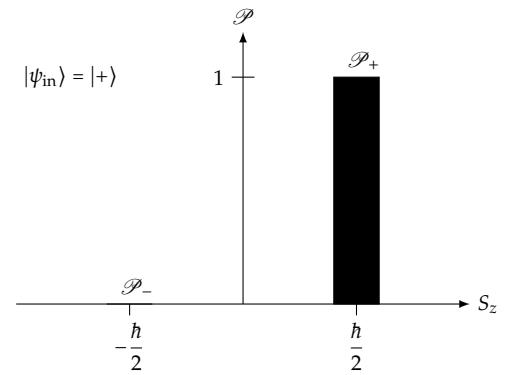


Figure 4.2: Histogram of S_z spin component measurements for Experiment 1 with $|\psi_{\text{in}}\rangle = |+\rangle$.

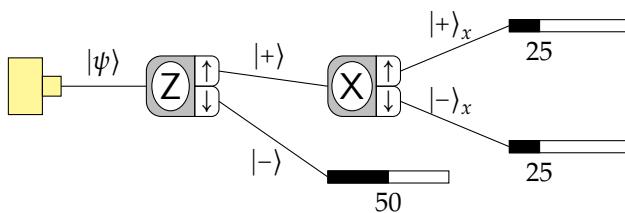


Figure 4.3: Original version of Experiment 2 which measures the spin component along the z -axis and then along the x -axis.

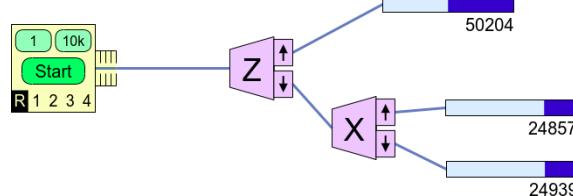


Figure 4.4: Alternative version of Experiment 2 in which the 1st analyser prepares system in state $|\psi_{\text{in}}\rangle = |-\rangle$

The results of both versions of Experiment 2 are summarised through the histograms in Figure 4.5.

We can not (yet) predict the results of **Experiment 2** just by looking at the set-up. This is because, we do not know enough about the states $|\pm\rangle_x$. The good news is that we can use Experiment 2 and the 4th postulate of quantum mechanics, to learn more about $|\pm\rangle_x$.

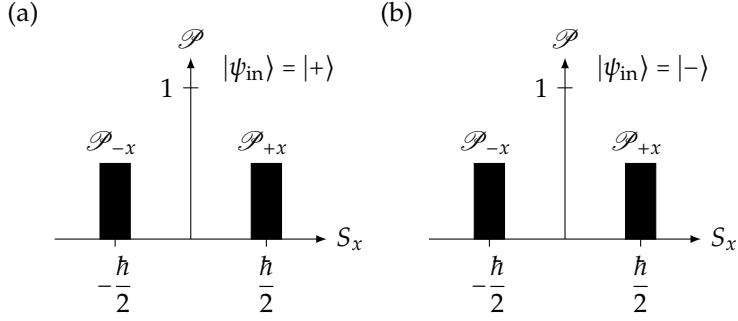


Figure 4.5: Histograms of S_x spin component measurements for Experiment 2 for different input states (a) $|\psi_{\text{in}}\rangle = |+\rangle$ and (b) $|\psi_{\text{in}}\rangle = |-\rangle$.

In the first version of Experiment 2, we look at atoms coming out of the upper port of the z -analyser so, the atoms are prepared in the $|\text{in}\rangle$ state: $|\psi_{\text{in}}\rangle = |+\rangle$. Since the second analyser is aligned along the x -direction, the two possible “out” states are $|+\rangle_x$ with $S_x = +\hbar/2$ and $|-\rangle_x$ with $S_x = -\hbar/2$. From the experiment we see that the probability, \mathcal{P}_{+x} , of measuring $S_x = +\hbar/2$ is $\frac{1}{2}$. This tells us something about the inner product between $|+\rangle$ and $|+\rangle_x$ since the 4th postulate of QM allows us to determine that in this case

$$\begin{aligned}\mathcal{P}_{+x} &= |\langle_x (+|\psi_{\text{in}}\rangle)|^2 \\ &= |\langle_x (+|+)\rangle|^2 \\ &= \frac{1}{2}\end{aligned}\quad (4.1)$$

In a similar fashion, using the observation that in this experiment that the probability, \mathcal{P}_{-x} , of measuring $S_x = -\hbar/2$ is $\frac{1}{2}$, we can conclude that

$$\begin{aligned}\mathcal{P}_{-x} &= |\langle_x (-|\psi_{\text{in}}\rangle)|^2 \\ &= |\langle_x (-|+)\rangle|^2 \\ &= \frac{1}{2}\end{aligned}\quad (4.2)$$

Experiment 2(b) is very similar to experiment 2(a) except that for the second version we have $|\psi_{\text{in}}\rangle = |-\rangle$. This means that, with a slight modification of the previous calculations one obtains

$$\begin{aligned}\mathcal{P}_{+x} &= |\langle_x (+|\psi_{\text{in}}\rangle)|^2 \\ &= |\langle_x (+|-)\rangle|^2 \\ &= \frac{1}{2}\end{aligned}\quad (4.3)$$

$$\begin{aligned}\mathcal{P}_{-x} &= |\langle_x (-|\psi_{\text{in}}\rangle)|^2 \\ &= |\langle_x (-|-)\rangle|^2 \\ &= \frac{1}{2}\end{aligned}\quad (4.4)$$

Now, armed with the inner products between $|\pm\rangle$ and $|\pm\rangle_x$, we can find expressions for $|\pm\rangle_x$ in the basis $\{|+\rangle, |-\rangle\}$. Starting from the assumption that $|+\rangle$ and $|-\rangle$ form a complete basis implies that $|\pm\rangle_x$ can be written in terms of them as follows:

$$\begin{aligned}|+\rangle_x &= a|+\rangle + b|-\rangle \\ |-\rangle_x &= c|+\rangle + d|-\rangle,\end{aligned}\quad (4.5)$$

where a, b, c and d are unknown constants. From (4.5) we get

$$\begin{aligned} {}_x\langle +| &= a^* \langle +| + b^* \langle -| \\ {}_x\langle -| &= c^* \langle +| + d^* \langle -| , \end{aligned} \quad (4.6)$$

so using our postulate, and the **mathematics** of kets

$$\begin{aligned} \mathcal{P}_{+x} &= |{}_x\langle +| + \rangle|^2 \\ &= |(a^* \langle +| + b^* \langle -|) \langle +| |^2 \quad (\text{Discussion 4.1}) \\ &= |a^*|^2 = |a|^2 \end{aligned} \quad (4.7)$$

Discussion 4.1

Show how we go from the 2nd to the 3rd line of the calculations above.

Finally, comparing the mathematical result, $\mathcal{P}_{+x} = |{}_x\langle +| + \rangle|^2 = |a|^2$, with the observation from **Experiment 2(a)** that $\mathcal{P}_{+x} = \frac{1}{2}$, we conclude $|a|^2 = \frac{1}{2}$.

Discussion 4.2

We just found $|a|^2 = \frac{1}{2}$ where

$$\begin{aligned} |+\rangle_x &= a|+\rangle + b|-\rangle , \\ |-\rangle_x &= c|+\rangle + d|-\rangle . \end{aligned}$$

Show in a similar fashion that $|b|^2 = |c|^2 = |d|^2 = \frac{1}{2}$.

The constants, a, b, c and d , are complex which means that they have an amplitude and a phase. For example, $a = |a|e^{i\theta}$ for some $\theta \in [0, 2\pi)$.

Discussion 4.3

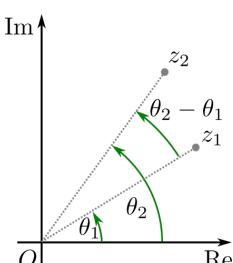
Explicitly write the complex number, $a = 1 + 2i$, in the form $|a|e^{i\theta}$.

Discussion 4.4

We will see that in general, the calculation of probabilities involves calculating the modulus of the inner product between two kets $|\langle \varphi|\psi\rangle|^2$. Show that the modulus $|\langle \varphi|\psi\rangle|^2$, is unaffected by changing a state $|\psi\rangle$ to a state $e^{i\beta}|\psi\rangle$. This tells us so that the overall phase (or a phase that we factorise out) of a state does not effect the probability of something happening. We say that an overall phase is **not physically relevant**.

The overall phase of a quantum state vector is not physically relevant (Discussion 4.4) which means that only the relative phases

Suppose we have,
 $|\psi\rangle = z_1|+\rangle + z_2|-\rangle = |z_1|e^{i\theta_1}|+\rangle + |z_2|e^{i\theta_2}|-\rangle$. Then factorising out the phase $e^{i\theta_1}$, we get $|\psi\rangle = e^{i\theta_1}(|z_1| |+\rangle + |z_2|e^{i(\theta_2-\theta_1)}|-\rangle)$. The overall phase is then $e^{i\theta_1}$. We call $e^{i(\theta_2-\theta_1)}$ the relative phase between the coefficients. It is the difference between the phases of z_1 and z_2 . As shown below, if we represent z_1 and z_2 as points on the complex plane, the relative phase is the angle between z_1 and z_2 .



between different components have any meaning. This means we can choose one of the coefficients of each vector to be real & positive (Discussion 4.5). This in turn tells us that we can take (Discussion 4.6):

$$\begin{aligned} |+\rangle_x &= \frac{1}{\sqrt{2}} (|+\rangle + e^{i\alpha} |-\rangle) \\ |-\rangle_x &= \frac{1}{\sqrt{2}} (|+\rangle + e^{i\beta} |-\rangle), \end{aligned} \quad (4.8)$$

where α and β are unknown relative phases.

Overall phase is irrelevant

Discussion 4.5

Explain how the statement “only the relative phases between different components of a state is physically meaningful” implies we “can choose one of the coefficients of each vector to be real & positive”

Discussion 4.6

Given $|a|^2 = |b|^2 = |c|^2 = |d|^2 = 1/2$, explain how we use the fact that we “can choose one of the coefficients of each vector to be real & positive” to go from

$$\begin{aligned} |+\rangle_x &= a |+\rangle + b |-\rangle \\ |-\rangle_x &= c |+\rangle + d |-\rangle \end{aligned} \quad (4.9)$$

to

$$\begin{aligned} |+\rangle_x &= \frac{1}{\sqrt{2}} (|+\rangle + e^{i\alpha} |-\rangle) \\ |-\rangle_x &= \frac{1}{\sqrt{2}} (|+\rangle + e^{i\beta} |-\rangle) \end{aligned} \quad (4.10)$$

We have now used up all the information from **Experiment 2** but still haven’t completely determined $|\pm\rangle_x$ – we need more information.

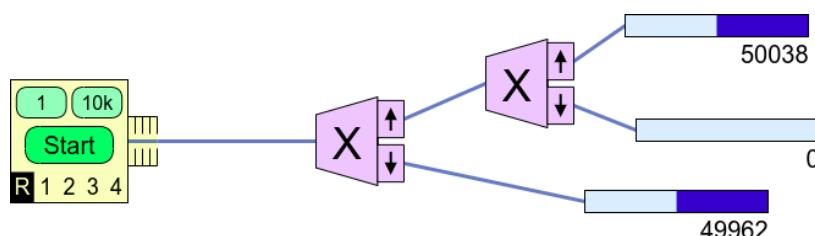


Figure 4.6: **Experiment 1** with both analysers aligned along the x -axis

If one performs **Experiment 1** with both analysers aligned along the x -axis, as in Figure 4.6, one finds that, all $|+\rangle_x$ states from the first analyser will be measured to have $S_x = \frac{1}{2}\hbar$ at the second analyser. In other words, all atoms exit the 2nd analyser in the $|+\rangle_x$ state. On the other hand, atoms entering the 2nd analyser with $|\psi_{in}\rangle = |+\rangle_x$, have a 0% chance to leave in the $|-\rangle_x$ state. This tells us

that

$$\mathcal{P}_{-x} = |_x\langle -|\psi_{\text{in}}\rangle|^2 = |_x\langle -|+\rangle_x|^2 = 0 ,$$

which means $|+\rangle_x$ and $|-\rangle_x$ are orthogonal to each other, so that,

$$\begin{aligned} 0 &= _x\langle -|+\rangle_x \\ &= \frac{1}{\sqrt{2}} \left(\langle +| + e^{-i\beta} \langle -| \right) \frac{1}{\sqrt{2}} \left(|+ \rangle + e^{i\alpha} |-\rangle \right) \\ &= \frac{1}{2} \left(1 + e^{i(\alpha-\beta)} \right) \\ \implies e^{i(\alpha-\beta)} &= -1 \\ \implies e^{i\alpha} &= -e^{i\beta} \end{aligned} \tag{4.11}$$

Discussion 4.7

Show how we go from the 2nd to the 3rd line of (4.11).

We found, $e^{i\alpha} = -e^{i\beta}$, but we *still* need some more information to determine their values. Unfortunately there is no more information on hand, but we have not specified exactly how the x -axis is oriented in the plane orthogonal to the z -axis. It turns out that choosing a particular value for α picks out a particular direction for the x -axis. For simplicity we take $\alpha = 0$, so that

$$\begin{aligned} |+\rangle_x &= \frac{1}{\sqrt{2}} \left(|+ \rangle + e^{i\alpha} |-\rangle \right) \\ |-\rangle_x &= \frac{1}{\sqrt{2}} \left(|+ \rangle - e^{i\alpha} |-\rangle \right) , \end{aligned} \tag{4.12}$$

becomes

$$\begin{aligned} |+\rangle_x &= \frac{1}{\sqrt{2}} (|+ \rangle + |-\rangle) \\ |-\rangle_x &= \frac{1}{\sqrt{2}} (|+ \rangle - |-\rangle) . \end{aligned} \tag{4.13}$$

A choice of basis

Usually, in this course we will use the S_z basis, $|\pm\rangle$. In principle, we could use any basis we like. To use the S_x basis, $|\pm\rangle_x$, you need to solve (4.13) for $|\pm\rangle$ which gives

$$\begin{aligned} |+\rangle &= \frac{1}{\sqrt{2}} (|+\rangle_x + |-\rangle_x) \\ |-\rangle &= \frac{1}{\sqrt{2}} (|+\rangle_x - |-\rangle_x) . \end{aligned} \tag{4.14}$$

Discussion 4.8

Show (4.14) by solving (4.13) for $|\pm\rangle$.

Looking at (4.14), we see that the $|+\rangle$ state is a combination of the $|+\rangle_x$ and $|-\rangle_x$ states. The coefficients in this expression tell us that there's a 50/50 chance to measure the spin as either up or down along the x -axis if we have a $|+\rangle$ state. Next we will now carefully talk about such combinations of states.

Superposition states and Mixed states

Classical physics is **deterministic**. This means that if we know the initial conditions¹ of all the particles in a system, using Newton's laws, we can *in principle* predict exactly what the system will be doing at some later time. *In practise* this is often not possible. We may be ignorant about the exact initial conditions of the system or some other property of the system like, the masses of all its particles or even the number of particles in the system. Another problem might be that the equations of motion are too complicated for us to solve². Think about flipping a coin. If we knew the initial velocity, angular momentum and moment of inertial of the coin we could predict whether it would land on heads or tails³. In practise, this is rather difficult to do and coin flips *appear* to be random, but, since the underlying physics is deterministic, this **effective randomness** is a function of our ignorance.

Quantum physics does not appear to be **deterministic**. Even if we know everything about a system we may not be able to predict the outcome of some experiments with absolute certainty. For example, if we know a particle is in the state $|+\rangle$, we can not, be certain about what will happen if we measure the x -component of its spin. This **inherent randomness** appears to be a fundamental feature of quantum mechanics⁴.

We can write a general spin- $\frac{1}{2}$ state vector $|\psi\rangle$ as a combination of basis kets:

$$|\psi\rangle = a|+\rangle + b|-\rangle .$$

Such a state is called a **superposition state**. To understand why such states are important, we need to look at another kind of state, a **mixed state**.

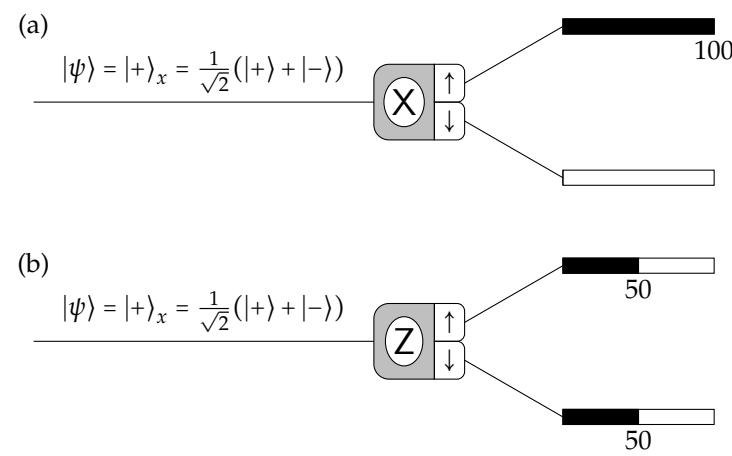


Figure 4.7 shows the results of measurements on the superposition state $|\psi\rangle = \frac{1}{\sqrt{2}}(|+ \rangle + | - \rangle)$. From the previous section, you should recognise that $|\psi\rangle = |+ \rangle_x$. For this state, if we measure S_x , we *always* find $S_x = \hbar/2$. On the other hand if we measure S_z for this state we find a 50/50 split between $S_z = \hbar/2$ and $S_z = -\hbar/2$. For the measurement in Figure 4.7 (b), it looks as if a beam of atoms in the $|\psi\rangle = |+ \rangle_x$

¹ The initial conditions that we need are the positions and velocities of every particle.

² Another factor may be chaos. In chaotic systems, a tiny change in the initial conditions, can result in large changes in the behaviour of the system after sometime. This exquisite sensitivity to initial conditions, means that, even if we know the initial conditions very accurately, we will only be able to predict how the system will behave for a limited time period.

³ We would also need to be able to accurately calculate the effect of air resistance

⁴ Hidden variable theories are an attempt to recast the randomness of quantum mechanics as being effective randomness due to our ignorance about possible hidden variables. We will return to these ideas when we look at Bell's Theorem.

Figure 4.7: Superposition state measurements

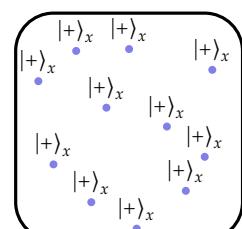


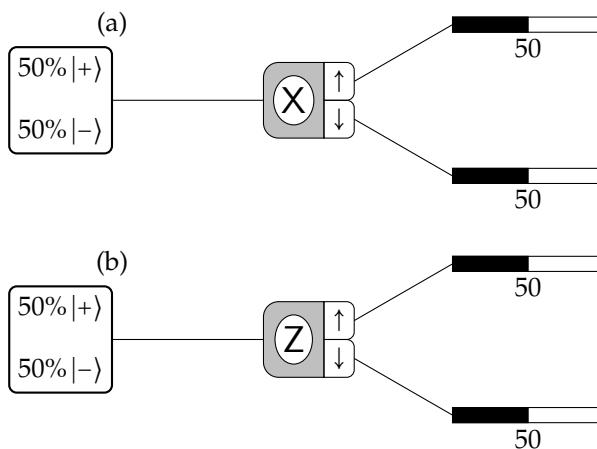
Figure 4.8: Representation of an ensemble of atoms in the $|+ \rangle_x$ -state

state, is made up of a beam with a 50/50 mixture of atoms in the $|+\rangle$ and $|-\rangle$ states.

When we talk about a state in the Stern-Gerlach experiment, we are really talking about an ensemble or collections of atoms on which the experiment is prepared. For a superposition state, every member of the ensemble is in the same state. For example, fig. 4.8 schematically shows an ensemble of atoms in the $|+\rangle_x$ state while fig. 4.9 shows an ensemble of atoms in the $|\psi\rangle = a|+\rangle + b|-\rangle$ state.

We could also have an ensemble in which not each atom is in the same state – we call this a mixed state. For example fig. 4.10 shows an ensemble with a 50/50 mix of $|+\rangle$ and $|-\rangle$ states. In a mixed state, we don't know what state a particular atom is in – there is some probability for an atom to be in one state or another. This is different from the probabilities we've seen so far which relate to the probability of making a particular spin-component measurement for a particular *given* state. As we shall see, mixed states are very different from a superposition state.

Figure 4.11 shows the results of measurements on a mixed state. Atoms in the input beam in Figure 4.11 have a 50% chance to be in the $|+\rangle$ state and a 50% chance to be in the $|-\rangle$ state. Comparing Figure 4.7 and Figure 4.11 we see that the superposition state, $|+\rangle_x$, does not produce the same experimental results and a 50/50 mix of the $|\pm\rangle$ states.



Let's use what we know about the $|\pm\rangle$ states to understand Figure 4.11. Looking at the top part of Figure 4.7, we know that measuring S_z for atoms in the $|+\rangle$ state we always find $S_z = \hbar/2$. Similarly atoms in the $|-\rangle$ state are always measured to have $S_z = -\hbar/2$. This means that if I have a 50/50 mixture of atoms in the $|+\rangle$ and $|-\rangle$ states, 50% of the time I will measure $S_z = \hbar/2$ and 50% of the time I will measure $S_z = -\hbar/2$. This is the same as what is observed for the $|+\rangle_x$ in the top part of Figure 4.7.

The difference between $|+\rangle_x$ and the mixed state in Figure 4.11 is evident when we look at measurements of S_x . We saw before that atoms in the state $|+\rangle$ are found to have, on average, $S_x = \hbar/2$ half of the time and $S_x = -\hbar/2$ the other 50% of the time. The same split

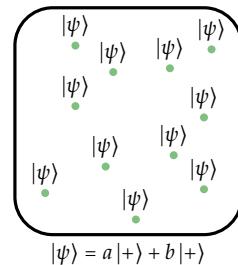


Figure 4.9: An ensemble of atoms in the $(a|+\rangle + b|-\rangle)$ -state

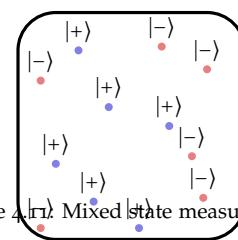


Figure 4.10: Ensemble of atoms that is a 50/50 mix of $|+\rangle$ and $|-\rangle$ states.

was seen for $|-\rangle$ state atoms. Putting this together we will observe a 50/50 split between $S_x = \hbar/2$ & $S_x = -\hbar/2$ for atoms in the mixed state. This is very different from the $|+\rangle_x$ state which we always measure to have $S_x = \hbar/2$. In conclusion, experiments show $|+\rangle_x$ isn't a mixed state with a 50/50 split of atoms in the $|+\rangle$ & $|-\rangle$ states, it is a superposition of the $|+\rangle$ and $|-\rangle$ states.

The default "R" state for the SPINS program is a mixed state with a 50/50 mix of the $|\pm\rangle$ states. How would you confirm this using the program? For the rest of this course, we will mainly concentrate on superposition states.

Example 4.1:

Problem Consider the state $|\psi\rangle = 3|+\rangle + 4|-\rangle$. Normalise it and find the probabilities of measuring $S_z = \pm\hbar/2$.

Solution Introducing a normalisation constant N , we write $|\psi\rangle = N(3|+\rangle + 4|-\rangle)$. Then requiring $1 = \langle\psi|\psi\rangle$ leads to

$$\begin{aligned} 1 &= \langle\psi|\psi\rangle \\ &= N^* (3|+| + 4|-|) N (3|+| + 4|-|) \\ &= |N|^2 (9(+|+) + 12(+|-) + 12(-|+) + 16(-|-)) \\ &= 25|N|^2 \\ \implies |N|^2 &= \frac{1}{25} \end{aligned} \quad (4.15)$$

Using the fact that the overall phase of a state is physically irrelevant, we take N to be real and positive. This means that we can take $N = \frac{1}{5}$ so that

$$|\psi_{\text{in}}\rangle = \frac{3}{5}|+\rangle + \frac{4}{5}|-\rangle .$$

Using the 4th postulate tells us that, the probability of measuring $S_z = +\hbar/2$ for this state is $\mathcal{P}_+ = \left|\frac{3}{5}\right|^2 = \frac{9}{25} = 0.36$, where as the probability of measuring $S_z = -\hbar/2$ for this state is $\mathcal{P}_- = \left|\frac{4}{5}\right|^2 = \frac{16}{25} = 0.64$. This is summarised in Figure 4.12.

Discussion 4.9

What are \mathcal{P}_{+x} and \mathcal{P}_{-x} for $|\psi_{\text{in}}\rangle = \frac{3}{5}|+\rangle + \frac{4}{5}|-\rangle$? Draw the corresponding histogram.

Discussion 4.10

Write a summary of the important points and formulae in this chapter.

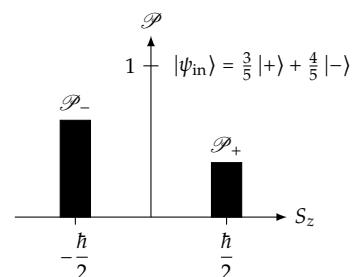


Figure 4.12: Histogram of S_z spin component measurements for $|\psi_{\text{in}}\rangle = \frac{3}{5}|+\rangle + \frac{4}{5}|-\rangle$.

5

New notation & Generalisations

Matrix notation

This reading is based on [McIntyre et al. 2012, Section 1.3-1.5]

So far, we have defined kets in terms of their inner products with other kets. In general we write a ket as

$$|\psi\rangle = (\langle +|\psi\rangle) |+\rangle + (\langle -|\psi\rangle) |-\rangle, \quad (5.1)$$

so, for example

$$\begin{aligned} |+\rangle_x &= (\langle +|+\rangle_x) |+\rangle + (\langle -|+\rangle_x) |-\rangle \\ &= \frac{1}{\sqrt{2}} |+\rangle + \frac{1}{\sqrt{2}} |-\rangle \end{aligned} \quad (5.2)$$

These kets are expressed in terms of the basis kets $|+\rangle$ and $|-\rangle$. Once we have agreed on these as our basis kets, we just need the two coefficients to specify our quantum state, namely $\langle +|\psi\rangle$ and $\langle -|\psi\rangle$. This means we can simplify our notation by just using two numbers. We can represent kets just using a column vector with the two coefficients. For example, we can represent $|+\rangle_x$ as

$$|+\rangle_x \doteq \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad (5.3)$$

where \doteq means “is represented by” and it is understood that we are using the $\{|+\rangle, |-\rangle\}$ basis.

Discussion 5.1

Explain why we do not just use *equals* in the formula $|+\rangle_x \doteq \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$. Explain why we *do* use an equal sign in the formula $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Our convention for the ordering of the amplitudes in our column vector representation of kets is to put the spin up amplitude first. For example,

$$|-\rangle_x \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \leftarrow |+\rangle \quad \leftarrow |-\rangle,$$

With this convention, the basis kets take on a very simple form

$$\begin{aligned} |+\rangle &\doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ |-\rangle &\doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned} \quad (5.4)$$

When written as column vectors, such a basis, in which all the entries of a vector are zero except for one entry which is 1, is called a **standard basis** (also called a natural basis or canonical basis).

💬 Discussion 5.2

The text book says that writing basis kets as:

$|+\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $|-\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ “demonstrates the important feature that *basis kets are unit vectors when written in their own basis.*” Although a bit clunky, I think it would be better to say “*basis kets constitute a standard basis in their own basis.*” Why do you think I have a problem with the author’s statement? **Hint:** my issue related to the usage of the word “when”.

This short-hand, in which we conveniently represent kets as a set of coefficients is called a **representation**. As we have used the S_z basis kets, this particular representation is called the S_z representation. Basis kets always have a simple form as in (5.4) in their own representation.

💬 Discussion 5.3

What do $|+\rangle_x$ and $|-\rangle_x$ look like in the S_x basis?

In this notation, a general ket, $|\psi\rangle = (\langle +|\psi\rangle) |+\rangle + (\langle -|\psi\rangle) |-\rangle$, is written

$$|\psi\rangle \doteq \begin{pmatrix} \langle +|\psi\rangle \\ \langle -|\psi\rangle \end{pmatrix}.$$

A nice feature of this notation, is that we can use matrix multiplication to write down **inner products**. To this end, we write bras as **row vectors**. Given a ket

$$\begin{aligned} |\psi\rangle &= a|+\rangle + b|-\rangle \\ &\doteq \begin{pmatrix} a \\ b \end{pmatrix}, \end{aligned} \quad (5.5)$$

the corresponding bra is

$$\begin{aligned} \langle\psi| &= a^* \langle +| + b^* \langle -| \\ &\doteq \begin{pmatrix} a^* & b^* \end{pmatrix}. \end{aligned} \quad (5.6)$$

With this notation, we can write inner products as the matrix

multiplication of a row and column vector

$$\langle \psi | \psi \rangle = \begin{pmatrix} a^* & b^* \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = |a|^2 + |b|^2 \quad (5.7)$$

Discussion 5.4

Explain why we **do** use *equals* (and not \doteq) in the formula

$$\langle \psi | \psi \rangle = \begin{pmatrix} a^* & b^* \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$$

Practise

To practise our new “matrix” notation, we’ll use the results of **Experiment 2** to determine $|+\rangle_y$ and $|-\rangle_y$ (in terms of the $|\pm\rangle$ basis). We’ll use a modified version of **Experiment 2** with the 2nd analyser aligned along the y -axis as shown in Figure 5.1.

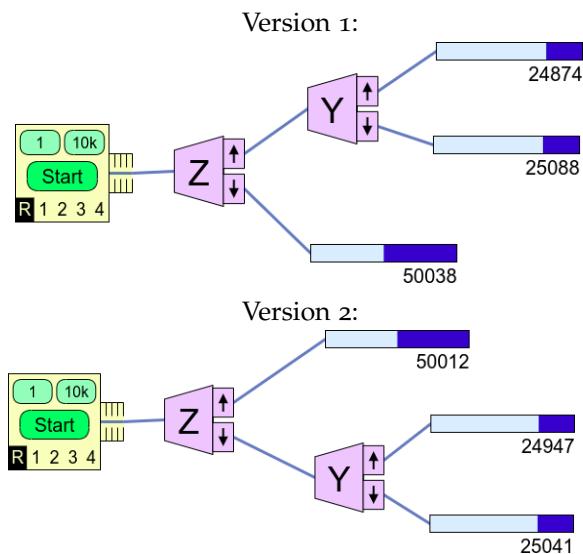


Figure 5.1: Experiment 2 repeated with second analyser along the y -axis.

The results are analogous to what we found before when the 2nd analyser was aligned along the x -direction:

$$\begin{aligned}\mathcal{P}_{1,+y} &= |_y\langle +|+\rangle|^2 = \frac{1}{2} \\ \mathcal{P}_{1,-y} &= |_y\langle -|+\rangle|^2 = \frac{1}{2} \\ \mathcal{P}_{2,+y} &= |_y\langle +|-|-\rangle|^2 = \frac{1}{2} \\ \mathcal{P}_{2,-y} &= |_y\langle -|-|-\rangle|^2 = \frac{1}{2}.\end{aligned}\quad (5.8)$$

The discussion and calculation will follow the same lines as we had for $|\pm\rangle_x$ – that is until we took $\alpha = 0$. This means we have

$$\begin{aligned}|+\rangle_y &= \frac{1}{\sqrt{2}} (|+\rangle + e^{i\alpha} |-\rangle) \\ |-\rangle_y &= \frac{1}{\sqrt{2}} (|+\rangle - e^{i\alpha} |-\rangle).\end{aligned}\quad (5.9)$$

Now we need more information to fix α , so we repeat **Experiment 2** with the first analyser along the x -direction as in Figure 5.2.

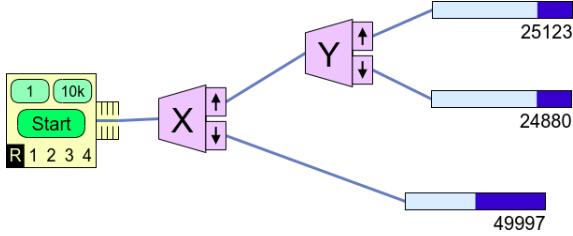


Figure 5.2: Experiment 2 repeated with 1st analyser along the x -axis and second analyser along the y -axis.

Looking at the \uparrow -output port of the y -analyser in Figure 5.2, we conclude that $\mathcal{P}_{+y} = |\langle \text{out} | \text{in} \rangle|^2 = |\langle y | + \rangle_x|^2 = \frac{1}{2}$. To use matrix algebra, we first need to express our bras and kets in matrix notation. So,

$$|+\rangle_y = \frac{1}{\sqrt{2}} (|+\rangle + e^{i\alpha} |-\rangle) \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ e^{i\alpha} \end{pmatrix},$$

which tells us that

$$_y \langle + | \doteq \frac{1}{\sqrt{2}} (1 \quad e^{-i\alpha}).$$

We also need

$$|+\rangle_x = \frac{1}{\sqrt{2}} (|+\rangle + |-\rangle) \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Now we can calculate

$$\begin{aligned} _y \langle + | + \rangle_x &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & e^{-i\alpha} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \frac{1}{2} (1 + e^{-i\alpha}), \end{aligned} \quad (5.10)$$

which means that

$$\begin{aligned} \mathcal{P}_{+y} &= \frac{1}{2} = |_y \langle + | + \rangle_x|^2 \\ &= \frac{1}{2} (1 + e^{-i\alpha}) \frac{1}{2} (1 + e^{i\alpha}) \\ &= \frac{1}{4} (1 + 1 + e^{-i\alpha} + e^{i\alpha}) \\ &= \frac{1}{2} + \frac{1}{2} \underbrace{\left(\frac{e^{-i\alpha} + e^{i\alpha}}{2} \right)}_{\cos \alpha}. \end{aligned} \quad (5.11)$$

Bringing it all together, we see that

$$\mathcal{P}_{+y} = \frac{1}{2} = \frac{1}{2} (1 + \cos \alpha) \implies \cos \alpha = 0 \implies \alpha = \pm \pi/2.$$

It turns out that $\alpha = +\pi/2$ corresponds to a right-handed coordinate system, so this is the solution we choose. This leads to

$$e^{i\alpha} = e^{i\pi/2} = i. \quad (5.12)$$

Substituting (5.12) into our expressions for $|\pm\rangle_y$ gives

$$\begin{aligned} |+\rangle_y &= \frac{1}{\sqrt{2}} (|+\rangle + i|-\rangle) \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \\ |-\rangle_y &= \frac{1}{\sqrt{2}} (|+\rangle - i|-\rangle) \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}. \end{aligned} \quad (5.13)$$

Notice that complex coefficients are **required** in our expressions. On the other hand, probabilities, which involve the complex squares, $|\langle \text{out} | \text{in} \rangle|^2$, will *always* be **real** and **non-negative** as expected.

General quantum systems

We can extend the tools we've learnt about for spin- $\frac{1}{2}$ systems. For example, suppose we have an observable which gives us a_n possible results with $n = 1, 2, 3, \dots$, or N . We can represent this schematically, for with $N = 3$, as in Figure 5.3 .

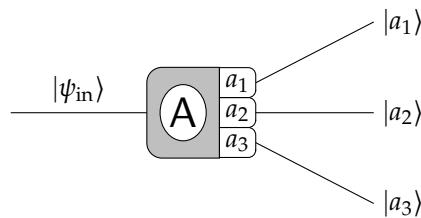


Figure 5.3: Generic depiction of the quantum mechanical measurement of observable A .

In Figure 5.3, A labels the measurement devices, while a_1 , a_2 and a_3 are possible measurements which we use to label the output ports.

General rules for $|\text{Kets}\rangle$

First we need to define the **Kronecker delta**, δ_{ij} :

$$\delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}, \quad (5.14)$$

which allows us to compactly express the **orthonormality** condition for the basis kets:

$$\langle a_i | a_j \rangle = \delta_{ij}.$$

We also have the property of **completeness** which means that any generic state can be written in terms of our basis kets:

$$|\psi\rangle = \sum_{i=1}^N \langle a_i | \psi \rangle |a_i\rangle$$

Generalisation of the 4th postulate

The probability of measuring a_n for some state $|\psi\rangle$ is
 $\mathcal{P}_{a_n} = |\langle a_n | \psi \rangle|^2$.

Example 5.1:

Problem Let's look at a system with the observable A which can be measured to have 3 possible values: a_1 , a_2 or a_3 . Suppose we have the state

$$|\psi\rangle = C(2|a_1\rangle - 3|a_2\rangle + 4i|a_3\rangle).$$

- (i) Normalise $|\psi\rangle$
- (ii) Determine the probability of measuring a_1 , a_2 , or a_3 .

Solution

- (i) For a state to be normalised we must have $\langle\psi|\psi\rangle = 1$ so,

$$\begin{aligned}
 1 &= \langle\psi|\psi\rangle \\
 &= C^* (2\langle a_1 | - 3\langle a_2 | + 4i\langle a_3 |) C (2|a_1\rangle - 3|a_2\rangle + 4i|a_3\rangle) \\
 &= |C|^2 (4\langle a_1 | a_1 \rangle - 6\langle a_1 | a_2 \rangle + 8i\langle a_1 | a_3 \rangle - 6\langle a_2 | a_1 \rangle \\
 &\quad + 9\langle a_2 | a_2 \rangle - 12i\langle a_2 | a_3 \rangle - 8i\langle a_3 | a_1 \rangle + 12i\langle a_3 | a_2 \rangle + 16\langle a_3 | a_3 \rangle) \\
 &= |C|^2 (4 + 9 + 16) = |C|^2 29 \\
 \implies C &= 1/\sqrt{29}
 \end{aligned}$$

So the normalised ket is

$$|\psi\rangle = \frac{1}{\sqrt{29}} (2|a_1\rangle - 3|a_2\rangle + 4i|a_3\rangle)$$

- (ii) Using the generalised 4th postulate,

$$\begin{aligned}
 \mathcal{P}_{a_1} &= |\langle a_1 | \psi \rangle|^2 \\
 &= \left| \langle a_1 | \left(\frac{1}{\sqrt{29}} (2|a_1\rangle - 3|a_2\rangle + 4i|a_3\rangle) \right) \right|^2 \\
 &= \left| \frac{2}{\sqrt{29}} \underbrace{\langle a_1 | a_1 \rangle}_{\delta_{11}=1} - \frac{3}{\sqrt{29}} \underbrace{\langle a_1 | a_2 \rangle}_{\delta_{12}=0} + \frac{4i}{\sqrt{29}} \underbrace{\langle a_1 | a_3 \rangle}_{\delta_{13}=0} \right|^2 = \left| \frac{2}{\sqrt{29}} \right|^2 \\
 &= \frac{4}{29} \\
 \mathcal{P}_{a_2} &= |\langle a_2 | \psi \rangle|^2 \\
 &= \left| \langle a_2 | \left(\frac{1}{\sqrt{29}} (2|a_1\rangle - 3|a_2\rangle + 4i|a_3\rangle) \right) \right|^2 \\
 &= \left| -\frac{3}{\sqrt{29}} \right|^2 \\
 &= \frac{9}{29} \\
 \mathcal{P}_{a_3} &= |\langle a_3 | \psi \rangle|^2 \\
 &= \left| \langle a_3 | \left(\frac{1}{\sqrt{29}} (2|a_1\rangle - 3|a_2\rangle + 4i|a_3\rangle) \right) \right|^2 \\
 &= \left| \frac{4i}{\sqrt{29}} \right|^2 \\
 &= \frac{16}{29}
 \end{aligned}$$

Figure 5.4 shows a summary of the calculated probabilities for

$$|\psi\rangle = \frac{1}{\sqrt{29}} (2|a_1\rangle - 3|a_2\rangle + 4i|a_3\rangle)$$

Discussion 5.5

Repeat the calculation of the normalisation of $|\psi\rangle = C (2|a_1\rangle - 3|a_2\rangle + 4i|a_3\rangle)$ and the probability \mathcal{P}_{a_1} using matrix notation. You should work in the representation

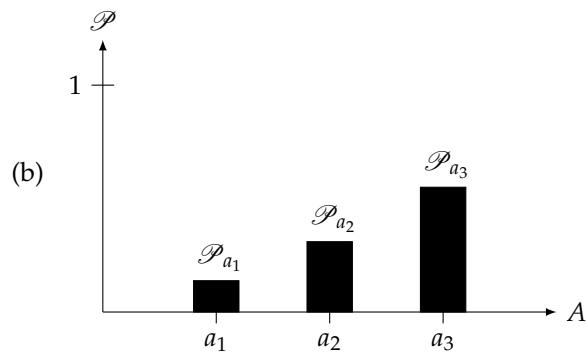
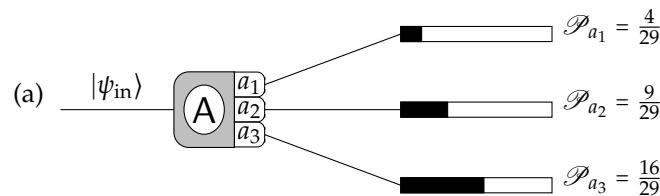


Figure 5.4: (a) Schematic diagram of the measurement of observable A and (b) histogram of the predicted measurement probabilities.

in which, $|a_1\rangle \doteq \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $|a_2\rangle \doteq \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and $|a_3\rangle \doteq \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

5.1 Postulates of Quantum Mechanics

So far we have seen two of the postulates of quantum mechanics. These postulates give us the recipe for how quantum mechanics works & how to use it. These postulates are not **proven** but **tested** by many, many experiments to an incredibly high degree of accuracy. In principle new results could force a reevaluation...

Although you still need to learn a bit more before you can understand all of the postulates, they are listed below to give you a flavour of what they entail:

Postulates of Quantum Mechanics

1. The state of a quantum mechanical system, including all the information you can know about it, is represented mathematically by a normalized ket $|\psi\rangle$.
2. A physical observable is represented mathematically by an operator \hat{A} that acts on kets.
3. The only possible result of a measurement of an observable is one of the eigenvalues a_n of the corresponding operator \hat{A} .
4. The probability of obtaining the eigenvalue a_n in a measurement of the observable A on the system in the state $|\psi\rangle$ is $\mathcal{P}_{a_n} = |\langle a_n | \psi \rangle|^2$ where $|a_n\rangle$ is the normalized eigenvector of \hat{A} corresponding to the eigenvalue a_n .
5. After a measurement of A that yields the result a_n , the quantum system is in a new state that is the normalized projection of the original system ket onto the ket (or kets) corresponding to the result of the measurement:

$$|\psi'\rangle = \frac{\hat{P}_n |\psi\rangle}{\sqrt{\langle \psi | \hat{P}_n | \psi \rangle}} .$$
6. The time evolution of a quantum system is determined by the Hamiltonian or total energy operator $\hat{H}(t)$ through the Schrödinger equation $i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H}(t) |\psi\rangle$.

Discussion 5.6

Write a summary of the important points and formulae in this chapter.

6

Operators and Measurement

Introduction

This reading is based on [McIntyre et al. 2012, Section 2.1]

In the first chapter, you learnt how to predict the results of Stern-Gerlach experiment with analysers aligned along the x , y or z -axis and all at 90° to each other. This is all very well, but what about predicting experiments we haven't performed – for example, with analysers along some arbitrary directions. To do this we need to learn about **quantum operators**.

6.1 Operators, Eigenvectors & Eigenvalues

So far we have concentrated on **quantum state vectors** or kets. We have used kets to calculate **probabilities**. Each observable quantity (S_x , S_y and S_z) has been associated with a pair of kets corresponding to possible measurements. The **observables** themselves have not been included. To fill this gap we need new tools to relate observables and states.

The role of observables is provided by two new postulates. The first one we have is:

QM Postulate 2

A physical observable is represented mathematically by an **operator**, \hat{A} , that **acts** on kets.

To understand this we need some new terminology.

Definition

An **operator** is a mathematical object that takes a ket and gives a ket back to you, for example $\hat{A}|\psi\rangle = |\varphi\rangle$.

Technically we require our operators to be linear which means that $\hat{A}(\lambda|\psi\rangle) = \lambda\hat{A}|\psi\rangle$ and $\hat{A}(|\psi\rangle + |\gamma\rangle) = \hat{A}|\psi\rangle + \hat{A}|\gamma\rangle$.

Operators

Schematically, an operator is a map or function on Hilbert space:



Example 6.1: Matrices and vectors

You may have seen operators before in the context of vectors and matrices. If we represent the vector, $\vec{a} = a_x \hat{i} + a_y \hat{j} + a_z \hat{k}$, using a 3×1 column vector,

$$\vec{a} = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix}, \quad (6.1)$$

then multiplying the right hand side of (6.1) by a 3×3 matrix produces a 3×1 column vector. This means that 3×3 matrices are **operators** that act on 3×1 column vectors by matrix multiplication. For example, the matrix

$$\mathbf{P} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

will flip the vector since

$$\mathbf{P} \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} = \begin{pmatrix} -a_x \\ -a_y \\ -a_z \end{pmatrix}.$$

The way that \mathbf{P} operates on a vector is illustrated in Figure 6.1. Another example is the rotation matrix

$$\mathbf{R}_z = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

which rotates vectors by an angle θ about the z -axis. The way that \mathbf{R}_z operates on a vector (in the xy -plane) is illustrated in Figure 6.2.

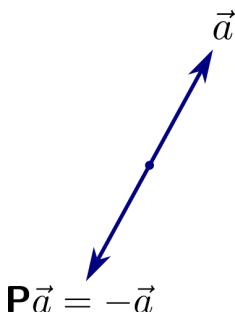


Figure 6.1: The action of the operator \mathbf{P} on a vector \vec{A}

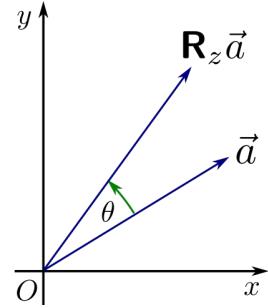


Figure 6.2: The action of the operator \mathbf{R}_z on a vector \vec{A}

Example 6.2: Differential operators

Another class of operators you may be familiar with are *differential* operators. For example, the operator, $\hat{D}_a = \frac{d}{dx}$, acts on (differentiable) functions, as in say

$$\hat{D}_a f(x) = \frac{d}{dx} f(x) = f'(x).$$

An example of a more complicated differential operators is, $\hat{D}_b = \frac{d^2}{dx^2} + l^2$, where l^2 is a constant. \hat{D}_b acts on a function $f(x)$ as follows

$$\hat{D}_b f(x) = \left(\frac{d^2}{dx^2} + l^2 \right) f(x) = f''(x) + l^2 f(x).$$

Notice that we often use a “hat”, $\hat{}$, to denote operators.

There are *special kets* that are not changed by the action of a *particular operator* – except that is, up to being multiplied by an overall constant:

$$\hat{A} |\psi\rangle = \underbrace{\lambda}_{\text{constant}} |\psi\rangle. \quad (6.2)$$

¹ or eigenkets

Such special kets are called¹ **eigenvectors of the operator \hat{A}** . We call the multiplicative constants the **eigenvalues** of the operator and (6.2) is called an **eigenvalue equation**. A geometric interpretation of (6.2), in the context of spacial vectors, is shown in Figure 6.3.

Example 6.3: Eigenvectors and Eigenvalues

To get a feel for eigenvectors and eigenvalues, let's look at some examples in a more familiar context. Any vector is an eigenvector of the matrix **P** with eigenvalue -1 , since

$$\mathbf{P}\vec{A} = \underbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}}_{\text{Operator}} \underbrace{\begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix}}_{\text{eigenvector}} = \underbrace{-1}_{\text{eigenvalue}} \underbrace{\begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix}}_{\text{eigenvector}}.$$

On the other hand, vectors pointing in the z -direction are eigenvectors of the rotation matrix, **R_z** with eigenvalue 1 since

$$\mathbf{R}_z \vec{A} = \underbrace{\begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\text{Operator}} \underbrace{\begin{pmatrix} 0 \\ 0 \\ A_z \end{pmatrix}}_{\text{eigenvector}} = \underbrace{1}_{\text{eigenvalue}} \underbrace{\begin{pmatrix} 0 \\ 0 \\ A_z \end{pmatrix}}_{\text{eigenvector}}$$

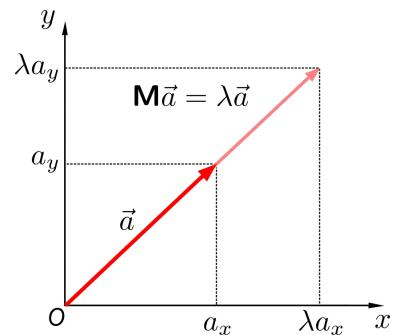


Figure 6.3: The matrix **M** stretches the vector \vec{a} but does not change its line of action meaning \vec{a} is an eigenvector of **M**.

Modification of this image
Attribution: Lyudmil Antonov, CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0/>, via Wikimedia Commons

Discussion 6.1

Show that 2×1 column vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ are eigenvectors of the matrix $\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$. What are the corresponding eigenvalues?

Discussion 6.2

Show that the *complex* valued vectors $\begin{pmatrix} 1 \\ i \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -i \end{pmatrix}$ are eigenvectors of **R_z**. What are the corresponding eigenvalues? Use the Euler relation to simplify your answer.

Example 6.4: Eigenfunctions

In the context of differential operators, we usually talk about **eigenfunctions** rather than eigenvectors.

For example, e^{qx} is an eigenfunctions of $\hat{D}_a = \frac{d}{dx}$ with eigenvalue, q , since

$$\hat{D}_a e^{qx} = \underbrace{\frac{d}{dx}}_{\text{Operator}} \underbrace{e^{qx}}_{\text{eigenfunction}} = \underbrace{q}_{\text{eigenvalue}} \underbrace{e^{qx}}_{\text{eigenfunction}} .$$

Similarly, e^{-qx} is an eigenfunctions of \hat{D}_a with eigenvalue, $-q$, since $\hat{D}_a e^{-qx} = -qe^{-qx}$.

Discussion 6.3

Write out the eigenvalue equation and show that e^{ikx} , e^{-ikx} , $\cos(kx)$ and $\sin(kx)$ are eigenfunctions of $\hat{D}_c = \frac{d^2}{dx^2}$. What are the corresponding eigenvalues?

Notice that the eigenvalue equations for differential operators are differential equations.

What are the eigenfunctions and eigenvalues of the operator

$$\hat{D}_b = \frac{d^2}{dx^2} + l^2 ?$$

Going back to quantum mechanics, eigenvalues are important because of postulate 3:

QM Postulate 3

The *only* possible result of a measurement of an observable is one of the *eigenvalues* a_n of the corresponding operator \hat{A} .

This is the promised mathematical relationship between physical observables, eg. S_z , possible measurements, eg. $S_z = \pm \hbar/2$, and kets, eg. $|\pm\rangle$. The relationship,

$$\underbrace{\hat{A}}_{\text{operator}} \underbrace{|\psi\rangle}_{\text{eigenvector}} = \underbrace{\hat{a}}_{\text{eigenvalue}} \underbrace{|\psi\rangle}_{\text{eigenvector}}, \quad (6.3)$$

is a quantum mechanical **eigenvalue equation**. So for example, for \hat{S}_z and $|+\rangle$ we have the eigenvalue equation

$$\hat{S}_z |+\rangle = +\hbar/2 |+\rangle . \quad (6.4)$$

The other eigenvalue equation for \hat{S}_z is

$$\hat{S}_z |-\rangle = -\hbar/2 |-\rangle . \quad (6.5)$$

Notation

We use hats to differentiate between operators and the observables that they correspond to. This means that \hat{S}_z is the operator corresponding to the observed quantity $S_z = \pm\hbar/2$.

Warning: the textbook McIntyre et al. [2012] is often not careful about making this distinction and you may have to work out from the context whether they are referring to an operator or an observable.

We could just work with kets², but it will be useful to use matrix equations to represent operators. Writing (6.4) as a matrix equation leads to:

$$\begin{array}{c} \hat{S}_z \\ \sim \end{array} \begin{array}{c} |+\rangle \\ \sim \end{array} = \underbrace{\frac{\hbar}{2} |+\rangle}_{\begin{pmatrix} ? & ? \\ ? & ? \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}} . \quad (6.6)$$

Since the kets are represented by 2×1 columns, \hat{S}_z needs to be represented by a 2×2 matrix for (6.6) to work as a matrix equation.

Discussion 6.4

Suppose \hat{S}_z is represented by a $m \times n$ matrix. Remember m is the number of rows and n is the number of columns. Can you explain why we need $m = n = 2$ for (6.6) to work as a matrix equation? Hint 1: What does n have to be for matrix multiplication to work? Hint 2: What does m have to be so that we get a 2×1 column matrix on the RHS of (6.6) ?

To find the matrix representation of \hat{S}_z (with respect to the $|\pm\rangle$ basis) we assume that \hat{S}_z has the form of a general 2×2 matrix:

$$\hat{S}_z \doteq \begin{pmatrix} a & b \\ c & d \end{pmatrix} . \quad (6.7)$$

Just like we labelled parts of the ket column vector using our basis bras, we can do the same for operators:

\hat{S}_z	$ +\rangle$	$ -\rangle$	$ +\rangle$	$ -\rangle$
$\langle + $	a	b	1	0
$\langle - $	c	d	0	1

Now, using (6.7) allows us to write (6.4) and (6.5) in matrix form as follows

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = +\frac{\hbar}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \quad (6.8)$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\frac{\hbar}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} , \quad (6.9)$$

which implies that,

$$\begin{aligned} a &= +\frac{\hbar}{2}, & b &= 0, \\ c &= 0, & d &= -\frac{\hbar}{2} . \end{aligned} \quad (6.10)$$

² We can use (6.4) and (6.5) to work out how \hat{S}_z acts on a general ket $|\psi\rangle = a|+\rangle + b|-\rangle$ as follows

$$\begin{aligned} \hat{S}_z |\psi\rangle &= \hat{S}_z(a|+\rangle + b|-\rangle) \\ &= (a\hat{S}_z|+\rangle + b\hat{S}_z|-\rangle) \\ &\stackrel{(6.4)(6.5)}{=} a\frac{\hbar}{2}|+\rangle + b(-\frac{\hbar}{2})|-\rangle \\ &= \frac{\hbar}{2}(a|+\rangle - b|-\rangle) . \end{aligned}$$

In the above calculation we have assumed that we can take \hat{S}_z inside the brackets in the second step. Mathematicians would say that we have assumed that \hat{S}_z acts linearly. Strictly speaking, in QM Postulate 2, we should specify that our operators are linear.

Discussion 6.5

Show that multiplying out the eigenvalue equations for \hat{S}_z , (6.8) and (6.9), leads to, $\begin{pmatrix} a \\ c \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} b \\ d \end{pmatrix} = -\frac{\hbar}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, respectively. Then explain how this lets you conclude that, $a = \hbar/2$, $b = 0$, $c = 0$ and $d = -\hbar/2$.

Using (6.7) and (6.10) we finally get

$$\hat{S}_z \doteq \begin{pmatrix} +\hbar/2 & 0 \\ 0 & -\hbar/2 \end{pmatrix}. \quad (6.11)$$

Note

1. This representation of \hat{S}_z is diagonal. This means that it only has non-zero entries on the diagonal.
2. The diagonal elements are equal to the eigenvalues & in the same order as the corresponding eigenvectors

A general rule

The two points noted above are special cases of a general rule:

An operator's representation is always diagonal when expressed with respect to a basis constructed from its eigenvectors. This is similar to the fact that eigenvectors always look simple in their own basis, for example, $|+\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

Notation

Following, Griffiths [2016], we will use a bold sans-serif font when referring specifically to the matrix representation of an operator. For example, we can write (6.11) as

$$\mathbf{S}_z \doteq \begin{pmatrix} +\hbar/2 & 0 \\ 0 & -\hbar/2 \end{pmatrix}. \quad (6.12)$$

When writing things out by hand, you can denote matrices using a wavy underline \underline{S}_z .

Matrix representation of operators

A general operator \hat{A} describing some observable (in a spin- $1/2$ system), can be represented as a 2×2 matrix, \mathbf{A} , (with respect to say

the S_z basis):

$$\hat{A} \doteq \mathbf{A} \doteq \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \quad (6.13)$$

Let's look at how \hat{A} acts on a basis ket:

$$\underbrace{\hat{A}|+\rangle}_{\text{new ket}} \doteq \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \underbrace{\begin{pmatrix} a \\ c \end{pmatrix}}_{\doteq \hat{A}|+\rangle}. \quad (6.14)$$

The inner product of this new ket with $|+\rangle$ gives

$$\langle + | (\hat{A}|+\rangle) = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ c \end{pmatrix} = a. \quad (6.15)$$

Discussion 6.6

Given a general operator \hat{A} in the S_z representation with $\hat{A} \doteq \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, show that $\langle +|\hat{A}|-\rangle = b$, $\langle -|\hat{A}|+\rangle = c$ and $\langle -|\hat{A}|-\rangle = d$.

Matrix elements are bra, operator, ket “sandwiches” 🍔 :

$$\langle \text{bra} | \widehat{\text{Operator}} | \text{ket} \rangle. \quad (6.16)$$

They are very important in quantum mechanics calculations. Even when we have general states like, $\langle \psi | \hat{A} | \phi \rangle$, we still call them **matrix elements**:

$$\underbrace{\langle \psi |}_{\text{bra}} \underbrace{\hat{A}}_{\text{operator}} \underbrace{|\phi\rangle}_{\text{ket}}. \quad (6.17)$$

You should show (Discussion 6.6) that all elements of an operator's matrix representation can be found as matrix elements. For example,

$$\hat{A} \doteq \mathbf{A} \doteq \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \langle + | \hat{A} | + \rangle & \langle + | \hat{A} | - \rangle \\ \langle - | \hat{A} | + \rangle & \langle - | \hat{A} | - \rangle \end{pmatrix}. \quad (6.18)$$

We can write a table to show the column and row structure³ of \hat{A} :

\hat{A}	$ +\rangle$	$ -\rangle$
$\langle + $	$\langle + \hat{A} + \rangle$	$\langle + \hat{A} - \rangle$
$\langle - $	$\langle - \hat{A} + \rangle$	$\langle - \hat{A} - \rangle$

or more compactly

$$\underbrace{\langle m |}_{\text{row}} \underbrace{\hat{A}}_{\text{operator}} \underbrace{|n\rangle}_{\text{column}}. \quad (6.19)$$

More generally, in a higher dimensional Hilbert space, operators will be represented by larger matrices

$$\hat{A} \doteq \mathbf{A} \doteq \begin{pmatrix} A_{11} & A_{12} & A_{13} & \dots \\ A_{21} & A_{22} & A_{23} & \dots \\ A_{31} & A_{32} & A_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \quad (6.20)$$

³ Analogously we have for a general ket, $|\psi\rangle$,

$$\begin{array}{c|c} |\psi\rangle & \\ \hline \langle + | & \langle + | \psi \rangle \\ \langle - | & \langle - | \psi \rangle \end{array}$$

and a general bra, $\langle \varphi |$,

$$\begin{array}{c|cc} \langle \varphi | & |+ \rangle & |- \rangle \\ \hline \langle \varphi | + \rangle & & \langle \varphi | - \rangle \end{array}$$

In this case the matrix elements are, $A_{ij} = \langle i | A | j \rangle$, where we have labelled our basis states as $|i\rangle$. In A_{ij} , the i subscript labels the rows and the j subscript labels the columns.

Discussion 6.7

Given a general operator \hat{A} in an n -dimensional Hilbert space with the representation, (6.20), supposing $|a_1\rangle \doteq \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$,

what is $\hat{A}|a_1\rangle$ in this representation?

Using this matrix representation, (6.20) , we can look at the action of an operator, \hat{A} , on a general ket $|\psi\rangle = \sum_{i=1}^n c_i |i\rangle$:

$$\begin{aligned} \hat{A}|\psi\rangle &\doteq \begin{pmatrix} A_{11} & A_{12} & A_{13} & \dots \\ A_{21} & A_{22} & A_{23} & \dots \\ A_{31} & A_{32} & A_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \\ & & & \vdots \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \\ &= \begin{pmatrix} A_{11}c_1 + A_{12}c_2 + A_{13}c_3 + \dots \\ A_{21}c_1 + A_{22}c_2 + A_{23}c_3 + \dots \\ A_{31}c_1 + A_{32}c_2 + A_{33}c_3 + \dots \\ \vdots \end{pmatrix}. \end{aligned} \quad (6.21)$$

Writing this new ket, $|\phi\rangle = \hat{A}|\psi\rangle$, in terms of the basis $|i\rangle$, as

$|\phi\rangle = \sum_{i=1}^n b_i |i\rangle$, we can write (6.21) more compactly as

$$b_i = \sum_{j=1}^n A_{ij}c_j . \quad (6.22)$$

Diagonalisation of operators

For \hat{S}_z , we used experimental results and the eigenvalue equation to find the matrix elements (in the S_z basis). It is common to work the other way round. In other words, given the matrix representation of an operator, we would like to be able to determine the possible results of a measurement. The 3rd postulate tells us that possible measurements are eigenvalues of the operator and the quantum states are the corresponding eigenvectors. This means that, given the matrix representation of an operator, we can find the possible results of a measurement of the corresponding observable by determining the eigenvalues of the matrix⁴.

The eigenvalue equation for a general operator is

$$\text{operator} \quad \widehat{\hat{A}} \quad \underbrace{|a_n\rangle}_{\text{eigenvector}} = \text{eigenvalue} \quad \widehat{a_n} \quad \underbrace{|a_n\rangle}_{\text{eigenvector}} . \quad (6.23)$$

⁴ On the other hand, when we have differential operators corresponding to an observable, finding the results of possible measurements boils down to solving a differential equation.

For a two state system, we can represent (6.23) in matrix notation as

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} c_{n_1} \\ c_{n_2} \end{pmatrix} = a_n \begin{pmatrix} c_{n_1} \\ c_{n_2} \end{pmatrix}, \quad (6.24)$$

where c_{n_1} and c_{n_2} are the unknown coefficients of the eigenvector, $|a_n\rangle$. For example, for a spin-½ system in the S_z representation we would have, $|a_n\rangle = c_{n_1}|+\rangle + c_{n_2}|-\rangle$. You should be able to show that (6.24) is equivalent to the two equations:

$$(A_{11} - a_n)c_{n_1} + A_{12}c_{n_2} = 0, \quad (6.25)$$

$$A_{21}c_{n_1} + (A_{22} - a_n)c_{n_2} = 0. \quad (6.26)$$

The rules of linear algebra imply that a set of homogeneous equations like (6.25) and (6.26), only has non-trivial solutions if the determinant of coefficients vanish:

$$\begin{vmatrix} A_{11} - \lambda & A_{12} \\ A_{21} & A_{22} - \lambda \end{vmatrix} = 0. \quad (6.27)$$

It is common to use λ for eigenvalues and we can write (6.27) as

$$\det(\mathbf{A} - \lambda \mathbf{1}) = 0, \quad (6.28)$$

where $\mathbf{1}$ is the 2×2 identity matrix $\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Discussion 6.8

Show that (6.27) can be more compactly written as (6.28).

Equation (6.28) is called the **secular** or **characteristic** equation. For 2×2 matrices the secular equation is a quadratic equation which implies that it has 2 possible roots. These roots are the eigenvalues a_1 and a_2 . Once we have the eigenvalues, we can substitute them back into (6.24) to find the coefficients of the eigenvector. Since our eigenvectors correspond to quantum states, we may also need to normalise them.

Example 6.5:

Problem Suppose we know that the matrix representation for \hat{S}_y (in the S_z basis) is

$$\mathbf{S}_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},$$

find the eigenvalues and normalised eigenvectors.

Solution The general eigenvalue equation is, $\hat{S}_y|\lambda\rangle = \lambda|\lambda\rangle$, for which we need to solve the secular equation $\det(\mathbf{S}_y - \lambda \mathbf{1}) = 0$. Writing out the secular equation explic-

itly gives us

$$\begin{vmatrix} -\lambda & -i\hbar/2 \\ i\hbar/2 & -\lambda \end{vmatrix} = 0$$

which implies

$$\begin{aligned} \lambda^2 - (-i\hbar/2)(i\hbar/2) &= 0 \\ \implies \lambda &= \pm\hbar/2 . \end{aligned} \quad (6.29)$$

This result is expected since we know that the result of measuring any of the spin components is $\pm\hbar/2$. So, the eigenvalue equation for the $\hbar/2$ eigenvalue is, $\hat{S}_y |+\rangle_y = +\hbar/2 |+\rangle_y$, which in matrix notation becomes

$$\underbrace{\frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}}_{\hat{S}_y} \underbrace{\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}}_{\doteq |+\rangle_y} = \frac{\hbar}{2} \underbrace{\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}}_{\doteq |+\rangle_y}, \quad (6.30)$$

$$\implies \frac{\hbar}{2} \begin{pmatrix} -ic_2 \\ ic_1 \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \quad (6.31)$$

which only gives us one independent equation

$$c_2 = ic_1 . \quad (6.32)$$

Now normalisation of our eigenvector requires

$$|c_1|^2 + |c_2|^2 = 1 \xrightarrow{(6.32)} |c_1|^2 + |ic_1|^2 = 1 \implies 2|c_1|^2 = 1 ,$$

so that

$$|c_1|^2 = \frac{1}{2} \implies c_1 = \frac{1}{\sqrt{2}} \xrightarrow{(6.32)} c_2 = \frac{i}{\sqrt{2}} ,$$

which means that

$$\begin{aligned} |+\rangle_y &\doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \\ &= \frac{1}{\sqrt{2}} |+ \rangle + \frac{i}{\sqrt{2}} |- \rangle . \end{aligned} \quad (6.33)$$

The procedure for finding eigenvalues and eigenvectors is called **diagonalisation**. Diagonalisation is a key step in many quantum mechanics problems.

Discussion 6.9

Using the matrix representation, $\mathbf{S}_y \doteq \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, show, from

the eigenvalue equation, $\hat{S}_y |- \rangle_y = -\frac{\hbar}{2} |- \rangle_y$, that $|- \rangle_y \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}$

The matrix representation of the spin operators and eigenvectors can be shown to be:

$$\begin{aligned}\mathbf{s}_x &\doteq \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & |+\rangle_x &\doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} & |-\rangle_x &\doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ \mathbf{s}_y &\doteq \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} & |+\rangle_y &\doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} & |-\rangle_y &\doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \\ \mathbf{s}_z &\doteq \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} & |+\rangle &\doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} & |-\rangle &\doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix}\end{aligned}$$

❖ Discussion 6.10

Write a summary of the important points and formulae in this chapter.

7

New operators

Spin Component in a General Direction

We will now use the operators \hat{S}_x , \hat{S}_y and \hat{S}_z to find the operator, \hat{S}_n , for the spin along some direction¹. The direction \hat{n} is determined by two angles θ and ϕ as shown in Figure 7.1.

The unit vector \hat{n} has components

$$\hat{n} = (\sin \theta \cos \phi) \hat{i} + (\sin \theta \sin \phi) \hat{j} + (\cos \theta) \hat{k} . \quad (7.1)$$

Discussion 7.1

Show how you can use Figure 7.1 to obtain (7.1).

We can find the component of the spin in the direction, \hat{n} , which we write as S_n , by taking the dot product of \vec{S} and \hat{n} , i.e. $S_n = \hat{n} \cdot \vec{S}$. A simple example of that would be to use \hat{i} to get $S_x = \hat{i} \cdot \vec{S}$. So, using (7.1);

$$\begin{aligned} S_n &= \vec{S} \cdot \hat{n} \\ &= S_x \sin \theta \cos \phi + S_y \sin \theta \sin \phi + S_z \cos \theta . \end{aligned} \quad (7.2)$$

which suggests that in terms of our operators

$$\begin{aligned} \hat{S}_n &= \hat{\vec{S}} \cdot \hat{n} \\ &= \hat{S}_x \sin \theta \cos \phi + \hat{S}_y \sin \theta \sin \phi + \hat{S}_z \cos \theta . \end{aligned} \quad (7.3)$$

Now we use (7.3) and the matrix representations \mathbf{S}_x , \mathbf{S}_y and \mathbf{S}_z to find \mathbf{S}_n . The result of this calculation is

$$\mathbf{S}_n \doteq \frac{\hbar}{2} \begin{pmatrix} \cos \theta & \sin \theta e^{-i\phi} \\ \sin \theta e^{i\phi} & -\cos \theta \end{pmatrix} . \quad (7.4)$$

Discussion 7.2

Show, starting from (7.3), that in the S_z -basis, \mathbf{S}_n is given by (7.4).

This reading is based on [McIntyre et al. 2012, Section 2.2]

¹ Here the “hat” on \hat{n} tells us that we have a unit vector and does not denote an operator \hat{n} . You may have seen the unit vector \hat{n} is also referred to as $\hat{e}_r = \hat{r} = \frac{\vec{r}}{r}$ where \vec{r} is the position vector extending from the origin to the point (x, y, z) .

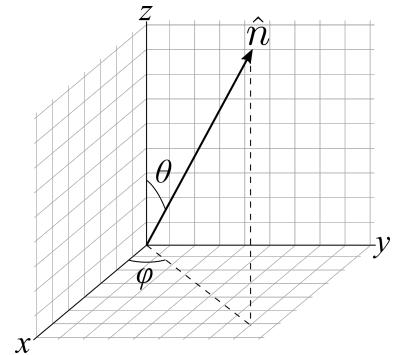


Figure 7.1: General direction along which to measure the spin component.

Modification of this image
Attribution: Andeggs, Public domain, via Wikimedia Commons

💬 Discussion 7.3

What values of θ and ϕ correspond to \mathbf{S}_x , \mathbf{S}_y and \mathbf{S}_z respectively? Substitute these values into (7.4) and check that you recover the expected matrices.

Just like the other spin components, \hat{S}_n has eigenvalues $\pm\hbar/2$.

💬 Discussion 7.4

Show that the matrix, \mathbf{S}_n , in (7.4), has eigenvalues $\pm\hbar/2$.

We can also use the matrix representation to find the corresponding eigenvectors of \hat{S}_n :

$$|+\rangle_n = \cos \frac{\theta}{2} |+ \rangle + \sin \frac{\theta}{2} e^{i\phi} |- \rangle \quad (7.5)$$

$$|-\rangle_n = \sin \frac{\theta}{2} |+ \rangle - \cos \frac{\theta}{2} e^{i\phi} |- \rangle . \quad (7.6)$$

Note that the convention that the first coefficient is real and positive has been enforced in (7.5) and (7.6). Any possible ket in a spin- $\frac{1}{2}$ system can be written in the form (7.5) (or (7.6)), for appropriate angles² with $0 \leq \theta < \pi$ and $0 \leq \phi < 2\pi$.

² Note that in higher dimensional Hilbert spaces we need more than two angles to specify a general state.

💬 Discussion 7.5

Start with a general ket, $|\psi\rangle = a|+\rangle + b|-\rangle$ with a and b complex. Explain how, using the convention that the first coefficient must be real and positive as well as requiring the state is normalised, means we can write this as $|\psi\rangle = |a||+\rangle + \sqrt{1 - |a|^2} e^{i\phi} |-\rangle$. How does this relate to (7.5)?

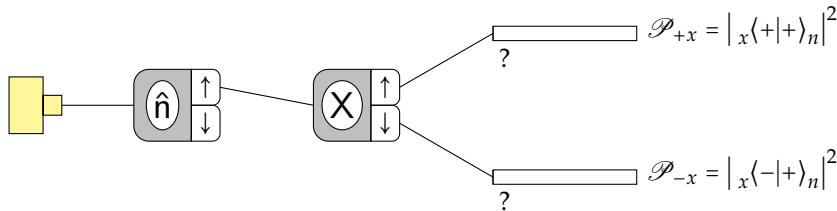


Figure 7.2: Measurement of the spin component after state preparation in a new direction.

📖 Example 7.1:

Problem Find the probabilities of the S_x measurements assuming that the first Stern-Gerlach analyser is aligned along the direction \hat{n} with $\theta = 2\pi/3$ and $\phi = \pi/4$ as shown in Figure 7.2.

Solution First we need to determine the state, $|\psi_{\text{in}}\rangle = |+\rangle_n$, that exits the upper port of the \hat{n} analyser and enters the

X-analyser. Using (7.5), we have

$$\begin{aligned} |\psi_{\text{in}}\rangle &= |+\rangle_n = \cos \frac{\theta}{2} |+\rangle + \sin \frac{\theta}{2} e^{i\phi} |-\rangle \\ &= \cos \frac{\pi}{3} |+\rangle + \sin \frac{\pi}{3} e^{i\pi/4} |-\rangle \\ &= \frac{1}{2} |+\rangle + \frac{\sqrt{3}}{2} e^{i\pi/4} |-\rangle \end{aligned} \quad (7.7)$$

$$\begin{aligned} \mathcal{P}_{+x} &= |_x \langle +|+\rangle_n|^2 \\ &= \left| \left(\frac{1}{\sqrt{2}} \langle +| + \frac{1}{\sqrt{2}} \langle -| \right) \left(\frac{1}{2} |+\rangle + \frac{\sqrt{3}}{2} e^{i\pi/4} |-\rangle \right) \right|^2 \\ &= \left| \frac{1}{2\sqrt{2}} + \frac{\sqrt{3}}{2\sqrt{2}} e^{i\pi/4} \right|^2 \\ &= \frac{1}{8} (1 + \sqrt{3} e^{i\pi/4}) (1 + \sqrt{3} e^{-i\pi/4}) \\ &= \frac{1}{8} (1 + 3 + \sqrt{3} (e^{i\pi/4} + e^{-i\pi/4})) \\ &= \frac{1}{8} (4 + 2\sqrt{3} \cos(\pi/4)) = \frac{1}{2} + \frac{\sqrt{3}}{4\sqrt{2}} \\ &\approx 0.806 . \end{aligned} \quad (7.8)$$

Now for \mathcal{P}_{-x} , remembering that $|-\rangle_x = \frac{1}{\sqrt{2}}(|+\rangle - |-\rangle)$ and using matrix notation for a change,

$$\begin{aligned} \mathcal{P}_{-x} &= |_x \langle -|+\rangle_n|^2 \\ &= \left| \frac{1}{\sqrt{2}} (1 \quad -1) \frac{1}{2} \begin{pmatrix} 1 \\ \sqrt{3} e^{i\pi/4} \end{pmatrix} \right|^2 \\ &= \left| \frac{1}{2\sqrt{2}} - \frac{\sqrt{3}}{2\sqrt{2}} e^{i\pi/4} \right|^2 \\ &= \frac{1}{8} (1 - \sqrt{3} e^{i\pi/4}) (1 - \sqrt{3} e^{-i\pi/4}) \\ &= \frac{1}{8} (1 + 3 - \sqrt{3} (e^{i\pi/4} + e^{-i\pi/4})) \\ &= \frac{1}{8} (4 - 2\sqrt{3} \cos(\pi/4)) = \frac{1}{2} - \frac{\sqrt{3}}{4\sqrt{2}} \\ y &\approx 0.194 . \end{aligned} \quad (7.9)$$

We can check that, $\mathcal{P}_{+x} + \mathcal{P}_{-x} = 1$, as you should have expected (why?). A histogram of these results is shown in Figure 7.3.

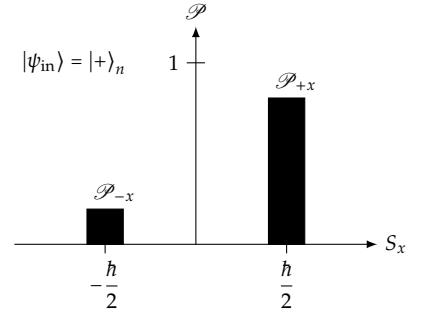


Figure 7.3: Histogram of spin component S_x measurement.

Projection Operators

So far we have looked at operators corresponding to components of the spin vector. In this section we'll look at a new type of operator. Consider a ket written in the S_z basis

$$|\psi\rangle = a|+\rangle + b|-\rangle = \underbrace{(\langle +|\psi\rangle)}_a |+\rangle + \underbrace{(\langle -|\psi\rangle)}_b |-\rangle \quad (7.10)$$

The 1st term of (7.10) can be written as a number times a ket or as a ket times a number:

$$(\langle +|\psi\rangle)|+\rangle = |+\rangle(\langle +|\psi\rangle) . \quad (7.11)$$

Separating out the inner product, the 2nd term in (7.11) can be written as

$$|+\rangle(\langle +|\psi\rangle) = (|+\rangle\langle +|)|\psi\rangle \quad (7.12)$$

The object in brackets on the right hand side of (7.12) is the product of a bra and a ket³. The new object, $|+\rangle\langle +|$, acts on a ket $|\psi\rangle$ and produces another ket $(\langle +|\psi\rangle)|+\rangle$. This means it is an operator. Some people call this a ket-bra and it is also known as an **outer product**.

We can write $|\psi\rangle$ in terms of ket-bras:

$$\begin{aligned} |\psi\rangle &= (\langle +|\psi\rangle)|+ + (\langle -|\psi\rangle)|- \\ &= |+\rangle\langle +|\psi\rangle + |-\rangle\langle -|\psi\rangle \\ &= (|+\rangle\langle +| + |-\rangle\langle -|)|\psi\rangle. \end{aligned} \quad (7.13)$$

Once again we see that the object in brackets on the right hand side of (7.13) is an operator. If you look closely at (7.13) you should notice that, $|+\rangle\langle +| + |-\rangle\langle -|$, is a very special operator – it gives us back our original ket – this means that it is the **identity operator**. We shall use $\hat{\mathbf{1}}$ to denote the identity operator⁴ so that in summary we have:

$$|+\rangle\langle +| + |-\rangle\langle -| = \hat{\mathbf{1}} \quad (7.14)$$

Equation (7.14) is known as the **completeness relation** or **closure**.

In matrix notation we see that:

$$\begin{aligned} |+\rangle\langle +| + |-\rangle\langle -| &\doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} \\ &\doteq \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ &\doteq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{1}. \end{aligned} \quad (7.15)$$

Now, we'll look at $|+\rangle\langle +|$ and $|-\rangle\langle -|$ which are called **projection operators**. In general, projection operators give us the component of a state in a particular “direction”⁵. In the S_z -basis, \hat{P}_{\pm} have the form

$$\begin{aligned} \hat{P}_+ &= |+\rangle\langle +| \doteq \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \\ \hat{P}_- &= |-\rangle\langle -| \doteq \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \quad (7.16)$$

³ With the order reversed if we compare it to an inner product.

⁴ The identity operator has the property that, for any ket, $|\psi\rangle$:

$$\hat{\mathbf{1}}|\psi\rangle = |\psi\rangle.$$

Note that,

$$\hat{\mathbf{1}} \doteq \mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

⁵ Mathematically projection operators are defined as operators which satisfy, $\hat{P}^2 = \hat{P}$. In other words, once you have projected a state, projecting again gives you the same state.

Discussion 7.6

Using (7.16), show that $(\mathbf{P}_+)^2 = \mathbf{P}_+$ and $(\mathbf{P}_-)^2 = \mathbf{P}_-$. Can you show directly that $\hat{P}_{\pm}^2 = \hat{P}_{\pm}$?

Discussion 7.7

Find the matrix representation, in the z -representation, of the projection operator, $\hat{P}_{+x} = |+\rangle_x \langle +|$. You may find it useful to recall that in z -representation $|+\rangle_x = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Check that $(\mathbf{P}_{+x})^2 = \mathbf{P}_{+x}$.

We can use these operators to write the completeness relation

$$\hat{P}_+ + \hat{P}_- = \hat{1}. \quad (7.17)$$

The projection operators, \hat{P}_{\pm} , produce new kets that align along the particular + or - eigenstate respectively. The amplitude of the new ket is the amplitude for the original ket, $|\psi\rangle$, to be in that eigenstate:

$$\begin{aligned} \hat{P}_+ |\psi\rangle &= |+\rangle \langle +|\psi\rangle = (\langle +|\psi\rangle) |+\\ \hat{P}_- |\psi\rangle &= |-\rangle \langle -|\psi\rangle = (\langle -|\psi\rangle) |- . \end{aligned} \quad (7.18)$$

In particular, a projector acting on its corresponding eigenstate gives us that eigenstate back, for example,

$$\hat{P}_+ |+\rangle = |+\rangle \langle +|+ = |+\rangle .$$

On the other hand a projector acting on a orthogonal state, gives us zero, for example,

$$\hat{P}_+ |-\rangle = |+\rangle \langle +|- = 0 .$$

We can use projection operators to rewrite probabilities. For example.

$$\begin{aligned} \mathcal{P}_+ &= |\langle +|\psi\rangle|^2 \\ &= \langle +|\psi\rangle^* \langle +|\psi\rangle \\ &= \langle \psi|+ \rangle \langle +|\psi\rangle \\ &= \langle \psi|\hat{P}_+|\psi\rangle \end{aligned} \quad (7.19)$$

Expressing (7.19) in words, we can say that the probability of measuring $S_z = +\hbar/2$, can be found by calculating a matrix element of the projection operator. The relevant matrix element involves the input state $|\psi\rangle$ and the projection operator \hat{P}_+ corresponding to the result.

In chapter 1 we saw how measurement disturbs a system. For example if a state $|\psi\rangle$ is measured to have $S_z = +\hbar/2$, then its state is changed to $|+\rangle$ by the measurement itself. Projection operators give us a way of expressing this mathematically. In (7.18) we see that when we feed a general state to \hat{P}_+ , we get $|+\rangle$ times the probability amplitude, $\langle +|\psi\rangle$. This means, to get just $|+\rangle$, we also need to divide out the amplitude. Taking the square root of (7.19), gives us what we need⁶ $\langle +|\psi\rangle = \sqrt{\langle \psi|\hat{P}_+|\psi\rangle}$. Writing the above paragraph in terms of formulae

$$\hat{P}_+ |\psi\rangle \stackrel{(7.18)}{=} (\langle +|\psi\rangle) |+ \rangle$$

⁶ Strictly speaking, there could be a phase, $\langle +|\psi\rangle = |\langle +|\psi\rangle| e^{i\phi}$, that we are not worrying about here. Since this will end-up being an overall phase it is of no physical significance – see Discussion 7.8.

$$\begin{aligned} \implies |+\rangle &= \frac{\hat{P}_+ |\psi\rangle}{\langle +|\psi\rangle} \\ &= \frac{\hat{P}_+ |\psi\rangle}{\sqrt{\mathcal{P}_+}} \\ &\stackrel{(7.19)}{=} \frac{\hat{P}_+ |\psi\rangle}{\sqrt{\langle \psi| \hat{P}_+ |\psi\rangle}}. \end{aligned} \quad (7.20)$$

More generally, we can express the effect of making a measurement using the relevant projection operator corresponding to the quantum state we observe. This is summarised in the fifth postulate, which is also called the **projection postulate**, below

QM Postulate 5

After a measurement of A that yields the result a_n , the quantum system is in a new state that is the normalized projection of the original system ket onto the ket (or kets) corresponding to the result of the measurement:

$$|\psi_{\text{new}}\rangle = \frac{\hat{P}_n |\psi\rangle}{\sqrt{\langle \psi| \hat{P}_n |\psi\rangle}} \quad (7.21)$$

The projection postulate is central to quantum mechanics. The effect of measurement is also called the **collapse** of the quantum state vector. The postulate makes it clear that *measurement disturbs a system*. The one exception to this is if the initial state is the same as the measured state. The possibility of collapse makes quantum mechanics, unlike classical mechanics, *fundamentally irreversible*.

Discussion 7.8

The denominator in eq. (7.21) may leave behind an irrelevant overall phase. For example, consider the state $|\psi\rangle = \cos(\theta/2)|+\rangle + \sin(\theta/2)e^{i\phi}|-\rangle$ with $\hat{P}_- = |-\rangle\langle -|$.

Show that

$$|\psi_{\text{new}}\rangle = \frac{\hat{P}_- |\psi\rangle}{\sqrt{\langle \psi| \hat{P}_- |\psi\rangle}} = e^{i\phi} |-\rangle$$

Explain why this phase is irrelevant.

8

New operators (continued...)

We will now use the new language of projectors to model quantum mechanical measurement as shown in Figure 8.1. In this model, each port of the Stern-Gerlach analyser is modelled as a projection operator.

This reading is based on [McIntyre et al. 2012, Section 2.2]

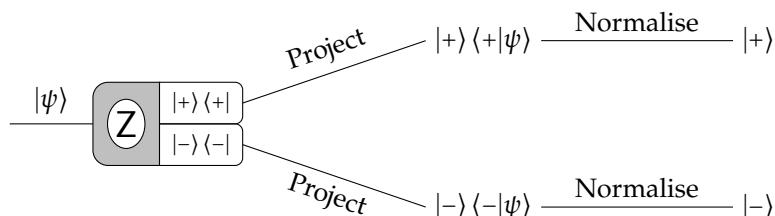


Figure 8.1: Schematic diagram of the role of the projection operator in a Stern-Gerlach spin measurement.

Looking at the upper port in Figure 8.1, the input state, $|\psi_{\text{in}}\rangle$, is acted on by the projection operator $\hat{P}_+ = |+\rangle\langle +|$. This leads to the output state, $|\psi_{\text{projected}}\rangle = |+\rangle(\langle +|\psi_{\text{in}}\rangle)$. The probability of this happening is given by $\mathcal{P}_+ = |\langle +|\psi_{\text{in}}\rangle|^2$. Notice that the output ket is not normalised, so the next step in our model is to normalise the state. We find that

$$|\psi_{\text{out}}\rangle = \frac{|\psi_{\text{projected}}\rangle}{\sqrt{\mathcal{P}_+}} = |+\rangle . \quad (8.1)$$

Discussion 8.1

Show why, to normalise $|\psi_{\text{projected}}\rangle$ in (8.1), we need to divide by $\sqrt{\mathcal{P}_+}$.

QM Postulate 5 tells us how to determine the state of a quantum system after a measurement, **but** it does not really tell us what is happening with the measurement process. Trying to understand measurement in quantum mechanics is an ongoing research problem. Nonetheless, despite having to treat measurement a bit like a “black box” as in Figure (8.1), quantum theory very accurately predicts experimental results.

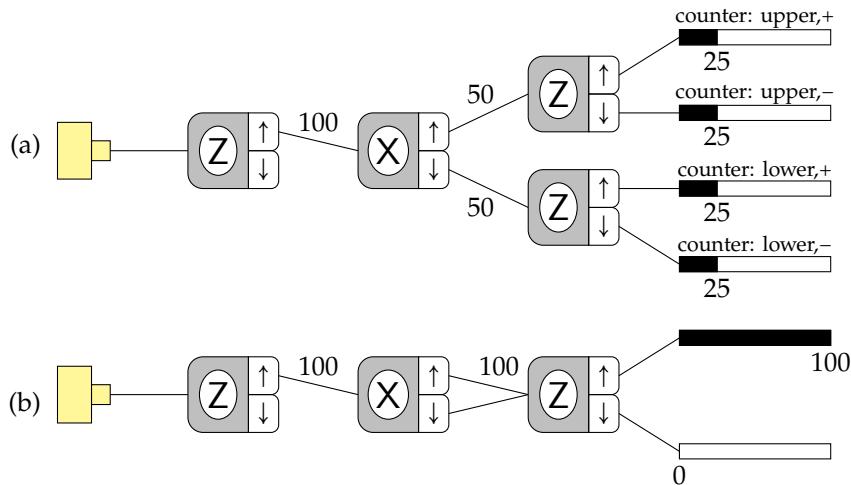


Figure 8.2: (a) Hybrid Experiment 4a and 4b, and (b) Experiment 4c.

Returning to Experiment 3 & 4

Armed with new tools, we now return to analyse Experiments 3 & 4 from chapter 2, as shown in Figure 8.2. Firstly you may remember that Experiment 3 is the same as Experiment 4a. Secondly in Figure 8.2, the hybrid experiment in part (a) combines Experiment 4a and Experiment 4b. Experiment 4a corresponds to atoms leaving upper port of the X-analyser while Experiment 4b corresponds to atoms leaving the lower port of the X-analyser. Finally in Figure 8.2, part (b) corresponds to Experiment 4(c).

In all these experiments, atoms are prepared in the $|+\rangle$ state by the first analyser. We would now like to look at the probability of atoms ending up in each of the counters at the end of the experiments. This involves looking at the measurements at the two intermediate analysers using our new tools.

Let's start by looking at the probability that atoms reaches top counter in Figure 8.2(a) (*i.e.* counter: upper, +). We first need to determine the probability that an atom leaves upper port of 2nd analyser. Since the atom left the upper port of the first Z-analyser, the input state for the second analyser is $|+\rangle$. To reach the upper analyser, the atom must leave the upper port of the second X-analyser which means the output state is $|+\rangle_x$. So, the probability it reaches the upper analyser is, $\mathcal{P}_{\text{upper}} = |\langle \text{out} | \text{in} \rangle|^2 = |x\langle +|+\rangle|^2$. Now we need to find the probability that an atom leaves the upper port of the third upper Z-analyser. Having left the second X-analyser the input state is $|+\rangle_x$. If it is leaving via the upper port of the third upper Z-analyser the output state is $|+\rangle$. So, the probability it exits the upper port is, $\mathcal{P} = |\langle \text{out} | \text{in} \rangle|^2 = |(+|+)\rangle_x|^2$. Finally the total probability will be the product of the two probabilities we have just found:

$$\begin{aligned}\mathcal{P}_{\text{upper},+} &= (\mathcal{P}_{\text{leaving the } \uparrow \text{ port of upper Z-analyser}})(\mathcal{P}_{\text{leaving the } \uparrow \text{ port of X-analyser}}) \\ &= |(+|+)\rangle_x|^2 |x\langle +|+\rangle|^2.\end{aligned}\tag{8.2}$$

Notation

As in (8.2) will read the probabilities from right to left.

In a similar way we can find the probability for atom to reach the next counter down (*i.e.* counter: upper, $-$) or in other words the probability of first measuring, $S_x = \hbar/2$ again, followed by a measurement of, $S_z = -\hbar/2$. The only difference with the previous case, (8.2), is that the output state from the final analyser is $|-\rangle$ rather than $|+\rangle$ so that

$$\begin{aligned}\mathcal{P}_{\text{upper},-} &= (\mathcal{P}_{\text{leaving the } \downarrow \text{ port of upper Z-analyser}})(\mathcal{P}_{\text{leaving the } \uparrow \text{ port of X-analyser}}) \\ &= |\langle -|+\rangle_x|^2 |\langle +|+\rangle_x|^2.\end{aligned}\quad (8.3)$$

Once again the calculation for the lower counters is much the same except for the fact that the intermediate state is $|-\rangle_x$ rather than $|+\rangle_x$ leading to.

$$\mathcal{P}_{\text{lower},+} = |\langle +|- \rangle_x|^2 |\langle -|+\rangle_x|^2 \quad (8.4)$$

$$\mathcal{P}_{\text{lower},-} = |\langle -|- \rangle_x|^2 |\langle -|+\rangle_x|^2 \quad (8.5)$$

Now, in Experiment 4(c) shown in Figure 8.2 (b), something different happens at the 2nd analyser. In this case both beams are sent to the 3rd analyser. This means that there is a 100% chance atoms reach the final Z-analyser and we only need to look at probabilities for the 3rd analyser. To calculate probabilities at the 3rd analyser, we first need to find the input state. We can find the input state using the projection postulate. Since both output ports feed into 3rd analyser, the projection operator we need is the sum of projection operators for each port. In other words we need, $\hat{P}_{+x} + \hat{P}_{-x}$, where $\hat{P}_{+x} = |+\rangle_x \langle +|$ and $\hat{P}_{-x} = |-\rangle_x \langle -|$. So, using the projection postulate, after leaving the second analyser, we have the state

$$|\psi_2\rangle = \frac{(\hat{P}_{+x} + \hat{P}_{-x})|\psi_1\rangle}{\sqrt{\langle\psi_1|(\hat{P}_{+x} + \hat{P}_{-x})|\psi_1\rangle}} \quad (8.6)$$

$$= \frac{(\hat{P}_{+x} + \hat{P}_{-x})|+\rangle}{\sqrt{\langle +|(\hat{P}_{+x} + \hat{P}_{-x})|+\rangle}}. \quad (8.7)$$

In this case, the projector, $\hat{P}_{+x} + \hat{P}_{-x}$, is just the identity operator since the two states constitute a complete basis. This makes the calculation rather easy and we find

$$|\psi_2\rangle = \frac{\hat{1}|+\rangle}{\sqrt{\langle +|\hat{1}|+\rangle}} = |+\rangle. \quad (8.8)$$

Rather than jumping ahead to the final answer, we can learn something about Experiment 4(c) by looking at the numerator in detail

and simplifying the denominator:

$$\begin{aligned}
 |\psi_2\rangle &= \frac{(\hat{P}_{+x} + \hat{P}_{-x})|+\rangle}{\sqrt{\langle+|+}\rangle} \\
 &= (|+\rangle_x \langle+| + |-\rangle_x \langle-|)|+\rangle \\
 &= |+\rangle_x (\langle+x|+)\rangle + |-\rangle_x (\langle+x|-\rangle) \\
 &= (\langle+x|+)\rangle|+\rangle_x + (\langle+x|-\rangle)|-\rangle_x
 \end{aligned} \tag{8.9}$$

From this expression we see that $|\psi_2\rangle$ is a coherent superposition of eigenstates of the 2nd analyser, namely $|+\rangle_x$ and $|-\rangle_x$. Using the form of $|\psi_2\rangle$ in (8.9) we can calculate probabilities of measurements at 3rd analyser. The probability of measuring spin up is

$$\begin{aligned}
 \mathcal{P}_+ &= |\langle+|\psi_2\rangle|^2 \\
 &= |\langle+|+\rangle_x \langle+x|+ + \langle+|-\rangle_x \langle+x|-\rangle|^2,
 \end{aligned} \tag{8.10}$$

and the probability of measuring spin down is

$$\begin{aligned}
 \mathcal{P}_- &= |\langle-|\psi_2\rangle|^2 \\
 &= |\langle-|+\rangle_x \langle+x|+ + \langle-|-\rangle_x \langle+x|-\rangle|^2.
 \end{aligned} \tag{8.11}$$

Notice that both probabilities, \mathcal{P}_\pm , are the square of a sum of probability amplitudes. Each of these amplitudes are in turn the product of the amplitudes of two particular measurements happening one after another. For example in \mathcal{P}_- , reading from right to left, the amplitude $\langle-|+\rangle_x \langle+x|+$, comes from an initial state $|+\rangle$ which is first measured to be in the $|+\rangle_x$ state and then measured to be in the $|-\rangle$ state. The path corresponding to this amplitude is shown in Figure 8.3. The other path contributing to \mathcal{P}_- is shown in Figure 8.4.

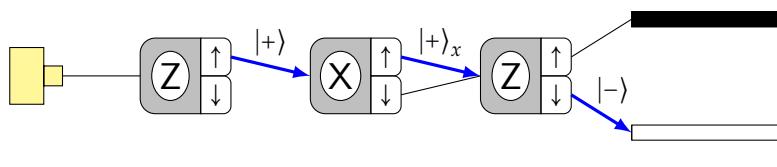


Figure 8.3: Path corresponding to the amplitude $\langle-|+\rangle_x \langle+x|+$ in experiment 4(c). This is the path “upper, –” ie. the particle leaves the upper port of analyser 2 and then is measured to be in the $|+\rangle$ state.

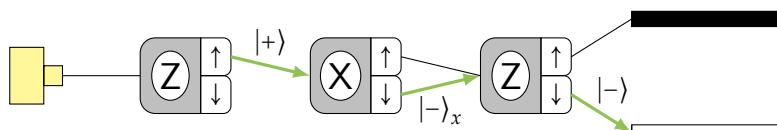


Figure 8.4: Path corresponding to the amplitude $\langle-|-\rangle_x \langle+x|-$ in experiment 4(c). This is the path “lower, –” ie. the particle leaves the lower port of analyser 2 and then is measured to be in the $|-\rangle$ state.

We add these two amplitudes together to get the total amplitude for the measurement of $S_z = -\hbar/2$. Having found the amplitude we then have to square it to find the probability. This leads to four

terms as follows:

$$\begin{aligned}
 \mathcal{P}_- &= |\langle -|+\rangle_x \langle +|+ \rangle + \langle -|- \rangle_x \langle -|- \rangle|^2 \\
 &= |\langle -|+\rangle_x \langle +|+ \rangle|^2 + |\langle -|- \rangle_x \langle -|- \rangle|^2 \\
 &\quad + \langle -|+\rangle_x \langle +|+ \rangle (\langle -|- \rangle_x)^* (\langle -|- \rangle_x)^* \\
 &\quad + (\langle -|+\rangle_x)^* (\langle +|+ \rangle)^* \langle -|- \rangle_x \langle -|- \rangle \\
 &= \mathcal{P}_{\text{upper},-} + \mathcal{P}_{\text{lower},-} + \text{interference terms}
 \end{aligned} \tag{8.12}$$

From the last line of (8.12), we see that the probability of finding an atom in the spin down state is equal to the probability that the particle followed upper path plus the probability the particle followed lower path *plus* additional terms. These additional terms are called **interference terms**. Unlike the first two, the interference terms are not complex squares which means they may be positive or negative. This means that its possible for the interference terms to be negative and cancel the other terms which is precisely what happens in this case.

This should remind you of the double slit experiment – at certain places, the amplitudes for different paths the electromagnetic wave follows cancel leading to a dark spot on the screen. This is an example of destructive interference.

The interference in Experiment 4c is due to the nature of the superposition state that enters the third analyser. To understand why this is the case let's look at what happens if we had used a mixed state leaving the second analyser. In a superposition state every atom in the beam is in the same state. In a mixed state, there is a statistical mixture of atoms in the beam in different states.

As we mentioned, Experiment 4c has superposition state entering 3rd analyser. We can change this state by watching to see which output port of the 2nd analyser a particular atom goes through. One could do this by, for example shining a light on the atoms, the scattering of the light would tell us which port a particular atom left from. Such an experiment is called a “Which Path” or “Welcher Weg” experiment.

If we know which path an atom takes, the state is not in the superposition state $|\psi_2\rangle$ since we have measured it to be either in the $|+\rangle_x$ state or the $|-\rangle_x$ state. This means that the chance an atom is detected as spin down at last detector is just the sum of the probabilities for the two possible paths. In other words, we add independent probabilities for the path $|+\rangle \rightarrow |+\rangle_x \rightarrow |- \rangle$ and the path $|+\rangle \rightarrow |-\rangle_x \rightarrow |- \rangle$:

$$\begin{aligned}
 \mathcal{P}_{\text{watch},-} &= |\langle -|+\rangle_x \langle +|+ \rangle|^2 + |\langle -|- \rangle_x \langle -|- \rangle|^2 \\
 &= \mathcal{P}_{\text{upper},-} + \mathcal{P}_{\text{lower},-}
 \end{aligned} \tag{8.13}$$

The important thing to notice about (8.13) is that there are **no** interference terms. This example shows the important difference between a coherent superposition and statistical mixed state. In a coherent superposition there is a definite relative phase between

different states which means that there can be phase dependent interference effects. In a mixed state the phase relationship is destroyed and interference effects are washed out. Incidentally, this also tells us what it means to combine beams “properly” – we have to preserve the relative phase. Randomising the phase destroys the superposition leading to a statistical mixture. If the beams are properly combined so that superposition is unaffected, the results of Experiment 4(c) are same as if no measurement had happened. Even though we have a measuring device in the middle of Experiment 4(c), effectively no measurement has been made.

Summary

- If no measurement is made on intermediate states, we first add the amplitudes and then square to find the probability. In this case you can have interference.
- If a measurement is made, we first square the amplitudes and then add to find the total probability. In this case there isn’t interference.

Discussion 8.2

Write a summary of the important points and formulae in this chapter.

9

Hermitian Operators

We will now look at a special class of operators called **hermitian operators** which play a *very* important role in quantum mechanics. Before we look at these operators we need to understand **hermitian conjugation**.

This reading is based on [McIntyre et al. 2012, Section 2.2]

Hermitian Conjugation

In our matrix representation, the formula $\hat{A}|\psi\rangle = |\phi\rangle$, corresponds to a matrix multiplying a column vector from the left. We say that the operator acts on the ket from the left to produce a new ket. On the other hand if the operator is on the right of our ket, as in $|\psi\rangle \hat{A}$, we get something which is not defined in our matrix representation (see Discussion 9.1).

Discussion 9.1

Show that, $|\psi\rangle \hat{A}$, does not make sense given the matrix representation, $|\psi\rangle = \begin{pmatrix} a \\ b \end{pmatrix}$, and, $\hat{A} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$.

Conversely, if we want to act with an operator on a bra, the operator must be on the right, as in $\langle\psi| \hat{A}$ (see Discussion point 9.2):

Discussion 9.2

Show that, $\langle\psi| \hat{A}$, **does** make sense given the matrix representation, $\langle\psi| = \begin{pmatrix} a^* & b^* \end{pmatrix}$, and, $\hat{A} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, whereas, $\hat{A}\langle\psi|$, is **not** defined using the matrix representation.

Notation

We define the hermitian conjugate, \dagger , which takes us from a ket, $|\psi\rangle$, to its associated bra, $\langle\psi|$, as follows

$$(|\psi\rangle)^\dagger = \langle\psi|. \quad (9.1)$$

Conversely hermitian conjugation can also take us from a

bra to a ket:

$$(\langle \psi |)^\dagger = |\psi\rangle . \quad (9.2)$$

💬 Discussion 9.3

Explain why $(|\psi\rangle)^{\dagger\dagger} = |\psi\rangle$.

💬 Discussion 9.4

Show that in the matrix representation, given ,

$$|\psi\rangle \doteq \begin{pmatrix} a \\ b \end{pmatrix},$$

then

$$(|\psi\rangle)^\dagger = \langle \psi | \doteq \left(\begin{pmatrix} a \\ b \end{pmatrix}^T \right)^* = \left(\begin{pmatrix} a \\ b \end{pmatrix}^* \right)^T$$

where T corresponds to taking the transpose. What is the matrix representation of eq. (9.2)?

From Discussion 9.4 we see that, in the matrix representation, hermitian conjugation corresponds to taking the transpose and complex conjugate, i.e. $\dagger = T^* = *T$.

💬 Discussion 9.5

Given the representation, $|\psi\rangle \doteq \begin{pmatrix} a \\ b \end{pmatrix}$, and $\hat{A} \doteq \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, if, $|\phi\rangle = \hat{A} |\psi\rangle$, show explicitly using the matrix representation, that, $\langle \phi | = (\hat{A} |\psi\rangle)^\dagger$, is not necessarily the same as $\langle \psi | \hat{A}$. What goes wrong?

Following on from Discussion 9.5 we define the **hermitian adjoint**, \hat{A}^\dagger , of an operator \hat{A} which satisfies

$$(\hat{A} |\psi\rangle)^\dagger = \langle \psi | \hat{A}^\dagger , \quad (9.3)$$

for all states $|\psi\rangle$.

💬 Discussion 9.6

Using eq. (9.3) and the property that if $|\alpha\rangle = b |\beta\rangle + c |\gamma\rangle$, then $\langle \alpha | = b^* \langle \beta | + c^* \langle \gamma |$, show that

$$\hat{A} = b \hat{B} + c \hat{C} \implies \hat{A}^\dagger = b^* \hat{B}^\dagger + c^* \hat{C}^\dagger , \quad (9.4)$$

where $a, b \in \mathbb{C}$.

 Discussion 9.7

Revisiting your calculations from Discussion 9.5, show that in the matrix representation, if $\hat{A} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ then $\hat{A}^\dagger = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{T*}$, or in other words

$$\mathbf{A}^\dagger = \mathbf{A}^{T*}. \quad (9.5)$$

 Example 9.1:

Problem Find \mathbf{M}^\dagger if $\mathbf{M} = \begin{pmatrix} 3 & i \\ 1 & 2i \end{pmatrix}$.

Solution

$$\begin{aligned} \mathbf{M}^\dagger &= (\mathbf{M}^T)^* \\ &= \left(\begin{pmatrix} 3 & i \\ 1 & 2i \end{pmatrix}^T \right)^* \\ &= \begin{pmatrix} 3 & 1 \\ i & 2i \end{pmatrix}^* \\ &= \begin{pmatrix} 3 & 1 \\ -i & -2i \end{pmatrix} \end{aligned}$$

More generally, consider the inner product between the ket, $|\phi\rangle = \hat{A}|\alpha\rangle$, and some arbitrary state, $|\beta\rangle$. Using the properties of the inner product, we know that

$$\begin{aligned} \langle \alpha | \beta \rangle &= (\langle \beta | \alpha \rangle)^* \\ \stackrel{(9.3)}{\implies} [\langle \alpha | \hat{A}^\dagger] |\beta\rangle &= (\langle \beta | [\hat{A} | \alpha \rangle])^* \\ \implies \langle \alpha | \hat{A}^\dagger | \beta \rangle &= (\langle \beta | \hat{A} | \alpha \rangle)^*. \end{aligned} \quad (9.6)$$

 Discussion 9.8

Explain why eq. (9.6) is a more general form of eq. (9.5).

 Discussion 9.9

Explain how we can conclude from eq. (9.6) that

$$\hat{A}^{\dagger\dagger} = \hat{A}. \quad (9.7)$$

Notation

When an operator \hat{A} acts on a state $|\psi\rangle$ we get a new state $\hat{A}|\psi\rangle$. We will sometimes write this new state as $|\hat{A}\psi\rangle$:

$$|\hat{A}\psi\rangle := \hat{A}|\psi\rangle , \quad (9.8)$$

where “ $:=$ ” is short hand for “is defined as”. We also write

$$(|\hat{A}\psi\rangle)^\dagger := \langle \hat{A}\psi| . \quad (9.9)$$

Discussion 9.10

Explain why

$$\langle \hat{A}\psi | = \langle \psi | \hat{A}^\dagger \quad (9.10)$$

Discussion 9.11

Show that

$$\langle \alpha | \hat{A} | \beta \rangle = (\langle \beta | \hat{A}^\dagger | \alpha \rangle)^*, \quad (9.11)$$

for arbitrary states $|\alpha\rangle$ and $|\beta\rangle$.

Discussion 9.12

Using eqs. (9.8) and (9.11) show that

$$\langle \alpha | \hat{A} \beta \rangle = \langle A^\dagger \alpha | \beta \rangle , \quad (9.12)$$

for arbitrary states $|\alpha\rangle$ and $|\beta\rangle$.

Discussion 9.13

Show that

$$(\hat{A}\hat{B})^\dagger = B^\dagger A^\dagger \quad (9.13)$$

Hint: make use of eq. (9.12).

Hermitian Operators

We say that an operator is **hermitian** if it satisfies

Hermitian operators

$$\hat{A} = \hat{A}^\dagger \quad (9.14)$$

 Discussion 9.14

Show that for hermitian operators

$$\langle \alpha | \hat{A} \beta \rangle = \langle \hat{A} \alpha | \beta \rangle . \quad (9.15)$$

Hermitian operators will be very important for us, since, it turns out that, in quantum mechanics all operators corresponding to observables are **Hermitian**. The fact that observables correspond to hermitian operators is related to two important mathematical properties:

Important properties of hermitian operators

1. Hermitian matrices have real eigenvalues, which ensures that results of measurements are always real
2. The eigenvectors of a Hermitian matrix comprise a complete set of orthogonal basis states. This means that we can use eigenvectors of any observable as a valid basis.

 Discussion 9.15

1. Show that \mathbf{S}_y is hermitian.
2. Is $\mathbf{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ hermitian? What are the eigenvalues of \mathbf{J} ?
3. Show explicitly that $(\mathbf{S}_y \mathbf{J})^\dagger = \mathbf{J}^\dagger \mathbf{S}_y^\dagger$

 Discussion 9.16

Consider $\mathbf{U} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$. Clearly \mathbf{U} is not Hermitian (why?).

Show that \mathbf{U} has real eigenvalues but that the eigenvectors are not orthogonal.

10

Measurement, Observables and Uncertainty

10.1 Measurement

This reading is based on [McIntyre et al. 2012, Section 2.3-2.5]

We have seen that quantum mechanics involves probabilities. To study the probabilities experimentally we need to perform repeated experiments on identically prepared systems. To see how this can be done consider the schematic Stern-Gerlach like experiments for spin- $\frac{1}{2}$ particles shown in Figure 10.1. The atoms are prepared in some state $|\psi\rangle$ with the first analyser. The second analyser can then perform measurements on atoms which have all been prepared in the $|\psi\rangle$ state.

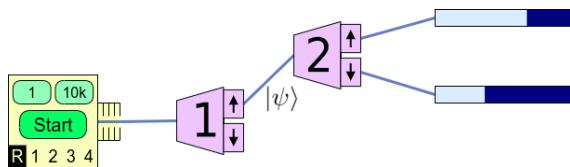


Figure 10.1: General Stern-Gerlach experiment

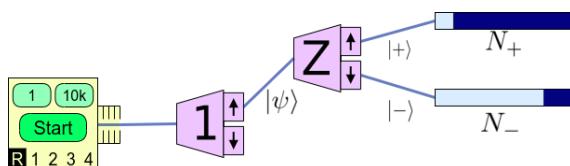


Figure 10.2: Stern-Gerlach experiment measuring S_z

For example, suppose we measure S_z for N identically prepared atoms as shown in Figure 10.2. In Figure 10.2, N_+ is the number of times we measure $S_z = \hbar/2$ and N_- = number of times we measure $S_z = -\hbar/2$. The total number of measurements is $N = N_+ + N_-$. Now, suppose we measure S_z for N identically prepared atoms. The 4th postulate of quantum mechanics implies that the probability to measure $S_z = \hbar/2$ is

$$\mathcal{P}_+ = |\langle + | \psi \rangle|^2 \quad (10.1)$$

For a large number of measurements we expect $\mathcal{P}_+ \approx N_+/N$. The two quantities are unlikely to be exactly the same due to statistical fluctuations, although as N increases, we expect N_+/N to get closer

and closer to \mathcal{P}_+ . Mathematically this can be expressed as:

$$\lim_{N \rightarrow \infty} \frac{N_+}{N} = \mathcal{P}_+ = |\langle +|\psi \rangle|^2 . \quad (10.2)$$

Similarly,

$$\lim_{N \rightarrow \infty} \frac{N_-}{N} = \mathcal{P}_- = |\langle -|\psi \rangle|^2 . \quad (10.3)$$

Characterising Data

You should be familiar with the fact that in an experiment we can characterise data using:

- the **mean** (or average)
 - The mean is the average of all measurements
- the **standard deviation**
 - The standard deviation tells us how the data is “spread” around the mean.

For example, in the experiment represented by Figure 10.2, the average of our S_z measurement would be

$$\begin{aligned} \text{mean } S_z &= \frac{\left(+\frac{\hbar}{2}\right) N_+ + \left(-\frac{\hbar}{2}\right) N_-}{N} \\ &= \left(\frac{\hbar}{2}\right) \left(\frac{N_+}{N}\right) + \left(-\frac{\hbar}{2}\right) \left(\frac{N_-}{N}\right) \end{aligned} \quad (10.4)$$

Now since, for a large number of measurements, $\mathcal{P}_{\pm} \approx N_{\pm}/N$, we expect

$$\text{mean } S_z \approx \left(\frac{\hbar}{2}\right) \mathcal{P}_+ + \left(-\frac{\hbar}{2}\right) \mathcal{P}_- . \quad (10.5)$$

For this reason, we define the **expected mean value**, $\langle \hat{S}_z \rangle$, by

$$\langle \hat{S}_z \rangle := \left(\frac{\hbar}{2}\right) \mathcal{P}_+ + \left(-\frac{\hbar}{2}\right) \mathcal{P}_- , \quad (10.6)$$

where the angle brackets are short hand for “mean value”. Once again, for a large number of measurements, we expect, $(\text{mean } S_z) \approx \langle \hat{S}_z \rangle$, but, the two quantities are unlikely to be exactly the same due to statistical fluctuations.

Now, using the rules of quantum mechanics, for a given state, $|\psi\rangle$

$$\begin{aligned} \langle \hat{S}_z \rangle &= \left(\frac{\hbar}{2}\right) \mathcal{P}_+ + \left(-\frac{\hbar}{2}\right) \mathcal{P}_- \\ &= \left(\frac{\hbar}{2}\right) |\langle +|\psi \rangle|^2 + \left(-\frac{\hbar}{2}\right) |\langle -|\psi \rangle|^2 \\ &= \left(\frac{\hbar}{2}\right) \langle \psi | + \rangle \langle +|\psi \rangle + \left(-\frac{\hbar}{2}\right) \langle \psi | - \rangle \langle -|\psi \rangle \\ &= \langle \psi | \left[\left(\frac{\hbar}{2}\right) |+\rangle \langle +|\psi \rangle + \left(-\frac{\hbar}{2}\right) |-\rangle \langle -|\psi \rangle \right] \\ &= \langle \psi | [\hat{S}_z |+\rangle \langle +|\psi \rangle + \hat{S}_z |-\rangle \langle -|\psi \rangle] \\ &= \langle \psi | \hat{S}_z [|+\rangle \langle +| + |-\rangle \langle -|] |\psi \rangle , \end{aligned} \quad (10.7)$$

💬 Discussion 10.1

Justify the third equality in (10.7)

💬 Discussion 10.2

Justify the fifth equality in (10.7)

Using the completeness relation, $|+\rangle\langle+| + |-\rangle\langle-| = \hat{1}$, (10.7) becomes

$$\langle \hat{S}_z \rangle = \langle \psi | \hat{S}_z | \psi \rangle , \quad (10.8)$$

which is another way to calculate the expected mean value that might be more useful for calculations. “Expected mean value” is rather a mouthful and we refer to $\langle \hat{S}_z \rangle$ as the **expectation value**. Note that, the expectation value is **not** the expected *mean* value of any single experiment but the expected *mean* value of a large number of experiments.

For a general observable

$$\langle \hat{A} \rangle := \langle \psi | \hat{A} | \psi \rangle = \sum_n a_n \mathcal{P}_{a_n} , \quad (10.9)$$

where a_n are the eigenvalues of \hat{A} , and \mathcal{P}_{a_n} is the probability of measuring a_n .

💬 Discussion 10.3

Justify the second equality in (10.9).

💬 Discussion 10.4

Show that $\langle \hat{A} + \hat{B} \rangle = \langle \hat{A} \rangle + \langle \hat{B} \rangle$

💬 Discussion 10.5

Show that for some constant λ , $\langle \lambda \hat{A} \rangle = \lambda \langle \hat{A} \rangle$.

To get a better feel for expectation values, we’ll look at some examples in spin- $1/2$ systems. Consider the state $|+\rangle$ and S_z :

$$\begin{aligned} \langle \hat{S}_z \rangle &= \langle + | \hat{S}_z | + \rangle \\ &= \langle + | \frac{\hbar}{2} | + \rangle \\ &= \frac{\hbar}{2} \langle + | + \rangle \\ &= \frac{\hbar}{2} . \end{aligned} \quad (10.10)$$

In retrospect, this result should be obvious since, we always measure $S_z = +\hbar/2$ for the $|+\rangle$ state, so the average of our measurements must be $+\hbar/2$.

Now let's look at $|+\rangle_x$ and S_z . Using matrix notation,

$$\begin{aligned}\langle \hat{S}_z \rangle &= {}_x\langle +|\hat{S}_z|+\rangle_x \\ &= \frac{1}{\sqrt{2}}(1 \quad 1) \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (10.11)\end{aligned}$$

$$= \frac{\hbar}{4}(1 \quad 1) \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0\hbar. \quad (10.12)$$

We can understand this fact from there is a 50/50 split between the two possible measurements of $\pm\hbar/2$ so that the average must be zero. Note that we never measure $0\hbar$ for any individual measurement – it is not an expected value but the value we expect for the average of a large number of measurements.

As previously mentioned, the standard deviation measures the spread of data around the expected value. It is defined as the **square root** of the **mean** of the **square** of deviations from the mean. In other words, for some observable A , the standard deviation, ΔA , is given by

$$\Delta A = \sqrt{\langle (\hat{A} - \hat{1}\langle \hat{A} \rangle)^2 \rangle} \quad (10.13)$$

The expression for ΔA can be simplified as follows

$$\begin{aligned}\Delta A &= \sqrt{\langle (\hat{A} - \hat{1}\langle \hat{A} \rangle)^2 \rangle} \\ &= \sqrt{\langle \hat{A}^2 - 2\hat{A}\langle \hat{A} \rangle + \langle \hat{A} \rangle^2 \rangle} \\ &= \sqrt{\langle \hat{A}^2 \rangle - 2\langle \hat{A} \rangle \langle \hat{A} \rangle + \langle \hat{A} \rangle^2} \\ &= \sqrt{\langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2} \quad (10.14)\end{aligned}$$

In (10.14) we need to be careful about the difference between the square of the mean $\langle \hat{A} \rangle^2$ and the mean of the square $\langle \hat{A}^2 \rangle$. We can calculate the mean of the square using

$$\langle \hat{A}^2 \rangle = \langle \psi | \hat{A}^2 | \psi \rangle, \quad (10.15)$$

In (10.15), the square of the operator, \hat{A}^2 , indicates that we apply the operator twice in a row

$$\hat{A}^2 |\psi\rangle = \hat{A}\hat{A} |\psi\rangle = \hat{A}(\hat{A}|\psi\rangle). \quad (10.16)$$

To get some practise with the standard deviation let's return to the state $|+\rangle$ and S_z . In this case

$$\begin{aligned}\langle \hat{S}_z^2 \rangle &= \langle +|\hat{S}_z^2|+\rangle = \langle +|\hat{S}_z\hat{S}_z|+\rangle = \langle +|\hat{S}_z \frac{\hbar}{2}|+\rangle \\ &= \langle +|\left(\frac{\hbar}{2}\right)^2|+\rangle \\ &= \left(\frac{\hbar}{2}\right)^2. \quad (10.17)\end{aligned}$$

Since we already found $\langle S_z \rangle$ for this state in (10.10) we can determine the standard deviation

$$\Delta S_z = \sqrt{\langle S_z^2 \rangle - \langle S_z \rangle^2}$$

$$\begin{aligned}
 &= \sqrt{\left(\frac{\hbar}{2}\right)^2 - \left(\frac{\hbar}{2}\right)^2} \\
 &= 0\hbar.
 \end{aligned} \tag{10.18}$$

Once again, in retrospect this makes sense since there is no spread in our measurements – see Figure 10.3 (a).

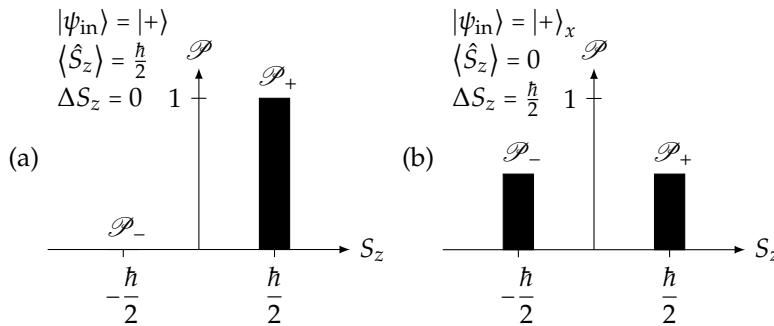


Figure 10.3: Idealised measurements of S_z with (a) a $|+\rangle$ input state and (b) with a $|+\rangle_x$ input state.

Now consider the state $|+\rangle_x$ and S_z . Once again we need

$$\begin{aligned}
 \langle \hat{S}_z^2 \rangle &= {}_x \langle + | \hat{S}_z^2 | + \rangle_x \\
 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \end{pmatrix} \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
 &= \frac{1}{2} \left(\frac{\hbar}{2} \right)^2 (1 \ 1) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
 &= \left(\frac{\hbar}{2} \right)^2.
 \end{aligned} \tag{10.19}$$

Since we already found $\langle S_z \rangle$ for this state in (10.12) we can determine the standard deviation,

$$\begin{aligned}
 \Delta S_z &= \sqrt{\langle \hat{S}_z^2 \rangle - \langle \hat{S}_z \rangle^2} \\
 &= \sqrt{\left(\frac{\hbar}{2} \right)^2 - 0\hbar^2} \\
 &= \frac{\hbar}{2}.
 \end{aligned} \tag{10.20}$$

Looking at Figure 10.3 (b) allows us to understand this result since both possible measurements differ from the mean, which is $0\hbar$, by the same amount: $\hbar/2$.

The standard deviation, ΔA , quantifies the uncertainty in experimental results. Uncertainty is a fundamental part of quantum mechanics. This means that no matter how well you design your experiment, a certain minimum amount of uncertainty may remain – this is the kind of thing we calculated in (10.18) and (10.19). We will look more closely at uncertainty when we consider the Heisenberg Uncertainty Principle in Section 10.3.

10.2 Commuting observables

Experiment 3 taught us we can not know about two incompatible observables at the same time and measurement of one observable seemed to erase knowledge of the other. In this section we will look

at what determines whether observables are compatible. To this end, we first need to define the **commutator**:

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \quad (10.21)$$

When, $[\hat{A}, \hat{B}] = 0$, we say the operators commute and if $[\hat{A}, \hat{B}] \neq 0$, we say the operators don't commute. Commutators play a very important role in quantum mechanics. When two operators commute

$$\begin{aligned} [\hat{A}, \hat{B}] &= 0 \\ \implies \hat{A}\hat{B} - \hat{B}\hat{A} &= 0 \\ \implies \hat{A}\hat{B} &= \hat{B}\hat{A}. \end{aligned} \quad (10.22)$$

Which tells that when we have commuting operators, the order of the operation does not matter. Conversely, for noncommuting operators, the order does matter.

Given two commuting operators \hat{A} and \hat{B} , consider an eigenstate $|a\rangle$ of the operator \hat{A} which has the eigenvalue a :

$$\hat{A}|a\rangle = a|a\rangle. \quad (10.23)$$

Acting on both sides of (10.23) with \hat{B} gives

$$\begin{aligned} \hat{B}\hat{A}|a\rangle &= \hat{B}a|a\rangle \\ \xrightarrow{[\hat{A}, \hat{B}] = 0} \hat{A}\hat{B}|a\rangle &= a\hat{B}|a\rangle \\ \implies \hat{A}(\hat{B}|a\rangle) &= a(\hat{B}|a\rangle) \end{aligned} \quad (10.24)$$

The last line of (10.24) tells us that $\hat{B}|a\rangle$ is an eigenstate of \hat{A} with the same eigenvalue a . If we assume that each eigenvalue is associated with a unique eigenstate¹, the state $\hat{B}|a\rangle$ must be equal to $|a\rangle$ times some number. In other words

$$\hat{B}|a\rangle = b|a\rangle, \quad (10.25)$$

for some number b . What we can conclude is that, assuming A and \hat{B} commute, implies that they have a common or **simultaneous set of eigenstates**².

Note

Commuting operators share common eigenstates.

¹ This is true if there is no degeneracy. You will look at degeneracy in more detail later.

² We also note from (10.25) that, while to operators have common eigenstates, the \hat{A} and \hat{B} eigenvalues for a given state may be different.

This has important implications for measurement in quantum mechanics. Suppose we start with a state $|\psi\rangle$ and measure A . After this measurement, the system will be in some eigenstate $|a\rangle$ of \hat{A} . Now we measure B , but since $|a\rangle$ is an eigenstate of \hat{B} , the measurement does not change the state $|a\rangle$. If we then make another measurement of \hat{A} we will get the same result, as shown in Figure 10.4 . We conclude that one can have information about commuting observables simultaneously. In other words measuring one of the observables does not erase information about the other observable. We say that commuting observables are **compatible** .

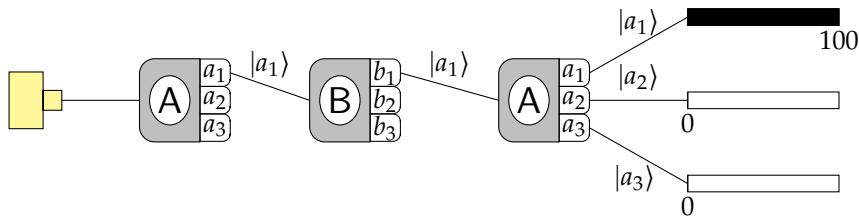


Figure 10.4: Successive measurements of commuting observables.

On the other hand, operators that do not commute cannot be measured or known simultaneously. We refer to these as incompatible observables. In Experiment 3, we saw that S_x and S_z are incompatible. Let's confirm that $[\hat{S}_z, \hat{S}_x]$ is non-zero – which is what we expect for incompatible observables;

$$\begin{aligned}
 [\hat{S}_z, \hat{S}_x] &\doteq \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\
 &\doteq \left(\frac{\hbar}{2}\right)^2 \left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right) \\
 &\doteq \left(\frac{\hbar}{2}\right)^2 \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix} \\
 \implies [\hat{S}_z, \hat{S}_x] &= i\hbar \hat{S}_y
 \end{aligned} \tag{10.26}$$

Discussion 10.6

Explain how we go from the second last to the last line of (10.26).

It turns out that the spin component operators do not commute with each other.

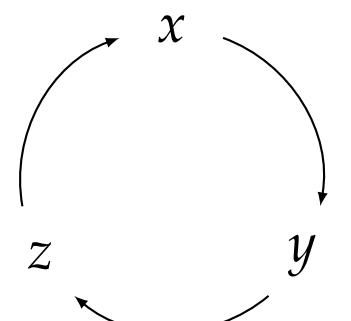
Spin Commutators

$$\begin{aligned}
 [\hat{S}_x, \hat{S}_y] &= i\hbar \hat{S}_z \\
 [\hat{S}_y, \hat{S}_z] &= i\hbar \hat{S}_x \\
 [\hat{S}_z, \hat{S}_x] &= i\hbar \hat{S}_y
 \end{aligned} \tag{10.27}$$

These relations are cyclic in the sense that we can go from one formula to next by sending $x \rightarrow y$, $y \rightarrow z$ and $z \rightarrow x$ as shown in Figure 10.5.

When represented as matrices we can sometimes tell if they commute easily. For example, two diagonal matrices will commute with each other:

$$\mathbf{AB} = \begin{pmatrix} a_1 & 0 & 0 & \dots \\ 0 & a_2 & 0 & \dots \\ 0 & 0 & a_3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} b_1 & 0 & 0 & \dots \\ 0 & b_2 & 0 & \dots \\ 0 & 0 & b_3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Figure 10.5: Cyclic permutation of x, y and z .

$$\begin{aligned}
&= \begin{pmatrix} a_1 b_1 & 0 & 0 & \dots \\ 0 & a_2 b_2 & 0 & \dots \\ 0 & 0 & a_3 b_3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \\
&= \begin{pmatrix} b_1 a_1 & 0 & 0 & \dots \\ 0 & b_2 a_2 & 0 & \dots \\ 0 & 0 & b_3 a_3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \\
&= \mathbf{BA}.
\end{aligned} \tag{10.28}$$

On the other hand, matrices may *still* commute if one or more of them are not diagonal.

Discussion 10.7

Write a summary of the important points and formulae in this chapter.

10.3 Uncertainty Principle

We've seen a connection between commutators and how accurately we can measure the properties of a quantum system. The **uncertainty principle** makes this precise.

Proof of the Uncertainty Principle

Before proceeding, we need a mathematical property of inner products, called the **Schwarz inequality**,

$$|\langle \phi | \xi \rangle|^2 \leq \langle \phi | \phi \rangle \langle \xi | \xi \rangle \tag{10.29}$$

Discussion 10.8

Prove (10.29). Hint: Let $|\gamma\rangle = |\xi\rangle - (\langle \phi | \xi \rangle / \langle \phi | \phi \rangle) |\phi\rangle$ and use $\langle \gamma | \gamma \rangle \geq 0$. Remember that for any state $|\psi\rangle$, $\langle \psi | \psi \rangle \geq 0$.

Consider two observables A and B . For A , given a state $|\psi\rangle$

$$\begin{aligned}
(\Delta A)^2 &= \left\langle \left(\hat{A} - \langle \hat{A} \rangle \hat{1} \right)^2 \right\rangle \\
&= \langle \psi | \left(\hat{A} - \langle \hat{A} \rangle \hat{1} \right)^2 | \psi \rangle \tag{10.30} \\
&= \langle \psi | \left(\hat{A} - \langle \hat{A} \rangle \hat{1} \right) \left(\hat{A} - \langle \hat{A} \rangle \hat{1} \right) | \psi \rangle
\end{aligned}$$

Now, if we let, $|\phi\rangle = (\hat{A} - \langle \hat{A} \rangle \hat{1}) |\psi\rangle$, then

$$\langle \phi | \stackrel{(9.3)}{=} \langle \psi | (\hat{A} - \langle \hat{A} \rangle \hat{1})^\dagger = \langle \psi | (\hat{A} - \langle \hat{A} \rangle \hat{1}) \tag{10.31}$$

Discussion 10.9

Explain the second equality in (10.31).

Then using (10.31), we can write (10.30) as

$$(\Delta A)^2 = \langle \phi | \phi \rangle . \quad (10.32)$$

Similarity, if we define $|\xi\rangle = (\hat{B} - \langle \hat{B} \rangle) |\psi\rangle$,

$$(\Delta B)^2 = \langle \xi | \xi \rangle . \quad (10.33)$$

Now using the Schwarz inequality, (10.29),

$$(\Delta A)^2 (\Delta B)^2 = \langle \phi | \phi \rangle \langle \xi | \xi \rangle \geq |\langle \phi | \xi \rangle|^2 \quad (10.34)$$

Letting $z = \langle \phi | \xi \rangle$, we use the fact that, for any complex number z ,

$$|\langle \phi | \xi \rangle|^2 = |z|^2 = (\operatorname{Re} z)^2 + (\operatorname{Im} z)^2 \geq (\operatorname{Im} z)^2 = \left(\frac{1}{2i} (z - z^*) \right)^2 . \quad (10.35)$$

💬 Discussion 10.10

Justify the second and final equality in (10.35).

Then combining (10.34) and (10.35) we conclude that

$$(\Delta A)^2 (\Delta B)^2 \geq \left(\frac{1}{2i} (\langle \phi | \xi \rangle - \langle \xi | \phi \rangle) \right)^2 \quad (10.36)$$

💬 Discussion 10.11

Explain how we get (10.36) from (10.34) and (10.35).

Now, using (10.31) and the definition of $|\xi\rangle$,

$$\begin{aligned} \langle \phi | \xi \rangle &= \langle \psi | (\hat{A} - \hat{A} \langle \hat{A} \rangle)(\hat{B} - \hat{B} \langle \hat{B} \rangle) |\psi\rangle \\ &= \langle \psi | (\hat{A}\hat{B} - \hat{A} \langle \hat{B} \rangle - \hat{B} \langle \hat{A} \rangle + \langle \hat{A} \rangle \langle \hat{B} \rangle) |\psi\rangle \\ &= \langle \psi | \hat{A}\hat{B} |\psi\rangle - \langle \hat{B} \rangle \langle \psi | \hat{A} |\psi\rangle - \langle \hat{A} \rangle \langle \psi | \hat{B} |\psi\rangle + \langle \hat{A} \rangle \langle \hat{B} \rangle \langle \psi | \psi \rangle \quad (10.37) \\ &= \langle \hat{A}\hat{B} \rangle - \langle \hat{B} \rangle \langle \hat{A} \rangle - \langle \hat{A} \rangle \langle \hat{B} \rangle + \langle \hat{A} \rangle \langle \hat{B} \rangle \\ &= \langle \hat{A}\hat{B} \rangle - \langle \hat{A} \rangle \langle \hat{B} \rangle . \end{aligned}$$

Similarly,

$$\langle \xi | \phi \rangle = \langle \hat{B}\hat{A} \rangle - \langle \hat{A} \rangle \langle \hat{B} \rangle . \quad (10.38)$$

💬 Discussion 10.12

Justify the steps in (10.37).

💬 Discussion 10.13

Explain how we obtain (10.38) in a similar fashion to (10.37).

Using (10.37) and (10.38) we have,

$$\langle \phi | \xi \rangle - \langle \xi | \phi \rangle = \langle \hat{A}\hat{B} \rangle - \langle \hat{B}\hat{A} \rangle = \langle [\hat{A}, \hat{B}] \rangle . \quad (10.39)$$

and combining (10.36) and (10.39) leads to³:

³ You may be worried about the factor of i in (10.40) but it turns out that $\langle [\hat{A}, \hat{B}] \rangle$ also comes with a factor of i to cancel it.

The Uncertainty Principle

$$(\Delta A)^2 (\Delta B)^2 \geq \left(\frac{1}{2i} \langle [\hat{A}, \hat{B}] \rangle \right)^2 \quad (10.40)$$

Applications of the Uncertainty principle

Applying the uncertainty principle to the S_x and S_y spin components

$$\begin{aligned} (\Delta S_x)^2 (\Delta S_y)^2 &\geq \left(\frac{1}{2i} \langle [\hat{S}_x, \hat{S}_y] \rangle \right)^2 \\ &\geq \left(\frac{1}{2i} \langle i\hbar S_z \rangle \right)^2 \\ &\geq \left(\frac{\hbar}{2} \langle S_z \rangle \right)^2 \\ \implies \Delta S_x \Delta S_y &\geq \frac{\hbar}{2} |\langle S_z \rangle|. \end{aligned} \quad (10.41)$$

Equation (10.41) tells us the minimum quantum mechanical uncertainty in any experiment measuring S_x and S_y . There may also be additional uncertainty due to, for example, properties of a particular state, experimenter error or apparatus error.

Applying (10.41) to experiment 3, whose results are summarised in Figure 10.6, we find

$$\Delta S_x \Delta S_y \geq \left(\frac{\hbar}{2} \right)^2. \quad (10.42)$$

It is important to understand that (10.42) applies to an ensemble of particles all prepared in the state $|\psi_{in}\rangle = |+\rangle$. While Experiment 3 measures S_z , we can also make many measurements of S_x , S_y on particles in the ensemble. The spread of measurements in S_x and S_y will depend on ΔS_x and ΔS_y .

Now, since (10.42) implies that the product $\Delta S_x \Delta S_y$ is not zero, we must have that each term individually is non-zero, or in other words, both $\Delta S_x \neq 0$ and $\Delta S_y \neq 0$. From this we can conclude that, when we know everything about one spin component, we can not have complete information about the other two. This in turn tells us, rather surprisingly that, spin doesn't really point in a definite direction – although we may say a particle is 'spin-up', we can only be certain about one component.

Discussion 10.14

Write a summary of the important points and formulae in this chapter.

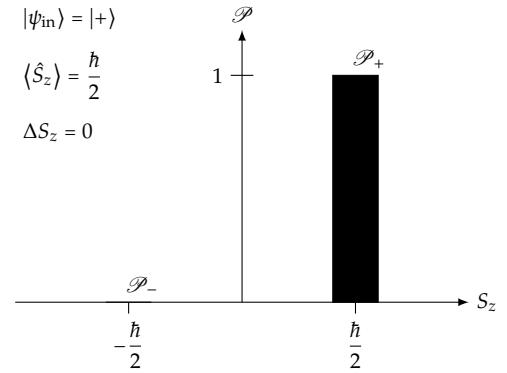


Figure 10.6: Idealised measurements of S_z with a $|+\rangle$ input state

11

S^2 , Spin-1 and General systems

\hat{S}^2 operator

This reading is based on [McIntyre et al. 2012, Section 2.6-2.8]

We can learn more about quantum spin by looking at a new operator corresponding to the magnitude of the spin vector:

$$\hat{S}^2 = \hat{\vec{S}} \cdot \hat{\vec{S}} = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2 . \quad (11.1)$$

In the z -representation, we find

$$\begin{aligned} \hat{S}^2 &\doteq \mathbf{S}_x^2 + \mathbf{S}_y^2 + \mathbf{S}_z^2 \\ &\doteq \left(\frac{\hbar}{2}\right)^2 \left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right) \\ &\doteq \left(\frac{\hbar}{2}\right)^2 \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \\ &\doteq \frac{3}{4}\hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{3}{4}\hbar^2 \mathbf{1} . \end{aligned} \quad (11.2)$$

Since \hat{S}^2 is proportional to the identity matrix it must commute with all the spin-components: \hat{S}_x , \hat{S}_y and \hat{S}_z . In addition, all states are eigenstates of \hat{S}^2 :

$$\hat{S}^2|\psi\rangle = \frac{3}{4}\hbar^2|\psi\rangle \quad (11.3)$$

Discussion 11.1

Show explicitly, in the z -representation, that (11.3) holds for some arbitrary state, $|\psi\rangle \doteq \begin{pmatrix} a \\ b \end{pmatrix}$.

Using (11.3) we find that

$$\langle \hat{S}^2 \rangle = \frac{3}{4}\hbar^2 . \quad (11.4)$$

Discussion 11.2

Show explicitly how (11.3) \implies (11.4).

Spin 1

We can do the Stern-Gerlach experiment on various atoms & particles. One always finds discrete beams exiting an analyser. As we have seen for the spin- $\frac{1}{2}$ case we get two beams. When one has three beams, the magnitude of the deflection of the beams is consistent with magnetic moments due to spin components of $1\hbar$, $0\hbar$ and $-1\hbar$. We call this a **spin-1** system and label states, based on their z -component as $|1\rangle$, $|0\rangle$ and $| -1\rangle$. As shown in, Figure 11.1, in the SPINS app, we label the states $|1\rangle$, $|0\rangle$ and $| -1\rangle$ as “+”, “0” and “-” respectively.

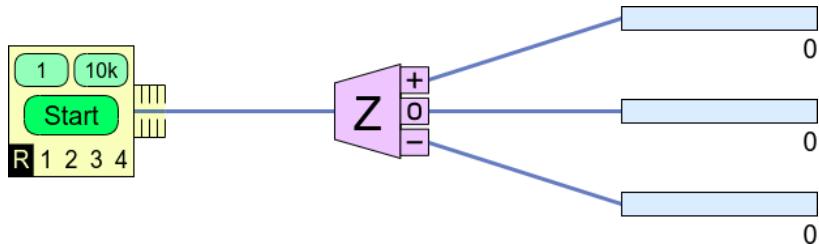


Figure 11.1: Spin 1 system modelled with the SPINS app

The three spin-1 eigenvalue equations for the operator S_z are

$$\begin{aligned}\hat{S}_z |1\rangle &= +1\hbar |1\rangle \\ \hat{S}_z |0\rangle &= 0\hbar |0\rangle \\ \hat{S}_z | -1\rangle &= -1\hbar | -1\rangle\end{aligned}\quad (11.5)$$

We will usually also use an S_z -basis for spin-1 systems. Using the fact that we can use eigenvectors to give us a standard basis¹, we can take

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad |0\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad | - 1\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (11.6)$$

¹ By standard basis, we mean that as a column vector, for each vector, one entry is 1 while all the other entries are 0.

Since operators are diagonal in their own basis, with the eigenvalues on the diagonal we can write down the S_z operator

$$\hat{S}_z = \begin{pmatrix} 1\hbar & 0 & 0 \\ 0 & 0\hbar & 0 \\ 0 & 0 & -1\hbar \end{pmatrix} = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \quad (11.7)$$

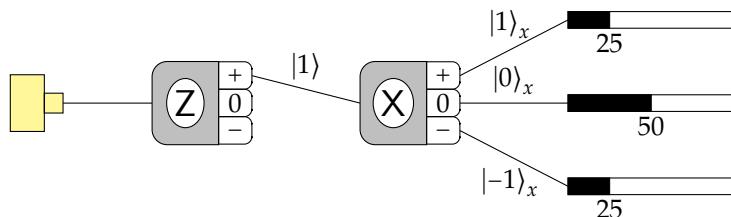


Figure 11.2: Experiment 2 in the spin-1 case.

We can repeat the equivalent of Experiment 2 for a spin 1. As shown in Figure 11.2, we prepare the system in the state $|1\rangle$ and measure the probability it ends up in $|1\rangle_x$ (ie. $S_x = 1\hbar$), $|0\rangle_x$ (ie. $S_x = 0\hbar$) or $| - 1\rangle_x$ (ie. $S_x = -1\hbar$). You might expect that, starting out

with the $|1\rangle$ state, one might guess that there is an even one third probability of measuring each of the possibly S_x components. In fact, as shown in, Figure 11.2, what one finds is

$$\begin{aligned}\mathcal{P}_{1x} &= \left| {}_x\langle 1|1\rangle \right|^2 = \frac{1}{4} \\ \mathcal{P}_{0x} &= \left| {}_x\langle 0|1\rangle \right|^2 = \frac{1}{2} \\ \mathcal{P}_{-1x} &= \left| {}_x\langle -1|1\rangle \right|^2 = \frac{1}{4}\end{aligned}\quad (11.8)$$

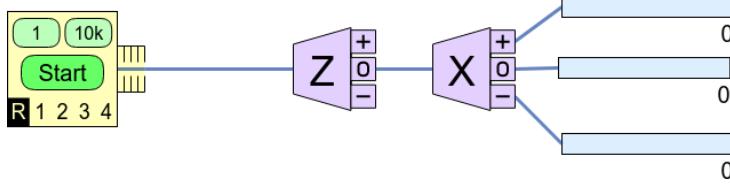
Just as we did for spin- $\frac{1}{2}$ one, we can use these results², to find expressions for the \hat{S}_x eigenstates in the \hat{S}_z basis:

$$\begin{aligned}|1\rangle_x &= \frac{1}{2}|1\rangle + \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{2}|-1\rangle \\ |0\rangle_x &= \frac{1}{\sqrt{2}}|1\rangle - \frac{1}{\sqrt{2}}|-1\rangle \\ |-1\rangle_x &= \frac{1}{2}|1\rangle - \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{2}|-1\rangle\end{aligned}\quad (11.9)$$

² and the results one obtains using the other outputs of the first Z-analyser

Discussion 11.3

Based on the form of (11.9), what do you expect for the outputs of the following set-up:



In a similar fashion, we can find expressions for the S_y eigenstates in the S_z basis:

$$\begin{aligned}|1\rangle_y &= \frac{1}{2}|1\rangle + i\frac{1}{\sqrt{2}}|0\rangle - \frac{1}{2}|-1\rangle \\ |0\rangle_y &= \frac{1}{\sqrt{2}}|1\rangle + \frac{1}{\sqrt{2}}|-1\rangle \\ |-1\rangle_y &= \frac{1}{2}|1\rangle - i\frac{1}{\sqrt{2}}|0\rangle - \frac{1}{2}|-1\rangle\end{aligned}\quad (11.10)$$

Discussion 11.4

How would you modify the set-up in Discussion 11.3 so that you could check/determine (11.10). What additional information would you need?

The matrix representations of \hat{S}_x and \hat{S}_y for spin-1 (in the S_z basis) are

$$\hat{S}_x \doteq \mathbf{S}_x = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \hat{S}_y \doteq \mathbf{S}_y = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad (11.11)$$

 Example 11.1:

Problem Suppose a system is prepared in the state $|\psi_{\text{in}}\rangle = \frac{2}{\sqrt{6}}|1\rangle - \frac{i}{\sqrt{6}}|0\rangle + \frac{i}{\sqrt{6}}|-1\rangle$. What are the chances of observing the various possible values of the z -component of \vec{S} ?

Solution The chance of finding $S_z = 1\hbar$ is

$$\begin{aligned}\mathcal{P}_1 &= |\langle 1 | \psi_{\text{in}} \rangle|^2 \\ &= \left| \langle 1 | \left(\frac{2}{\sqrt{6}}|1\rangle - \frac{i}{\sqrt{6}}|0\rangle + \frac{i}{\sqrt{6}}|-1\rangle \right) \right|^2 \\ &= \left| \frac{2}{\sqrt{6}} \langle 1|1 \rangle + \frac{-i}{\sqrt{6}} \langle 1|0 \rangle + \frac{i}{\sqrt{6}} \langle 1|-1 \rangle \right|^2 \\ &= \left| \frac{2}{\sqrt{6}} \right|^2 = \frac{2}{3}.\end{aligned}\quad (11.12)$$

The chance of finding $S_z = 0\hbar$ is

$$\begin{aligned}\mathcal{P}_0 &= |\langle 0 | \psi_{\text{in}} \rangle|^2 \\ &= \left| \langle 0 | \left(\frac{2}{\sqrt{6}}|1\rangle - \frac{i}{\sqrt{6}}|0\rangle + \frac{i}{\sqrt{6}}|-1\rangle \right) \right|^2 \\ &= \left| \frac{2}{\sqrt{6}} \langle 0|1 \rangle + \frac{-i}{\sqrt{6}} \langle 0|0 \rangle + \frac{i}{\sqrt{6}} \langle 0|-1 \rangle \right|^2 \\ &= \left| \frac{-i}{\sqrt{6}} \right|^2 = \frac{1}{6}.\end{aligned}\quad (11.13)$$

The chance of finding $S_z = -1\hbar$ is

$$\begin{aligned}\mathcal{P}_{-1} &= |\langle -1 | \psi_{\text{in}} \rangle|^2 \\ &= \left| \langle -1 | \left(\frac{2}{\sqrt{6}}|1\rangle - \frac{i}{\sqrt{6}}|0\rangle + \frac{i}{\sqrt{6}}|-1\rangle \right) \right|^2 \\ &= \left| \frac{2}{\sqrt{6}} \langle -1|1 \rangle + \frac{-i}{\sqrt{6}} \langle -1|0 \rangle + \frac{i}{\sqrt{6}} \langle -1|-1 \rangle \right|^2 \\ &= \left| \frac{i}{\sqrt{6}} \right|^2 = \frac{1}{6}.\end{aligned}\quad (11.14)$$

As expected, the probabilities sum to 1. A summary of these results are shown in Figure 11.3.

To talk about more general spin systems we need new notation. Firstly, s labels the spin. For example we could have spin- $\frac{1}{2}$, spin-1 or spin- $\frac{3}{2}$. For our original spin- $\frac{1}{2}$ system, we have $s = \frac{1}{2}$. Experimentally it is found that the number of beams leaving a Stern-Gerlach analyser is $2s+1$ with spin components ranging from $s\hbar$ to $-s\hbar$ in steps of \hbar . Secondly we use m to label the S_z component with $S_z = m\hbar$. It also turns out that the total spin squared is given by $S^2 = s(s+1)\hbar^2$. We will write a quantum state with specific s & m as $|sm\rangle$.

The eigenvalue equations for a state $|sm\rangle$ are

$$\begin{aligned}\hat{S}^2 |sm\rangle &= s(s+1)\hbar^2 |sm\rangle \\ \hat{S}_z |sm\rangle &= m\hbar |sm\rangle\end{aligned}\quad (11.15)$$

The label s denotes the spin and is called the **spin angular momentum quantum number** or **spin quantum number**. The label m tells

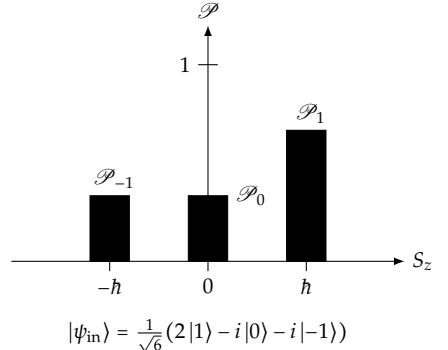


Figure 11.3: Histogram of measurements of the z -component of spin for spin-1 particle.

us about the z -component of the spin vector and is called the **spin component quantum number** or the **magnetic quantum number**. The name “magnetic quantum number” is related to its role in magnetic field experiments like Stern-Gerlach.

In this notation, the spin- $\frac{1}{2}$ states $|\pm\rangle$ are written as³

$$\begin{aligned} |+\rangle &= \left| \frac{1}{2}, \frac{1}{2} \right\rangle \\ |-\rangle &= \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \end{aligned} \quad (11.16)$$

and for the spin-1 case we have

$$\begin{aligned} |1\rangle &= |11\rangle \\ |0\rangle &= |10\rangle \\ |-1\rangle &= |1,-1\rangle . \end{aligned} \quad (11.17)$$

For the most part we will continue to use the $|\pm\rangle$ and $\{|-1\rangle, |0\rangle, |1\rangle\}$ notation for now.

As far as we know, every elementary particle has a specific and immutable value of s . For example⁴,

Higgs boson	:	spin 0
electron	:	spin $\frac{1}{2}$
photon	:	spin 1
Δ	:	spin $\frac{3}{2}$ (believed to be a composite of quarks)
graviton	:	spin 2

Discussion 11.5

Write a summary of the important points and formulae in this chapter.

³ We often put a comma in states like $|\frac{1}{2}, -\frac{1}{2}\rangle$ since $|\frac{1}{2} - \frac{1}{2}\rangle$ “looks” a bit funny.

⁴ The graviton is the hypothesised quantum of gravity analogous to the photon which is a quantum of light.

General quantum systems

We can extend what we’ve done so far to some general observable, \hat{A} , with possible quantised measurement results a_n . Using their eigenvalues to label each state, in this case the eigenvalue equation becomes,

$$\hat{A} |a_n\rangle = a_n |a_n\rangle . \quad (11.18)$$

If we use the eigenstates $|a_n\rangle$ as our basis, the operator A is represented by a diagonal matrix with the eigenvalues on the diagonal:

$$\hat{A} \doteq \mathbf{A} = \begin{pmatrix} a_1 & 0 & 0 & \dots \\ 0 & a_2 & 0 & \dots \\ 0 & 0 & a_3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

In their own basis, the eigenstates have the simple representation:

$$|a_1\rangle \doteq \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}, \quad |a_2\rangle \doteq \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}, \quad |a_3\rangle \doteq \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \end{pmatrix}, \dots \quad (11.19)$$

The projection operator corresponding to measurement of a_n is

$$\hat{P}_{a_n} = |a_n\rangle\langle a_n| . \quad (11.20)$$

Finally we have the completeness relation as a sum over the basis states

$$\sum_n \hat{P}_{a_n} = \sum_n |a_n\rangle\langle a_n| = \hat{1} . \quad (11.21)$$

 **Discussion 11.6**

Write a summary of the important points and formulae in this chapter.

12

Schrödinger Time Evolution

In classical mechanics, we can predict a particle's motion using Newton's second law

$$\vec{F} = m\vec{a} = m \frac{d^2\vec{r}}{dt^2}. \quad (1)$$

We can make predictions in quantum mechanics using the **Schrödinger equation**:

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H}(t) |\psi(t)\rangle. \quad (12.1)$$

The operator \hat{H} , which is called the **Hamiltonian** operator, represents the total energy of the system and comes from the classical Hamiltonian. By solving the Schrödinger equation, if we know the state $|\psi(t)\rangle$ at some initial time, t_0 , we can predict what it will be at some later time. This is analogous to the situation in classical mechanics, where if we know the initial position and velocity of a particle, we can predict its position and velocity at a later time by solving (1).

QM Postulate 6

The time evolution of a quantum system is determined by the Hamiltonian or total energy operator $H(t)$ through the Schrödinger equation

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H}(t) |\psi(t)\rangle.$$

Since it corresponds to the energy of the system, the eigenvalues of \hat{H} are the allowed energies of the system. This also means that the eigenstates of \hat{H} are the energy eigenstates of the system.

If we let E_n be the allowed energies, and denote the corresponding eigenstate as $|E_n\rangle$, the energy eigenvalue equation¹ is

$$\hat{H} |E_n\rangle = E_n |E_n\rangle. \quad (12.2)$$

Given a matrix representation of the Hamiltonian, \hat{H} we can diagonalise it. This means we can find the eigenvalues E_n and eigenvectors $|E_n\rangle$. For now, let's assume that we have found the eigenvalues E_n and eigenvectors $|E_n\rangle$.

Since the Hamiltonian is an **observable**, and hence hermitian, the eigenvectors $|E_n\rangle$ form a complete basis. We call this basis the

This reading is based on [McIntyre et al. 2012, Section 3.1]

¹ For reasons we will see later, (12.2), is also called the time-independent Schrödinger equation.

Energy basis. The eigenvectors will be orthonormal²

$$\langle E_n | E_m \rangle = \delta_{nm} \quad (12.3)$$

For now we assume the Hamiltonian operator, \hat{H} , does not depend on time. The eigenvectors, $|E_n\rangle$, come from diagonalising \hat{H} , so when \hat{H} doesn't depend on t , the $|E_n\rangle$ don't depend on t .

Since the **energy basis** is complete, we can write a general state as a sum over the energy eigenstates:

$$|\psi(t)\rangle = \sum_n c_n(t) |E_n\rangle . \quad (12.4)$$

Note that in (12.4), given that the $|E_n\rangle$ don't depend on t , we can conclude that only the coefficients, c_n , can depend on t .

We will now look at the Schrödinger Equation in the energy basis. Substituting, (12.4) into (12.1) we get:

$$i\hbar \frac{d}{dt} \sum_n c_n(t) |E_n\rangle = \hat{H} \sum_n c_n(t) |E_n\rangle = \sum_n c_n(t) \hat{H} |E_n\rangle . \quad (12.5)$$

Using (12.2), (12.5) simplifies to

$$i\hbar \sum_n \frac{dc_n(t)}{dt} |E_n\rangle = \sum_n c_n(t) E_n |E_n\rangle . \quad (12.6)$$

Taking the inner product of $\langle E_k |$ and (12.6) and using orthonormality gives

$$\begin{aligned} \langle E_k | i\hbar \sum_n \frac{d(c_n(t))}{dt} |E_n\rangle &= \langle E_k | \sum_n c_n(t) E_n |E_n\rangle \\ \implies i\hbar \sum_n \frac{d(c_n(t))}{dt} \langle E_k | E_n \rangle &= \sum_n c_n(t) E_n \langle E_k | E_n \rangle \\ \stackrel{(12.3)}{\implies} i\hbar \sum_n \frac{d(c_n(t))}{dt} \delta_{kn} &= \sum_n c_n(t) E_n \delta_{kn} \\ \implies i\hbar \frac{d(c_k(t))}{dt} &= c_k(t) E_k . \end{aligned} \quad (12.7)$$

⊕ Discussion 12.1

Explain, in your own words, how the Kronecker delta δ_{mn} , “collapses” sums. For example why $\sum_n c_n(t) E_n \delta_{kn} = c_k(t) E_k$.

The last line of (12.7) can be written as

$$\frac{dc_k(t)}{dt} = \frac{-iE_k}{\hbar} c_k(t) , \quad (12.8)$$

which gives us a first order differential equation for each of the coefficients $c_k(t)$. Equation (12.8) has the solution

$$c_k(t) = c_k(0) e^{-iE_k t / \hbar} , \quad (12.9)$$

where $c_k(0)$ is the initial value of $c_k(t)$ at $t = 0$.

² Strictly speaking, the energy eigenvectors will only be orthonormal if they are non-degenerate, *ie.* if they all have distinct eigenvalues. If we have degenerate eigenvectors, we can still construct an orthonormal basis using something from linear algebra called the **Gram-Schmidt process**. We will not need that in this course.

 **Discussion 12.2**

Show that (12.9) is the solution to the differential equation (12.8).

For simplicity, from now on we will write $c_k = c_k(0)$. Notice that the time dependence of each coefficient has the same exponential form albeit with a different frequency due to different values of the energy eigenvalues E_k .

In summary, if at time $t = 0$, we have the state,

$$|\psi(0)\rangle = \sum_n c_n |E_n\rangle \quad (12.10)$$

then at a later time t , our state is given by

$$|\psi(t)\rangle = \sum_n c_n e^{-iE_n t/\hbar} |E_n\rangle \quad (12.11)$$

In other words, the time-dependence of our state is found by multiplying each energy eigenstate by its particular phase factor $e^{-iE_n t/\hbar}$.

Note

Equation (12.11) **only** works when we write our state in terms of *energy* eigenstates.

To get a feel for time-dependence in quantum mechanics let's look at a few examples.

 **Example 12.1: Stationary States**

Suppose we start off, at $t = 0$ in the state,

$$|\psi(0)\rangle = |E_1\rangle, \quad (12.12)$$

then at a later time, folding in the relevant phase, we get

$$|\psi(t)\rangle = e^{-iE_1 t/\hbar} |E_1\rangle. \quad (12.13)$$

This is just the original state multiplied by an overall phase factor. We have seen before, that overall phases do not affect measurements. For example, for an observable \hat{A} , the probability of measuring the eigenvalue a_j , at some time t , is

$$\begin{aligned} \mathcal{P}_{a_j} &= |\langle a_j | \psi(t) \rangle|^2 \\ &= \left| \langle a_j | e^{-iE_1 t/\hbar} | E_1 \rangle \right|^2 \\ &= \left| \langle a_j | E_1 \rangle \right|^2, \end{aligned} \quad (12.14)$$

which doesn't depend on time and is the same as the probability as at $t = 0$. This means that we can not measure any time-dependence for this state. For this reason, we call energy eigenstates **stationary states**.

Discussion 12.3

Show that $|\langle a_j | e^{-iE_1 t/\hbar} | E_1 \rangle|^2 = |\langle a_j | E_1 \rangle|^2$.

Discussion 12.4

Explain why, for a spin- $\frac{1}{2}$ system, we can only have two possible orthogonal energy eigenstates.

Discussion 12.5

Suppose a spin half particle starts at $t = 0$, in an energy eigenstate state $|\psi(0)\rangle = |E_2\rangle$ with energy eigenvalue E_2 . What is $|\psi(t)\rangle$? Will $\langle \hat{S}_z \rangle$ depend on time? Will $\langle \hat{S}_x \rangle$ depend on time? What about $\langle \hat{A} \rangle$ if $\mathbf{A} = \begin{pmatrix} 0 & 1-i \\ 1+i & 0 \end{pmatrix}$?

Example 12.2: Simple Superposition State

Now we'll investigate the next simplest state, namely the superposition of two energy eigenstates.

$$|\psi(0)\rangle = c_1|E_1\rangle + c_2|E_2\rangle. \quad (12.15)$$

Using the same method of adding in the time dependence, we find that the time-evolution of the state is given by

$$|\psi(t)\rangle = c_1 e^{-iE_1 t/\hbar} |E_1\rangle + c_2 e^{-iE_2 t/\hbar} |E_2\rangle. \quad (12.16)$$

Now, the probability of measuring the energy to be E_1 at some time t is

$$\begin{aligned} \mathcal{P}_{E_1} &= |\langle E_1 | \psi(t) \rangle|^2 \\ &= \left| \langle E_1 | (c_1 e^{-iE_1 t/\hbar} |E_1\rangle + c_2 e^{-iE_2 t/\hbar} |E_2\rangle) \right|^2 \\ &= |c_1|^2. \end{aligned} \quad (12.17)$$

Once again this does not depend on time. The same applies to \mathcal{P}_{E_2} , so the probabilities of measuring particular energies **does not change**, as in the previous case.

Things get more interesting if we look at the probability of measuring another observable A . There are two cases to consider depending on whether \hat{A} and \hat{H} commute.

Case 1: \hat{A} and \hat{H} commute

If \hat{A} commutes with \hat{H} then \hat{A} and \hat{H} have common eigenstates. This means that measuring A is equivalent to measuring H in the sense that we use the same eigenstates to calculate probabilities. This means that, just like for energy, the probability of measuring any eigenstate of \hat{A} does not depend on time.

Case 2: \hat{A} and \hat{H} do not commute

If \hat{A} does not commute with \hat{H} , then \hat{A} and \hat{H} **do not** share eigenstates. Since the energy basis is complete, we at least know that the eigenstates of \hat{A} are made up of superposition of energy eigenstates. For simplicity suppose, the state $|a_1\rangle$ can be written as the superposition of just two energy eigenstates

$$|a_1\rangle = \alpha_1|E_1\rangle + \alpha_2|E_2\rangle ,$$

then

$$\begin{aligned} \mathcal{P}_{a_1} &= |\langle a_1|\psi(t)\rangle|^2 \\ &= \left| (\alpha_1^* \langle E_1 | + \alpha_2^* \langle E_2 |) (c_1 e^{-iE_1 t/\hbar} |E_1\rangle + c_2 e^{-iE_2 t/\hbar} |E_2\rangle) \right|^2 \\ &= \left| \alpha_1^* c_1 e^{-iE_1 t/\hbar} + \alpha_2^* c_2 e^{-iE_2 t/\hbar} \right|^2 . \end{aligned} \quad (12.18)$$

Factoring $e^{-iE_1 t/\hbar}$ out of (12.18), leads to

$$\begin{aligned} \mathcal{P}_{a_1} &= \left| e^{-iE_1 t/\hbar} \right|^2 \left| \alpha_1^* c_1 + \alpha_2^* c_2 e^{-i(E_2-E_1)t/\hbar} \right|^2 \\ &= |\alpha_1 c_1|^2 + |\alpha_2 c_2|^2 + 2 \operatorname{Re}(\alpha_1 c_1^* \alpha_2^* c_2 e^{-i(E_2-E_1)t/\hbar}) . \end{aligned} \quad (12.19)$$

which **does** depend on time. So in this case, $|a_1\rangle$ is not a stationary state. While an overall phase factors out, the relative *time-dependent* phase between the two terms remains in the final answer. The time dependence depends on the **difference** of energies of the two states in the superposition. The angular frequency, related to this difference,

$$\omega_{21} = \frac{E_2 - E_1}{\hbar} \quad (12.20)$$

is called the **Bohr frequency**.

💬 Discussion 12.6

Explain how we go from the second to the third line of (12.18).

💬 Discussion 12.7

Explain how we go from the first to the second line of (12.19).

💬 Discussion 12.8

Suppose a spin half particle starts at $t = 0$, in superposition state $|\psi(0)\rangle = \frac{1}{\sqrt{2}}|E_1\rangle + \frac{1}{\sqrt{2}}|E_2\rangle$. What is $|\psi(t)\rangle$? Suppose we have the Hamiltonian $\hat{H} = \omega_0 \hat{S}_z$. Will $\langle \hat{S}_z \rangle$ depend on time? Will $\langle \hat{S}_x \rangle$ depend on time? Will $\langle \hat{S}^2 \rangle$ depend on time?

Summary: Recipe for finding $|\psi(t)\rangle$

Problem Given a time-independent Hamiltonian \hat{H} and an initial state $|\psi(0)\rangle$, what is the probability that the eigenvalue a_j of the observable \hat{A} is measured at time t ?

Solution

1. Diagonalise \hat{H} (ie. find the eigenvalues E_n and eigenvectors $|E_n\rangle$).
2. Write $|\psi(0)\rangle$ in terms of the energy eigenstates $|E_n\rangle$.
3. Multiply each eigenstate coefficient by $e^{-iE_n t/\hbar}$ to get $|\psi(t)\rangle$
4. Calculate the probability $\mathcal{P}_{a_j} = |\langle a_j | \psi(t) \rangle|^2$.

💬 Discussion 12.9

Write a summary of the important points and formulae in this chapter.

13

Spin precession

In this section, we get some experience by applying the Schrödinger equation to a particular spin-½ system. We'll look at a spin-½ system in a **constant** magnetic field. Although this may sound rather simple, the system displays the phenomena of spin precession. Spin precession involves the rotation of an axis of rotation – see [an animation of the precession of a gyroscope](#). Spin precession was in the news in 2021 in relation to [Fermilab's Muon \$g - 2\$ experiment](#) which hints at a possible new force in nature.

In order to apply Schrödinger's equation we need to know the Hamiltonian (or total energy) of the system. Consider a magnetic dipole in a constant magnetic field. The potential energy of a magnetic dipole, $\vec{\mu}$, in a magnetic field, \vec{B} is:

$$H = -\vec{\mu} \cdot \vec{B} . \quad (13.1)$$

The magnetic moment of an electron is

$$\vec{\mu} = g \frac{q}{2m_e} \vec{S} = -\frac{e}{m_e} \vec{S} , \quad (13.2)$$

where $q = -e$ and the gyromagnetic ratio of the electron, g is approximately 2 (in this course, we will take $g = 2$). Substituting (13.2) into (13.1) gives

$$H = \frac{e}{m_e} \vec{S} \cdot \vec{B} . \quad (13.3)$$

13.1 Magnetic field in the z -direction

Firstly we'll look at a constant magnetic field in the z -direction:

$$\vec{B} = B_0 \hat{k} \quad (13.4)$$

then the Hamiltonian becomes

$$H = \frac{eB_0}{m_e} S_z = \omega_0 S_z \quad (13.5)$$

where we have defined the angular frequency, ω_0 , to be

$$\omega_0 = \frac{e}{m_e} B_0 , \quad (13.6)$$

to simplify our notation.

This reading is based on [McIntyre et al. 2012, Section 3.2.1]

This reading is based on [McIntyre et al. 2012, Section 3.2]

Now writing the Hamiltonian operator, in the S_z representation, using (13.5) gives

$$\hat{H} \doteq \omega_0 \mathbf{S}_z = \frac{\hbar\omega_0}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (13.7)$$

Notice that \hat{H} is already diagonal in this representation so that we don't need to diagonalise – step 1 of our Schrödinger time evolution recipe, from the last section, comes for free. It is easy to see that the eigenstates of \hat{H} are the basis states of the representation, namely $|\pm\rangle$, and the eigenvalues are along the diagonal of (13.7).

This means that the eigenvalue equations for \hat{H} are:

$$\begin{aligned}\hat{H}|+\rangle &= \omega_0 \hat{S}_z |+\rangle = \frac{\hbar\omega_0}{2} |+\rangle = E_+ |+\rangle \\ \hat{H}|-\rangle &= \omega_0 \hat{S}_z |-\rangle = -\frac{\hbar\omega_0}{2} |-\rangle = E_- |-\rangle,\end{aligned}\quad (13.8)$$

so that the eigenvectors and eigenvalues are:

$$\begin{aligned}|E_+\rangle &= |+\rangle \text{ with } E_+ = \frac{\hbar\omega_0}{2}, \\ |E_-\rangle &= |-\rangle \text{ with } E_- = -\frac{\hbar\omega_0}{2}.\end{aligned}\quad (13.9)$$

We often represent the information about energy eigenvalues diagrammatically as in Figure 13.1.

As shown in Figure 13.1, the energies of two states differ by $E_+ - E_- = \hbar\omega_0$. The angular frequency ω_0 (and \hbar) set the energy scale for the system. The state $|+\rangle$ has a magnetic moment that points against the magnetic field which means that it has a higher energy. Note that the fact that we have a negative charge, means that the electron's spin and its magnetic moment point in opposite directions.

We'll now look at a few examples to learn more about this system.

Example 13.1:

Given a magnetic field $\vec{B} = B_0 \hat{k}$, suppose that we initially have a spin-up state, or in other words,

$$|\psi(0)\rangle = |+\rangle. \quad (13.10)$$

This state is already in the energy basis (so step 2 of our Schrödinger recipe is done). Putting in the time dependent phase (step 3 of our recipe) we get

$$|\psi(t)\rangle = e^{-iE_+ t/\hbar} |+\rangle = e^{-i\omega_0 t/2} |+\rangle. \quad (13.11)$$

This is a particular example of what we saw in the last section: when the initial state is an energy eigenstate, the time-dependence just comes in as an overall phase which has no measurable effect. For example, the chance of measuring, $S_z = \hbar/2$, is

$$\mathcal{P}_+ = |\langle +|\psi(t)\rangle|^2 = \left| \langle +|e^{-i\omega_0 t/2} |+\rangle \right|^2 = 1, \quad (13.12)$$

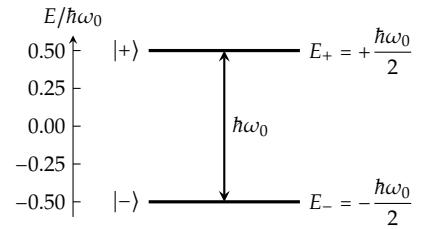


Figure 13.1: Energy level diagram of a spin-1/2 particle in a uniform magnetic field.

which does not depend on time, as expected for a stationary state. A model of this experiment is shown in Figure 13.2

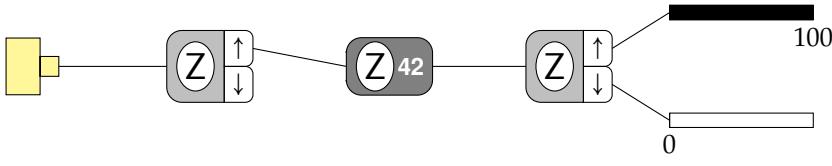


Figure 13.2: Schematic diagram of a Stern-Gerlach measurement with an applied uniform magnetic field represented by the box in the middle, with the number 42 proportional to the strength of the magnetic field, ie. $B_z = 42b_0$ where b_0 has units of tesla.

Example 13.2:

Given a magnetic field $\vec{B} = B_0 \hat{k}$, we now look at an initial state that is spin-up that is some general direction (specified by some θ and ϕ):

$$|\psi(0)\rangle = |+\rangle_n = \cos\left(\frac{\theta}{2}\right)|+\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|-\rangle \quad (13.13)$$

which can be written in matrix notation as

$$|\psi(0)\rangle \doteq \begin{pmatrix} \cos(\theta/2) \\ e^{i\phi}\sin(\theta/2) \end{pmatrix}. \quad (13.14)$$

The time-dependence is introduced by adding in the appropriate phase for each term:

$$\begin{aligned} |\psi(t)\rangle &\doteq \begin{pmatrix} e^{-iE_+t/\hbar}\cos(\theta/2) \\ e^{-iE_-t/\hbar}e^{i\phi}\sin(\theta/2) \end{pmatrix} \\ &\doteq \begin{pmatrix} e^{-i\omega_0 t/2}\cos(\theta/2) \\ e^{i\omega_0 t/2}e^{i\phi}\sin(\theta/2) \end{pmatrix} \\ &\doteq e^{-i\omega_0 t/2} \begin{pmatrix} \cos(\theta/2) \\ e^{i(\phi+\omega_0 t)}\sin(\theta/2) \end{pmatrix} \end{aligned} \quad (13.15)$$

Now that we have the state we can work out the probabilities and expectation values. The probability of measuring spin-up along the z -axis is

$$\begin{aligned} \mathcal{P}_+ &= |\langle +|\psi(t)\rangle|^2 \\ &= \left| (1 \ 0) e^{-i\omega_0 t/2} \begin{pmatrix} \cos(\theta/2) \\ e^{i(\phi+\omega_0 t)}\sin(\theta/2) \end{pmatrix} \right|^2 \\ &= \left| e^{-i\omega_0 t/2} \cos(\theta/2) \right|^2 \\ &= \cos^2(\theta/2), \end{aligned} \quad (13.16)$$

which does not depend on time. This is because, in this case, $[\hat{H}, \hat{S}_z] = 0$, so that, \hat{S}_z eigenstates are also energy eigenstates. On the other hand the probability for measuring spin up in the x -direction is

$$\mathcal{P}_{+x} = \left| {}_x\langle +|\psi(t)\rangle \right|^2$$

$$\begin{aligned}
&= \left| \frac{1}{\sqrt{2}} (1 \ 1) e^{-i\omega_0 t/2} \begin{pmatrix} \cos(\theta/2) \\ e^{i(\phi+\omega_0 t)} \sin(\theta/2) \end{pmatrix} \right|^2 \\
&= \frac{1}{2} \left| \cos(\theta/2) + e^{i(\phi+\omega_0 t)} \sin(\theta/2) \right|^2 \\
&= \frac{1}{2} \left(\cos^2(\theta/2) + \cos(\theta/2) \sin(\theta/2) \left(e^{i(\phi+\omega_0 t)} + e^{-i(\phi+\omega_0 t)} \right) \right. \\
&\quad \left. + \sin^2(\theta/2) \right), \\
&= \frac{1}{2} (1 + \sin \theta \cos(\phi + \omega_0 t)), \tag{13.17}
\end{aligned}$$

which does depend on time. This is because, in this case, the \hat{S}_x does not commute with the Hamiltonian so the states $|\pm\rangle_x$ are not stationary.

Discussion 13.1

- a) Show how we get from line 2 to line 3 in (13.17).
- b) Show how we get from line 3 to line 4 in (13.17).
- c) Show how we get from line 4 to line 5 in (13.17).

We can demonstrate the precession of the spin about the magnetic field by looking at the expectation values of the spin components. Starting with \hat{S}_z ,

$$\begin{aligned}
\langle \hat{S}_z \rangle &= \langle \psi(t) | \hat{S}_z | \psi(t) \rangle \\
&= e^{+i\omega_0 t/2} \begin{pmatrix} \cos(\theta/2) & e^{-i(\phi+\omega_0 t)} \sin(\theta/2) \end{pmatrix} \\
&\quad \times \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} e^{-i\omega_0 t/2} \begin{pmatrix} \cos(\theta/2) \\ e^{i(\phi+\omega_0 t)} \sin(\theta/2) \end{pmatrix} \\
&= \frac{\hbar}{2} (\cos^2(\theta/2) - \sin^2(\theta/2)) \\
&= \frac{\hbar}{2} \cos \theta, \tag{13.18}
\end{aligned}$$

then for \hat{S}_y we have

$$\begin{aligned}
\langle \hat{S}_y \rangle &= \langle \psi(t) | \hat{S}_y | \psi(t) \rangle \\
&= e^{+i\omega_0 t/2} \begin{pmatrix} \cos(\theta/2) & e^{-i(\phi+\omega_0 t)} \sin(\theta/2) \end{pmatrix} \\
&\quad \times \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} e^{-i\omega_0 t/2} \begin{pmatrix} \cos(\theta/2) \\ e^{i(\phi+\omega_0 t)} \sin(\theta/2) \end{pmatrix} \\
&= \frac{\hbar}{2} \sin \theta \sin(\phi + \omega_0 t), \tag{13.19}
\end{aligned}$$

and finally for \hat{S}_x

$$\langle \hat{S}_x \rangle = \langle \psi(t) | \hat{S}_x | \psi(t) \rangle = \frac{\hbar}{2} \sin \theta \cos(\phi + \omega_0 t). \tag{13.20}$$

Discussion 13.2

- a) Show how we get from line 2 to line 3 in (13.18).
- b) Show how we get from line 2 to line 3 in (13.19).
- c) Show how we get (13.20).

Figure 13.3 shows the expectation value $\langle \vec{S} \rangle$ of the spin components found in (13.18)-(13.20). We see that the expectation values of the spin components precesses about the magnetic field. The angular frequency of the precession is ω_0 . The spin vector's precession is called **Larmor precession** and its frequency is referred to as the **Larmor frequency**.

Discussion 13.3

We define the expectation value of the vector \vec{S} as

$$\langle \vec{S} \rangle = \langle S_x \rangle \hat{i} + \langle S_y \rangle \hat{j} + \langle S_z \rangle \hat{k}. \quad (13.21)$$

Explain how (13.18)-(13.20) corresponds to the motion shown in Figure 13.3.

Introducing Ehrenfest's theorem

An interesting feature of Example 13.2 is that, the expectation value of \hat{S} behaves exactly the way we would expect an object with a magnetic moment to behave classically. To see this, we need to review classical physics. Classically, a magnetic moment, $\vec{\mu}$, experiences torque,

$$\vec{\tau} = \vec{\mu} \times \vec{B}. \quad (13.22)$$

If the magnetic moment, $\vec{\mu}$, is due to the object's angular momentum,

$$\vec{\mu} = \frac{q}{2m} \vec{L}. \quad (13.23)$$

Then using the fact that for rotational motion, the torque is equal to the rate of change of momentum, we have that

$$\frac{d\vec{L}}{dt} = \vec{\tau}. \quad (13.24)$$

Substituting (13.22) and (13.23) into (13.24) leads to

$$\frac{d\vec{L}}{dt} = \frac{q}{2m} \vec{L} \times \vec{B}. \quad (13.25)$$

Equation (13.25) tells us that $\frac{d\vec{L}}{dt}$, which is the instantaneous change in \vec{L} , is perpendicular to both \vec{L} and to \vec{B} . This is what causes it to precess about \vec{B} . In addition, from (13.25), as shown below, one can conclude that the classical Larmor frequency of this rotation is $\omega_{cl} = qB/2m$.

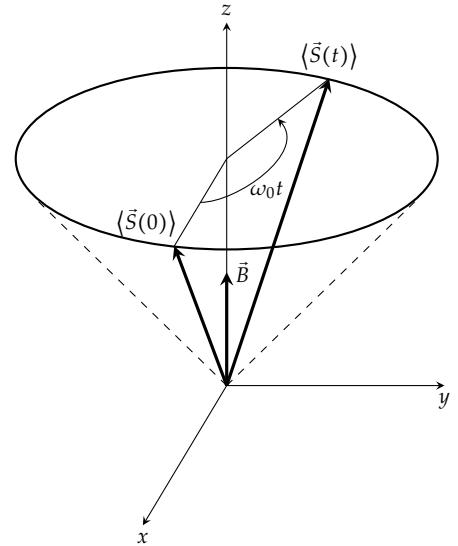


Figure 13.3: The expectation value of the spin vector precesses in a uniform magnetic field.

Example 13.3: Classical Larmor precession

Suppose we have $\vec{B} = B_0 \hat{k}$. Then

$$\vec{L} \times \vec{B} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ L_x & L_y & L_z \\ 0 & 0 & B_0 \end{vmatrix} = B_0(L_y \hat{i} - L_x \hat{j}) . \quad (13.26)$$

and (13.25) written in terms of components gives

$$\dot{L}_x = \frac{qB_0}{2m} L_y \quad (13.27)$$

$$\dot{L}_y = -\frac{qB_0}{2m} L_x \quad (13.28)$$

$$\dot{L}_z = 0 . \quad (13.29)$$

Equation (13.29) tells us that L_z is constant. Then, taking the time derivative of (13.27) gives

$$\ddot{L}_x = \frac{qB_0}{2m} \dot{L}_y \stackrel{(13.28)}{=} -\left(\frac{qB_0}{2m}\right)^2 L_x . \quad (13.30)$$

Similarly

$$\ddot{L}_y = -\left(\frac{qB_0}{2m}\right)^2 L_y . \quad (13.31)$$

You should recognise (13.30) and (13.31) as simple harmonic oscillator equations which tells us that L_x and L_y oscillate sinusoidally with angular frequency $\omega = \frac{qB_0}{2m}$.

From the preceding discussion, we see that $\langle \hat{S} \rangle$ behaves in the same way as angular momentum does classically. This is an example of **Ehrenfest's theorem** which states that quantum mechanical expectation values obey classical laws.

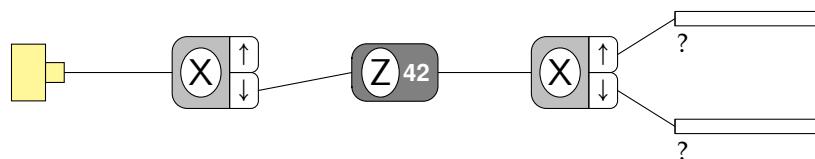


Figure 13.4: Spin precession experiment.

Example 13.4:

Problem As shown in Figure 13.4, a spin- $\frac{1}{2}$ electron, prepared in the state $|-\rangle_x$, is placed in a uniform magnetic field pointing in the z -direction after which we measure the spin components in the x -direction. What is the probability, \mathcal{P}_{+x} , of measuring the particle to be in spin-up in the x -direction?

Solution

- Following the recipe from section 12 we first determine

the eigenvectors and eigenvalues of the Hamiltonian. Since the magnetic field is in the z -direction, we know from (13.5) that $\hat{H} = \omega_0 \hat{S}_z$ which has energy eigenstates $|\pm\rangle$ with eigenvalues $\pm\hbar\omega_0/2$ with $\omega_0 = eB_0/m_e$.

2. The next step is to express the initial state, $|\psi(0)\rangle = |-\rangle_x$, with respect to the the energy basis found in the last step:

$$|\psi(0)\rangle = |-\rangle_x = \frac{1}{\sqrt{2}}|+\rangle - \frac{1}{\sqrt{2}}|-\rangle . \quad (13.32)$$

3. To find the time dependence we multiply each term in the expression from the last step by an appropriate phase:

$$\begin{aligned} |\psi(t)\rangle &= \frac{1}{\sqrt{2}}e^{-iE_+t/\hbar}|+\rangle - \frac{1}{\sqrt{2}}e^{-iE_-t/\hbar}|-\rangle \\ &= \frac{1}{\sqrt{2}}e^{-i\omega_0t/2}|+\rangle - \frac{1}{\sqrt{2}}e^{+i\omega_0t/2}|-\rangle . \end{aligned} \quad (13.33)$$

4. Finally we can calculate the probability:

$$\begin{aligned} \mathcal{P}_{+x} &= |{}_x\langle +|\psi(t)\rangle|^2 \\ &= \left| \left(\frac{1}{\sqrt{2}}(|+| + \frac{1}{\sqrt{2}}|-\rangle) \right) \left(\frac{1}{\sqrt{2}}e^{-i\omega_0t/2}|+\rangle - \frac{1}{\sqrt{2}}e^{+i\omega_0t/2}|-\rangle \right) \right|^2 \\ &= \frac{1}{4} \left| e^{-i\omega_0t/2} - e^{+i\omega_0t/2} \right|^2 \\ &= \sin^2(\omega_0 t / 2) \end{aligned} \quad (13.34)$$

A plot of the probability, \mathcal{P}_{+x} , that the system has spin up in the x direction is shown in Figure 13.5 (a). This probability is consistent with the model in Figure 13.5 (b) which shows the expectation value of the spin precessing in the xy -plane about the magnetic field, which points in the the z -direction.

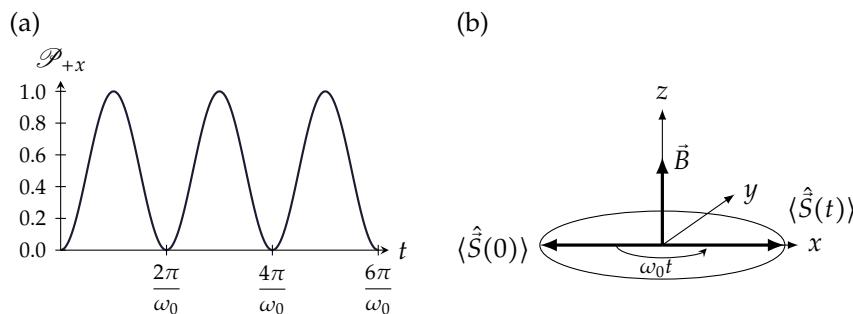


Figure 13.5: (a) Probability of a spin component measurement and (b) the corresponding precession of the expectation value of the spin.

Discussion 13.4

Write a summary of the important points and formulae in this chapter.

14

Magnetic Field in a (more) general direction

We now look at a more general magnetic field with a component in the x -direction:

$$\vec{B} = B_0 \hat{k} + B_1 \hat{i} \quad (14.1)$$

We could just tackle this problem by changing our axes so that \vec{B} aligns with the z -axis which reduces the problem to one we have done already. We will not do this for two reasons:

1. To get some practise finding solutions to the Schrödinger equation
2. We want to look at some questions in the original coordinate system

As shown in Figure 14.1, the magnetic field (14.1) is in the xz -plane. It makes an angle of θ with the z -axis.

In what follows, it is convenient to define two Larmor frequencies as follows:

$$\omega_0 := \frac{eB_0}{m_e}, \quad \omega_1 := \frac{eB_1}{m_e}, \quad (14.2)$$

so that the Hamiltonian, in this case, is

$$\begin{aligned} H &= -\vec{\mu} \cdot \vec{B} \\ &= \omega_0 S_z + \omega_1 S_x. \end{aligned} \quad (14.3)$$

Which means \hat{H} has the matrix representation

$$\hat{H} \doteq \frac{\hbar}{2} \begin{pmatrix} \omega_0 & \omega_1 \\ \omega_1 & -\omega_0 \end{pmatrix}. \quad (14.4)$$

Discussion 14.1

Explain how we go from the first to the second line of (14.3).

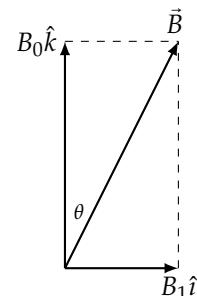


Figure 14.1: A uniform magnetic field in a general direction.

Discussion 14.2

Explain how we go from (14.3) to (14.4).

The first thing to note is that (14.4) is not diagonal – this tells us that its eigenstates are not $|\pm\rangle$ and we need to diagonalise it. Starting with the characteristic equation

$$\begin{vmatrix} \frac{\hbar}{2}\omega_0 - \lambda & \frac{\hbar}{2}\omega_1 \\ \frac{\hbar}{2}\omega_1 & -\frac{\hbar}{2}\omega_0 - \lambda \end{vmatrix} = 0$$

$$\implies -\left(\frac{\hbar}{2}\omega_0\right)^2 + \lambda^2 - \left(\frac{\hbar}{2}\omega_1\right)^2 = 0, \quad (14.5)$$

so

$$\lambda = \pm \frac{\hbar}{2} \sqrt{\omega_0^2 + \omega_1^2}. \quad (14.6)$$

💬 Discussion 14.3

How do we obtain the first line of (14.5)?

💬 Discussion 14.4

How do we go from the first to the second line of (14.5)?

Notice from (14.6) that when $\omega_1 = 0$, the eigenvalues are $\pm(\hbar\omega_0)/2$. Why is this what we expect from our solution to the previous problem?

We could find the eigenvectors directly but, by rewriting the Hamiltonian, the solution will be obvious. Firstly, from the Figure 14.1, we see that

$$\tan \theta = \frac{B_1}{B_0} = \frac{\omega_1}{\omega_0}. \quad (14.7)$$

💬 Discussion 14.5

Justify the second equality in (14.7).

Using (14.7), we can write the Hamiltonian operator as

$$\hat{H} \doteq \frac{\hbar}{2} \sqrt{\omega_0^2 + \omega_1^2} \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix} \quad (14.8)$$

💬 Discussion 14.6

Explain how we can get (14.8) from (14.4) using (14.7).

Recall the expression we found for \mathbf{S}_n in Reading 7:

$$\hat{S}_n \doteq \frac{\hbar}{2} \begin{pmatrix} \cos \theta & \sin \theta e^{-i\phi} \\ \sin \theta e^{i\phi} & -\cos \theta \end{pmatrix}, \quad (14.9)$$

so that if \hat{n} points in the direction of \vec{B} , comparing (14.8) and (14.9), we see that

$$\hat{H} = \left(\sqrt{\omega_0^2 + \omega_1^2} \right) \hat{S}_n, \quad (14.10)$$

with $\phi = 0$. We already found the eigenstates of \hat{S}_n which are $|\pm\rangle_n$:

$$\begin{aligned} |+\rangle_n &= \cos\left(\frac{\theta}{2}\right)|+\rangle + \sin\left(\frac{\theta}{2}\right)|-\rangle \\ |-\rangle_n &= \sin\left(\frac{\theta}{2}\right)|+\rangle - \cos\left(\frac{\theta}{2}\right)|-\rangle . \end{aligned} \quad (14.11)$$

Spin flip

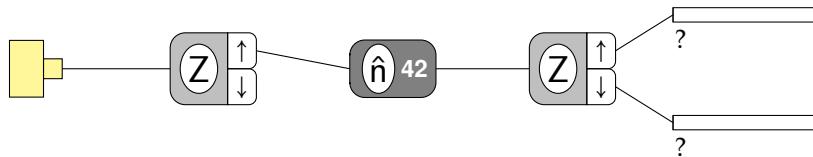


Figure 14.2: A spin precession experiment with a uniform magnetic field aligned in a general direction \hat{n} .

Now let's look at what happens if we start with a spin-up state along the z -axis. We would like to know the probability of a **spin flip**. In other words what is the chance that our particle will be found to be spin-down along the z -axis some time later? A schematic of a spin flip experiment is shown in Figure 14.2. Our initial state is $|\psi(0)\rangle = |+\rangle$, which, following our recipe, we would like to write in terms of the energy eigenstates of the system, $|\pm\rangle_n$. One way of doing this is to use completeness relation

$$\hat{1} = |+\rangle_n n|+| + |-\rangle_n n|-, \quad (14.12)$$

which gives

$$\begin{aligned} |\psi(0)\rangle &= |+\rangle = \hat{1}|+\rangle \\ &= (|+\rangle_n n|+| + |-\rangle_n n|-) |+\rangle \\ &= (n|+|+) |+\rangle_n + (n|-|+) |-\rangle_n \\ &= \cos\left(\frac{\theta}{2}\right)|+\rangle_n + \sin\left(\frac{\theta}{2}\right)|-\rangle_n \end{aligned} \quad (14.13)$$

Discussion 14.7

Explain how go we from the third to the fourth line of (14.13).

Having expressed our state in terms of energy eigenstates of the Hamiltonian, we can put in the appropriate phase factor for each term:

$$|\psi(t)\rangle = e^{-iE_+t/\hbar} \cos\left(\frac{\theta}{2}\right)|+\rangle_n + e^{-iE_-t/\hbar} \sin\left(\frac{\theta}{2}\right)|-\rangle_n , \quad (14.14)$$

where $E_{\pm} = \pm \frac{\hbar}{2} \sqrt{\omega_0^2 + \omega_1^2}$.

We can now calculate the probability of a spin flip, \mathcal{P}_{+-} ,

$$\begin{aligned} \mathcal{P}_{+-} &= |\langle -|\psi(t)\rangle|^2 \\ &= \left| \langle -| \left(e^{-iE_+t/\hbar} \cos\left(\frac{\theta}{2}\right)|+\rangle_n + e^{-iE_-t/\hbar} \sin\left(\frac{\theta}{2}\right)|-\rangle_n \right) \right|^2 \\ &= \left| e^{-iE_+t/\hbar} \cos\left(\frac{\theta}{2}\right) \langle -|+ \rangle_n + e^{-iE_-t/\hbar} \sin\left(\frac{\theta}{2}\right) \langle -|- \rangle_n \right|^2 \\ &= \left| e^{-iE_+t/\hbar} \cos\left(\frac{\theta}{2}\right) \sin\left(\frac{\theta}{2}\right) + e^{-iE_-t/\hbar} \sin\left(\frac{\theta}{2}\right) (-\cos\left(\frac{\theta}{2}\right)) \right|^2 \end{aligned}$$

$$\begin{aligned}
 &= \sin^2\left(\frac{\theta}{2}\right) \cos^2\left(\frac{\theta}{2}\right) \left|1 - e^{i(E_+ - E_-)t/\hbar}\right|^2 \\
 &= \sin^2(\theta) \sin^2\left(\frac{(E_+ - E_-)t}{2\hbar}\right)
 \end{aligned} \tag{14.15}$$

Discussion 14.8

Show how we go from the second last to the last line of (14.15).

We can rewrite \mathcal{P}_{+-} in terms of the Larmor frequencies to obtain **Rabi's formula** for the probability of a spin flip:

Rabi's formula

$$\mathcal{P}_{+-} = \frac{\omega_1^2}{\omega_0^2 + \omega_1^2} \sin^2\left(\frac{\sqrt{\omega_0^2 + \omega_1^2}}{2} t\right) \tag{14.16}$$

Discussion 14.9

Show how we go from (14.15) to (14.16).

Rabi's formula, (14.16), will have important applications in later on. To get a feel for the formula, let's look at two simple examples. Firstly if $\omega_1 = 0$, then $\mathcal{P}_{+-} = 0$ since the initial state is an eigenstate of the Hamiltonian and consequently stationary. Secondly, if $\omega_0 = 0$, then the probability oscillates with angular frequency of ω_1 between 0 and 1 (See Figure 14.3 (a)). As shown in Figure 14.3 (b), this case corresponds to spin precession in the xy -plane around a magnetic field along the x -direction with the spin certain to flip with an angular frequency of ω_1 .

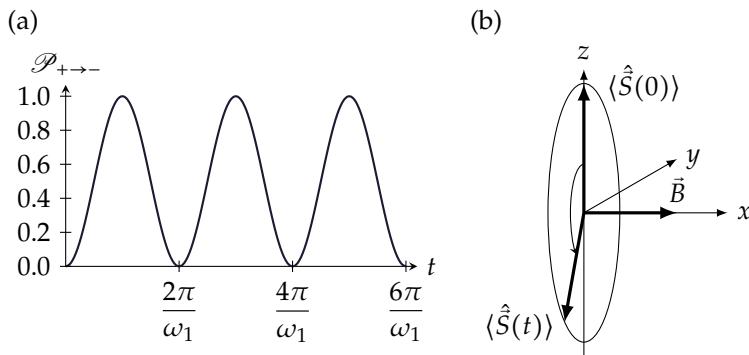


Figure 14.3: (a) Spin-flip probability for a uniform magnetic field in the x -direction and (b) the corresponding precession of the expectation value of the spin.

In general, when ω_0 and ω_1 are both non-zero, the probability of a spin-flip is never one, so we are never guaranteed a spin-flip at any time. If for example, the x -component of the field is much smaller than the z -component, or in other words if $\omega_1 \ll \omega_0$, then \mathcal{P}_{+-} oscillates between 0 and a small number, approximately at the angular frequency ω_0 (See Figure 14.4).

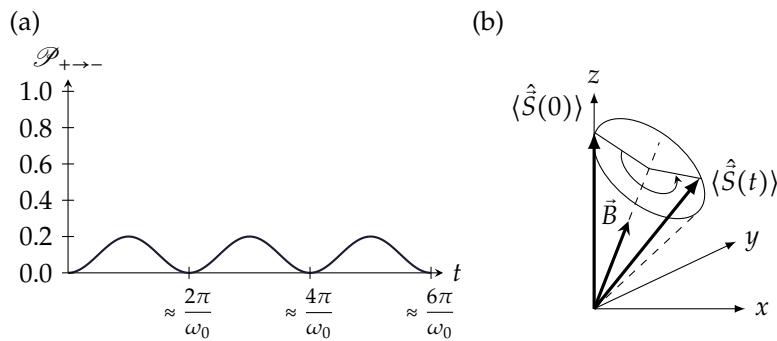


Figure 14.4: (a) Spin-flip probability for a uniform magnetic field with x - and z -components and (b) the corresponding precession of the expectation value of the spin.

Example 14.1:

A spin- $\frac{1}{2}$ particle with a magnetic moment is prepared in the state $|-\rangle$ and is subject to a uniform applied magnetic field $\vec{B} = B_0 \hat{j}$. Find the probability of measuring spin up in the z -direction after a time t .

The magnetic field is in the y direction which tells us that the Hamiltonian operator is $\hat{H} = \omega_0 \hat{S}_y$, which will have eigenstates $|\pm\rangle_y$ with energy eigenvalues $\pm\hbar\omega_0/2$.

Now writing the initial state in terms of the energy eigenstates gives

$$\begin{aligned} |\psi(0)\rangle &= |-\rangle \\ &= (|+\rangle_y + |-\rangle_y) |-\rangle \\ &= \left(\frac{1}{\sqrt{2}}(|+\rangle_y - |-\rangle_y)\right) |+\rangle_y + \left(\frac{1}{\sqrt{2}}(|+\rangle_y - |-\rangle_y)\right) |-\rangle_y \\ &= \frac{-i}{\sqrt{2}} |+\rangle_y + \frac{i}{\sqrt{2}} |-\rangle_y. \end{aligned} \quad (14.17)$$

Discussion 14.10

How do we go from

1. the first line to the second line of (14.17)
2. the third line to the fourth line of (14.17)

Now we can find $|\psi(t)\rangle$ by multiplying in the energy dependent phase factors giving,

$$\begin{aligned} |\psi(t)\rangle &= \frac{-i}{\sqrt{2}} e^{-iE_+ t/\hbar} |+\rangle_y + \frac{i}{\sqrt{2}} e^{-iE_- t/\hbar} |-\rangle_y \\ &= \frac{-i}{\sqrt{2}} e^{-i\omega_0 t/2} |+\rangle_y + \frac{i}{\sqrt{2}} e^{+i\omega_0 t/2} |-\rangle_y \end{aligned} \quad (14.18)$$

Discussion 14.11

How would the last line of (14.18) change if we had the same initial state, with $\vec{B} = -B_0 \hat{j}$?

$$\mathcal{P}_+ = |\langle +|\psi(t)\rangle|^2$$

$$\begin{aligned}
&= \left| \langle + | \left(\frac{-i}{\sqrt{2}} e^{-i\omega_0 t/2} |+\rangle_y + \frac{i}{\sqrt{2}} e^{+i\omega_0 t/2} |-\rangle_y \right) \right|^2 \\
&= \left| \frac{-i}{\sqrt{2}} e^{-i\omega_0 t/2} \langle + | + \rangle_y + \frac{i}{\sqrt{2}} e^{+i\omega_0 t/2} \langle + | - \rangle_y \right|^2 \\
&= \left| \frac{-i}{\sqrt{2}} e^{-i\omega_0 t/2} \left(\frac{1}{\sqrt{2}} \right) + \frac{i}{\sqrt{2}} e^{+i\omega_0 t/2} \left(\frac{1}{\sqrt{2}} \right) \right|^2 \\
&= \frac{1}{4} |-ie^{-i\omega_0 t/2} + ie^{+i\omega_0 t/2}|^2 \\
&= \frac{1}{4} |-2 \sin(\omega_0 t/2)|^2 \\
&= \sin^2(\omega_0 t/2). \tag{14.19}
\end{aligned}$$

As shown in Figure 14.5, the probability varies between 0 and 1 with angular frequency ω_0 consistent with the model shown in Figure 14.5 (b) of a spin vector precessing in the xz -plane about the magnetic field pointing along the y -axis.

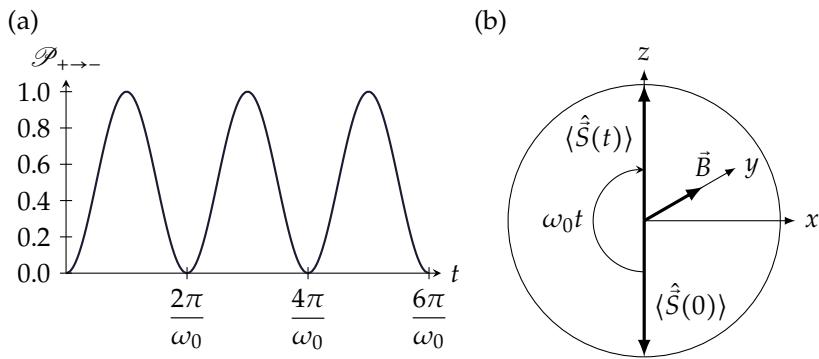


Figure 14.5: (a) Spin measurement probability and (b) the corresponding precession of the expectation value of the spin.

Comment on Rabi's formula

Rabi's formula, (14.16), does not *only* apply to spin-flips. If we can have a two level system with a Hamiltonian that we can write in the form (14.8), the Rabi formula will tell us the probability that the system will transition from one state to another.

Discussion 14.12

Write a summary of the important points and formulae in this chapter.

15

Neutrino Oscillations

Neutrinos are sub-atomic particles postulated to explain the apparent violation of the conservation of angular momentum and energy in beta decay: $n \rightarrow p + e^-$ (where n is the neutron, p is the proton and e^- is the electron). Neutrinos are uncharged, *very* light particles moving at relativistic speeds. They are extremely hard to detect as they apparently only interact through the **weak interaction**. The weak force is a very short range force 10^{-18} m that is involved in nuclear decay

Neutrinos are produced in nuclear beta decay. In beta decay, neutrons (n) or protons (p), in atomic nuclei, release either an electron (e^-) or a positron (e^+), as in:

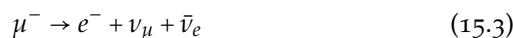


where ν_e is the electron neutrino and $\bar{\nu}_e$ is the electron antineutrino.

Neutrinos are produced in large numbers by thermonuclear reactions in the sun. The number of neutrinos that we have measured coming from the sun is much less than our expectation which is based on the predictions of our current models of stellar physics. This deficit of neutrinos from the sun is called the **solar neutrino problem**. Before looking at how this problem is now explained, we need to know that there are other reactions associated with other types (or *flavours*) of neutrinos. For example there is a subatomic particle, called the pion, (π), decays into a muon, (μ):



Pions are produced, for example, in high energy collisions between atomic nuclei in colliders. The muon, μ , behaves like an electron but has a larger mass and the ν_μ is the muon neutrino. The decay of muons to electrons also produces neutrinos



where ν_μ is a muon neutrino.

There is no known law of nature that forbids *flavour* changing neutrino reactions, $\nu_e \leftrightarrow \nu_\mu$, in which an electron neutrino changes into a muon neutrino (and vice versa). Such flavour changes are called **neutrino oscillations**.

This reading is based on [McIntyre et al. 2012, Section 3.3]

The neutrino association with electrons or muons is a property of the weak interaction. This means that $|\nu_\mu\rangle$ and $|\nu_e\rangle$ are eigenstates of weak interaction Hamiltonian. On the other hand, in free space, weak interaction, being a short range force, is not relevant. This means that the Hamiltonian affecting neutrinos in free space is *different* from inside nuclei. The main contribution to a neutrino's energy in free space is mainly due to its relativistic energy.

Now the key point is that since we have two different Hamiltonian there may be different eigenstates for the different Hamiltonians. We call the eigenstates of the free space Hamiltonian **mass eigenstates** and denote them by $|\nu_1\rangle$ and $|\nu_2\rangle$. The difference between these two sets of eigenstates is what makes neutrino oscillations possible. Our best explanation of the solar neutrino problem, namely the deficit of measured electron neutrinos coming from the sun, is through **neutrino oscillations**. The idea is that on their journey from the sun, some of the electron neutrinos change into muon neutrinos. It turns out that neutrino oscillation is analogous to a spin flip we studied previously and we can understand neutrino oscillations through an application of Rabi's formula¹.

Assuming both bases, $\{|\nu_\mu\rangle, |\nu_e\rangle\}$ and $\{|\nu_1\rangle, |\nu_2\rangle\}$, are complete², we can write one set of eigenstates in terms of the other. We assume this linear combination can be written

$$\begin{aligned} |\nu_e\rangle &= \cos\left(\frac{\theta}{2}\right)|\nu_1\rangle + \sin\left(\frac{\theta}{2}\right)|\nu_2\rangle \\ |\nu_\mu\rangle &= \sin\left(\frac{\theta}{2}\right)|\nu_1\rangle - \cos\left(\frac{\theta}{2}\right)|\nu_2\rangle \end{aligned} \quad (15.4)$$

The angle $\theta/2$ is called the **mixing angle** (the factor of a $1/2$ has been chosen to be consistent with what we had for spin- $1/2$ systems). In principle there could be a relative phase between the terms in (15.4) but for simplicity we will neglect it.

Discussion 15.1

Suppose we take $|\nu_e\rangle = a|\nu_1\rangle + b|\nu_2\rangle$. If we let $a = \cos(\theta/2)$, what range should we take for θ so that we can have $a \in [0, 1]$?

Discussion 15.2

Suppose we take $|\nu_e\rangle = \cos(\theta/2)|\nu_1\rangle + b|\nu_2\rangle$. Show that if we assume b is real and non-negative, the normalisation of $|\nu_e\rangle$ implies that $b = \sin(\theta/2)$.

Discussion 15.3

Suppose we take $|\nu_e\rangle = \cos(\theta/2)|\nu_1\rangle + \sin(\theta/2)|\nu_2\rangle$. If $|\nu_\mu\rangle$ is orthogonal to $|\nu_e\rangle$ and normalised, show that we can take $|\nu_\mu\rangle = \sin(\theta/2)|\nu_1\rangle - \cos(\theta/2)|\nu_2\rangle$.

¹ If you are interested you can find out more about the solar neutrino problem [here](#).

² Actually there is at least one other kind of neutrino that we know about – the tau neutrino – to get a feel for the physics, we'll neglect the tau neutrino for now. There may even be a fourth kind of neutrino, the *sterile* neutrino.

To examine neutrino oscillations we consider an electron neutrino, ν_e , created in some nuclear reaction and that then flies through empty space to a detector. Given this, we would like to know the probability of detecting a muon neutrino, ν_μ ? The first step will be to express our initial state in terms of the relevant energy eigenstates:

$$\begin{aligned} |\psi(0)\rangle &= |\nu_e\rangle \\ &= \cos\left(\frac{\theta}{2}\right)|\nu_1\rangle + \sin\left(\frac{\theta}{2}\right)|\nu_2\rangle, \end{aligned} \quad (15.5)$$

and the time evolution of our state will be

$$|\psi(t)\rangle = e^{-iE_1 t/\hbar} \cos\left(\frac{\theta}{2}\right)|\nu_1\rangle + e^{-iE_2 t/\hbar} \sin\left(\frac{\theta}{2}\right)|\nu_2\rangle. \quad (15.6)$$

Now, in free space, the energy eigenvalues are just the relativistic energies of the particles

$$E_i = \sqrt{(pc)^2 + (m_i c^2)^2}. \quad (15.7)$$

If the neutrinos are highly relativistic, which means that the energy related to their momentum is much larger than their rest mass, or in other words

$$mc^2 \ll pc, \quad (15.8)$$

we find that

$$\begin{aligned} E_i &= pc \left(1 + \left(\frac{m_i c^2}{pc} \right)^2 \right)^{\frac{1}{2}} \\ &\approx pc \left(1 + \frac{1}{2} \left(\frac{m_i c^2}{pc} \right)^2 \right) \end{aligned} \quad (15.9)$$

💬 Discussion 15.4

Show that $[E] = [mc^2] = [pc]$.

💬 Discussion 15.5

Show how we get from the first to the second line of (15.9).

Recycling Rabi

Mathematically the calculation of a spin flip in the spin- $\frac{1}{2}$ system we studied and neutrino oscillation problem is the same. This means that the formula for the probability of a spin flip and neutrino flavour change is the same:

$$\begin{aligned} \mathcal{P}_{\nu_e \rightarrow \nu_\mu} &= |\langle \nu_\mu | \psi(t) \rangle|^2 \\ &= \sin^2 \theta \sin^2 \left(\frac{(E_1 - E_2)t}{2\hbar} \right) \end{aligned} \quad (15.10)$$

❖ Discussion 15.6

Justify the statement

"Mathematically the calculation of a spin flip in the spin- $\frac{1}{2}$ system we studied and neutrino oscillation problem is the same".

Now, for the neutrinos

$$E_1 - E_2 \approx \frac{1}{2} \frac{(m_1 c^2)^2}{pc} - \frac{1}{2} \frac{(m_2 c^2)^2}{pc} = \frac{c^3}{2p} (m_1^2 - m_2^2) \quad (15.11)$$

❖ Discussion 15.7

Justify the first (approximate) equality in (15.11).

Highly relativistic neutrinos move at nearly the speed of light, so, if L is the distance they travel, then the time taken is $t \approx L/c$. Furthermore, we can approximate their momentum as $p \approx E/c$ so that

$$E_1 - E_2 \approx \frac{c^4}{2E} (m_1^2 - m_2^2), \quad (15.12)$$

and (15.10) becomes

$$\mathcal{P}_{\nu_e \rightarrow \nu_\mu} \approx \sin^2(\theta) \sin^2\left(\frac{(m_1^2 - m_2^2)Lc^3}{4E\hbar}\right). \quad (15.13)$$

❖ Discussion 15.8

Explain why, if we are working to first order in mc^2/pc , we can use the approximation, $p \approx E/c$, rather than $p \approx E/c - \frac{1}{2c} \left(\frac{m_i c^2}{pc} \right)^2$ in (15.12).

Looking at (15.13) we see that the probability oscillates as a function of L ranging from 0 to $\sin^2 \theta$. This means that if we measure fractions of different neutrinos³ at a known distance from the sun, and compare that to a model of expected fractions, we can infer the difference in the squares of the masses, $m_1^2 - m_2^2$, as well as θ . Results from solar neutrino observations and other experiments tell us that

$$m_1^2 - m_2^2 \approx 8 \times 10^{-5} \text{ eV}^2/\text{c}^4 \quad (15.14)$$

and

$$\theta \approx 69^\circ. \quad (15.15)$$

³ as a function of their energy

16

Time Dependent Hamiltonians

We have studied time-independent Hamiltonians. When the Hamiltonian depends on time, we **can not** use the recipe we developed **but** certain forms of time-dependence *can* be solved. We will look at sinusoidal time-dependence which is common.

This reading is based on [McIntyre et al. 2012, Section 3.4]

Magnetic resonance

Rabi's formula

$$\mathcal{P}_{+-} = \frac{\omega_1^2}{\omega_0^2 + \omega_1^2} \sin^2 \left(\frac{\sqrt{\omega_0^2 + \omega_1^2}}{2} t \right) \quad (16.1)$$

Looking at Rabi's formula, (16.1), for the probability of a spin-flip, \mathcal{P}_{+-} , you should be able to see that we can make \mathcal{P}_{+-} larger by taking $\omega_1 \gg \omega_0$. Now, recall that ω_1 is proportional to B_x which in turn is at 90° to the particle's initial spin-component: $S_z = \hbar/2$. In other words, we can optimise the chance of a spin flip by having the magnetic field orthogonal to spin of the particle, *i.e.* $\vec{B} \perp \vec{S}$. On the other hand we can also see from (16.1) that if the magnitude of the magnetic field, $|B_x|$, is small the flip is likely to take longer

It turns out we can efficiently induce a spin flip even with a small magnetic field *if* the magnetic field is time-dependent. In particular oscillations of the magnetic field should resonate with the Larmor frequency. This phenomenon is called **magnetic resonance**. The time-dependent field induces transitions between Energy states separated by $\hbar\omega_0$ as shown in Figure 16.1. An example is Nuclear magnetic resonance (NMR) used to study properties of various materials. We can use magnetic resonance to determine the Larmor frequency which in turn tells us about the magnetic moments of nuclei and internal magnetic fields. Magnetic resonance imaging (MRI) is used for medical diagnosis¹.

¹ MRI uses NMR and the original acronym was NMRI for but for marketing reasons the "N" for "nuclear" was taken out.

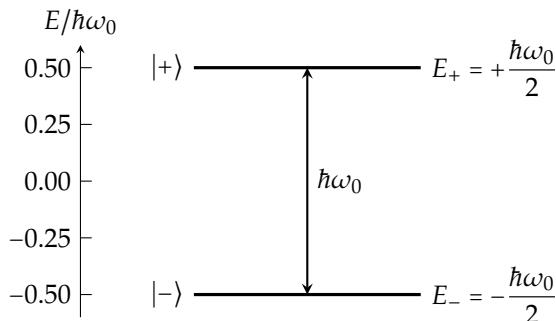


Figure 16.1: Energy level diagram of a spin- $\frac{1}{2}$ particle in a uniform magnetic field.

Discussion 16.1

Looking back at Reading 13, explain how an NMR measurement of the Larmor frequency will allow us to infer information about the magnetic moments of nuclei and internal magnetic fields.

Before studying the quantum system, it is useful to get a classical perspective. Due to the fact that it experiences a torque, $\vec{\tau} = \vec{\mu} \times \vec{B}$, an object with a magnetic moment will precess in a magnetic field. Now, imagine you are in the frame that moves with the magnetic moment, $\vec{\mu}$. In this frame, by definition, the magnetic moment, $\vec{\mu}$, does not move. The fact that $\vec{\mu}$ does not move means that it is not experiencing a torque which in turn implies that there is no magnetic field. To change the orientation of the spin, we can use a magnetic field. Since this will be the only field in the rotating frame, it can be small to have an effect. If there is a small field, there will be a small torque which in turn implies a long time for the flip.

Lets say we apply a constant magnetic field in the x -direction, in the rotating frame, $\vec{B} = B_1 \hat{i}$, to induce a spin flip. This field will not be changing in the rotating frame but it will be moving at the precession frequency in the original frame with the form.

$$\vec{B} = B_1 \cos(\omega t) \hat{i} + B_1 \sin(\omega t) \hat{j} \quad (16.2)$$

Note that, in what follows we will allow for ω not being the same as the Larmor frequency, ω_0 . This will often be the case in magnetic resonance experiments when the frequency of the applied magnetic field does not exactly match the Larmor frequency.

Discussion 16.2

How do we obtain (16.2) starting with $\vec{B} = B_1 \hat{i}$ in the rotating frame?

Motivated by our classical discussion, we now consider the quantum mechanical problem of a spin- $\frac{1}{2}$ field subject to the magnetic field

$$\vec{B} = B_0 \hat{k} + B_1 [\cos(\omega t) \hat{i} + \sin(\omega t) \hat{j}] \quad (16.3)$$

The magnetic field in the z -direction splits the energy of the $|+\rangle$ and



Figure 16.2: Perspective in a rotating reference frame. [Image source](#).

$|-\rangle$ states as shown in Figure 16.1. Note that the figure shows the split with $B_1 = 0$. The difference in energies will not be the same when B_1 is non-zero. So using (16.3), the Hamiltonian is

$$\begin{aligned} H &= -\vec{\mu} \cdot \vec{B} \\ &= \omega_0 S_z + \omega_1 (\cos(\omega t) S_x + \sin(\omega t) S_y) , \end{aligned} \quad (16.4)$$

where

$$\omega_0 = \frac{e}{m_e} B_0 , \quad \omega_1 = \frac{e}{m_e} B_1 . \quad (16.5)$$

So the Hamiltonian operator has the matrix representation

$$\hat{H} \doteq \frac{\hbar}{2} \begin{pmatrix} \omega_0 & \omega_1 e^{-i\omega t} \\ \omega_1 e^{i\omega t} & -\omega_0 \end{pmatrix} . \quad (16.6)$$

Discussion 16.3

Show how we go from (16.4) to (16.6).

Since (16.6) depends on time, we can not use earlier techniques, but as the $|\pm\rangle$ basis is complete, we can write

$$|\psi(t)\rangle = c_+(t)|+\rangle + c_-(t)|-\rangle \doteq \begin{pmatrix} c_+(t) \\ c_-(t) \end{pmatrix} . \quad (16.7)$$

Substituting (16.6) and (16.7) into the Schrödinger equation

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle , \quad (16.8)$$

gives

$$i\hbar \frac{d}{dt} \begin{pmatrix} c_+(t) \\ c_-(t) \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} \omega_0 & \omega_1 e^{-i\omega t} \\ \omega_1 e^{i\omega t} & -\omega_0 \end{pmatrix} \begin{pmatrix} c_+(t) \\ c_-(t) \end{pmatrix} \quad (16.9)$$

which leads to the coupled first order differential equations

$$\begin{aligned} i\hbar \dot{c}_+(t) &= \frac{\hbar \omega_0}{2} c_+(t) + \frac{\hbar \omega_1}{2} e^{-i\omega t} c_-(t) \\ i\hbar \dot{c}_-(t) &= \frac{\hbar \omega_1}{2} e^{i\omega t} c_+(t) - \frac{\hbar \omega_0}{2} c_-(t) , \end{aligned} \quad (16.10)$$

where the dot denotes a time derivative, ie $\dot{c}_+ = \frac{dc_+}{dt}$.

❖ Discussion 16.4

Show how we go from (16.9) to (16.10).

Our discussion of the classical problem, suggests that the problem will be simpler in the rotating frame. If we use, $|\tilde{\psi}(t)\rangle$, to denote the state vector in rotating frame then

$$|\tilde{\psi}(t)\rangle = c_+(t)e^{i\omega t/2}|+\rangle + c_-(t)e^{-i\omega t/2}|-\rangle \stackrel{def}{=} \begin{pmatrix} e^{i\omega t/2}c_+(t) \\ e^{-i\omega t/2}c_-(t) \end{pmatrix}. \quad (16.11)$$

Since we don't yet know how to change state vectors when we do a rotation, you can either take it on faith that (16.11) is the correct form, or just view it as a clever change of variables that makes the problem simpler. Calling the coefficients of the rotating state vector, $\alpha_{\pm}(t)$, we can write

$$|\tilde{\psi}(t)\rangle = \alpha_+(t)|+\rangle + \alpha_-(t)|-\rangle \stackrel{def}{=} \begin{pmatrix} \alpha_+(t) \\ \alpha_-(t) \end{pmatrix}. \quad (16.12)$$

Comparing (16.11) and (16.12), we conclude that

$$\begin{aligned} c_+(t) &= e^{-i\omega t/2}\alpha_+(t) \\ c_-(t) &= e^{i\omega t/2}\alpha_-(t) \end{aligned} \quad (16.13)$$

and substituting (16.13) into (16.10) leads to

$$\begin{aligned} i\hbar\dot{\alpha}_+(t) &= -\frac{\hbar\Delta\omega}{2}\alpha_+(t) + \frac{\hbar\omega_1}{2}\alpha_-(t) \\ i\hbar\dot{\alpha}_-(t) &= \frac{\hbar\omega_1}{2}\alpha_+(t) + \frac{\hbar\Delta\omega}{2}\alpha_-(t), \end{aligned} \quad (16.14)$$

where we have defined $\Delta\omega$ as short hand for the difference between the angular frequency of the rotating field and the first Larmor frequency, ie. $\Delta\omega = \omega - \omega_0$.

❖ Discussion 16.5

Show how substituting (16.13) into (16.10) leads to (16.14).

Now, (16.14) can be viewed as the Schrödinger equation for in the rotating state vector $|\tilde{\psi}(t)\rangle$ with respect to a new Hamiltonian \tilde{H} :

$$i\hbar\frac{d}{dt}|\tilde{\psi}(t)\rangle = \tilde{H}|\tilde{\psi}(t)\rangle, \quad (16.15)$$

where the new matrix representation of the new Hamiltonian is

$$\tilde{H} \stackrel{def}{=} \frac{\hbar}{2} \begin{pmatrix} -\Delta\omega & \omega_1 \\ \omega_1 & \Delta\omega \end{pmatrix}. \quad (16.16)$$

❖ Discussion 16.6

Show how substituting (16.12) and (16.16) into (16.15) leads to (16.14).

So by a clever change of variables, we have transmuted our original time-dependent problem to one with a time-independent Hamiltonian. What is more its a problem we have solved before – if we replace $(-\Delta\omega)$ with ω_0 we get the same Hamiltonian that we had in section 3.2b where we derived the Rabi formula. As in section 3.2b we are interested in finding the spin-flip probability, \mathcal{P}_{+-} , that is, the chance that a particle starting in the $|+\rangle$ state is later found in the $|-\rangle$ state. The rotational transformation does not change the $|\pm\rangle$ basis states so if

$$|\psi(0)\rangle = |+\rangle \quad (16.17)$$

then²

$$|\tilde{\psi}(0)\rangle = |+\rangle . \quad (16.18)$$

Given this initial state, the chance of a spin-flip is

$$\mathcal{P}_{+-} = |\langle -|\tilde{\psi}(t)\rangle|^2 = |c_-(t)|^2 . \quad (16.19)$$

² Strictly speaking, $|\tilde{\psi}(0)\rangle$, picks up an overall phase but this has no physical significance.

❖ Discussion 16.7

How does the second equality in (16.19) come about?

Then using (16.13), we get

$$\begin{aligned} |c_-(t)|^2 &= \left| e^{i\omega t/2} \alpha_-(t) \right|^2 \\ &= |\alpha_-(t)|^2 \\ &= |\langle -|\tilde{\psi}(t)\rangle|^2 \end{aligned} \quad (16.20)$$

❖ Discussion 16.8

How do the second equality and third equalities in (16.20) come about?

Equation (16.20) tells us that the probability we need is $|\langle -|\tilde{\psi}(t)\rangle|^2$, but $|\tilde{\psi}(t)\rangle$ is subject to the time-independent Hamiltonian (16.16), so we can recycle the Rabi formula exchanging $\Delta\omega$ for ω_0 :

$$\begin{aligned} \mathcal{P}_{+-} &= \frac{\omega_1^2}{\Delta\omega^2 + \omega_1^2} \sin^2 \left(\frac{\sqrt{\Delta\omega^2 + \omega_1^2}}{2} t \right) \\ &= \frac{\omega_1^2}{(\omega - \omega_0)^2 + \omega_1^2} \sin^2 \left(\frac{\sqrt{(\omega - \omega_0)^2 + \omega_1^2}}{2} t \right) . \end{aligned} \quad (16.21)$$

A spin-flip under the influence of a rotating field is called **Rabi flopping**. This is different to the spin flipping due to precession in a constant magnetic field in which there is no exchange of energy. In magnetic resonance, there is an exchange of energy due to the fact that the Hamiltonian is time-dependent.

Now the Rabi flopping probability (16.21), \mathcal{P}_{+-} , has the angular frequency

$$\Omega = \sqrt{(\omega - \omega_0)^2 + \omega_1^2} , \quad (16.22)$$

which is often called the **generalised Rabi frequency** whereas ω_1 is usually called the **Rabi frequency**. The two are equal, that is $\Omega = \omega_1$, when the rotating field is on resonance, *i.e.* when $\omega = \omega_0$. When $\omega = \omega_0$, the probability of a spin flip (16.21), simplifies:

$$\mathcal{P}_{+-} = \sin^2\left(\frac{\omega_1}{2}t\right), \quad (16.23)$$

which tells us that there is a 100% chance of a spin-flip with angular frequency ω_1 . As we move away from resonance, that is as ω moves away from ω_0 , the maximum value of \mathcal{P}_{+-} decreases. When \mathcal{P}_{+-} is plotted as function of ω , as in Figure 16.3, this resonant behaviour is evident. The shape in Figure 16.3 is called a **Lorentzian curve** which, in this case has a full width at half maximum (FWHM) of $2\omega_1$.

Discussion 16.9

What is the definition of full width at half maximum (FWHM)? Can you show that the FWHM of (16.21) is $2\omega_1$?

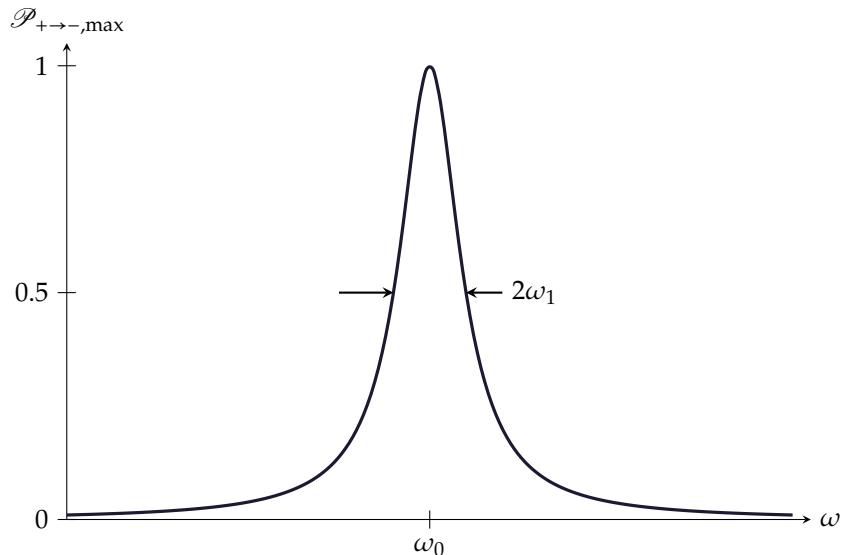


Figure 16.3: Magnetic resonance curve showing the maximum probability of a spin flip as a function of the applied frequency.

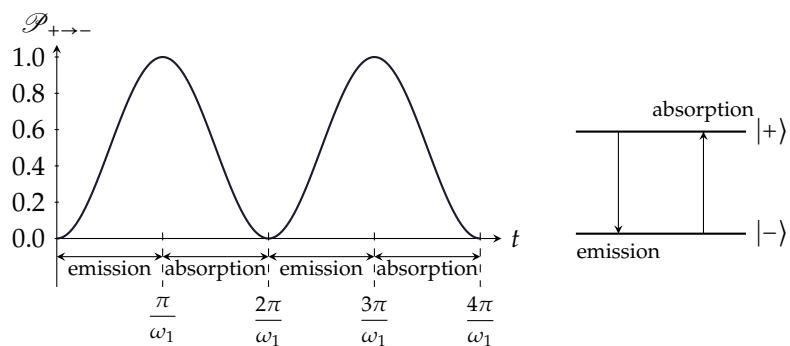


Figure 16.4: Rabi oscillations of the spin-flip probability for the resonance condition.

The spin flip probability when we have resonance, namely when $\omega = \omega_0$, as a function of time is plotted in Figure 16.4. Noting that

ω_1 is proportional to the rotating field magnitude, B_1 , we see that the rate of flipping increases as we increase the field, but even if the field is small there is still a 100% chance of a flip eventually. In fact, after time $t = \pi/\omega_1$, there is a 100% chance of spin-flip. This means we can produce a flip with just a pulse lasting $T = \pi/\omega_1$. This is called a **π -pulse** and despite the inherent uncertainty of Quantum Mechanics we can use a **π -pulse** to make a transition from one state to another with 100% probability. This is an example of Quantum Mechanical Engineering. The change from an upper to a lower state is called **emission** whereas the change from a lower to an upper state is called **absorption**, as shown in Figure 16.4.

Light-matter Interactions

The Rabi flopping Hamiltonian, (16.6), can also model how atoms absorb and emit light. When atoms interact with light, the electric dipole of atom interacts with the oscillating electric field of a light wave exchanging energy. Although atoms have many energy levels, if the frequency of the light is close to just one of the atom's Bohr frequencies we can model the atom as a two-level system.

Now consider two energy levels of an atom. We call the lower state $|g\rangle$ (for ground) & the higher state $|e\rangle$ (for excited) as shown in Figure 16.5.

The angular frequency, ω_0 , is defined by the energy difference $E_e - E_g = \hbar\omega_0$ and the applied light has frequency ω which is close but may not be equal to ω_0 .

We can split up the Hamiltonian (16.6) into two pieces, identifying the 1st term as the atomic Hamiltonian, H_{atom} , and the 2nd term is the Hamiltonian, H_{int} , due to the interaction of light & the atom:

$$\hat{H} \doteq \frac{\hbar}{2} \begin{pmatrix} \omega_0 & \omega_1 e^{-i\omega t} \\ \omega_1 e^{i\omega t} & -\omega_0 \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} \omega_0 & 0 \\ 0 & -\omega_0 \end{pmatrix} + \frac{\hbar}{2} \begin{pmatrix} 0 & \omega_1 e^{-i\omega t} \\ \omega_1 e^{i\omega t} & 0 \end{pmatrix} = \mathbf{H}_{\text{atom}} + \mathbf{H}_{\text{int}} . \quad (16.24)$$

In this context, ω_1 is the modulus of the off-diagonal matrix element of the interaction Hamiltonian connecting two states:

$$\omega_1 = \frac{2}{\hbar} |\langle e | \hat{H}_{\text{int}} | g \rangle| \quad (16.25)$$

The Rabi formula, (16.21) then gives us the probability for transitions. The two possible transitions correspond to absorption ($|g\rangle \rightarrow |e\rangle$) and emission ($|e\rangle \rightarrow |g\rangle$) of photons by the atom. Studying transitions teaches us about the energy levels of a system. This process is called **spectroscopy**. The ultimate goal is to determine the Hamiltonian.

Now, if the matrix element between two states is zero, i.e. $\langle e | \hat{H}_{\text{int}} | g \rangle = 0$, then no transition occurs. We refer to this as a **forbidden transition**. Whether a transition is allowed or forbidden is called a **selection rule**. It turns out that selection rules often relate to symmetries in the system.



Figure 16.5: Energy level diagram of a two-level atom interacting with an applied light field of frequency ω .

 **Discussion 16.10**

Write a summary of the important points and formulae in this chapter.

Quantum Spookiness

Many features of quantum mechanics are very different from the way we experience our macroscopic world. The fact that what and how we observe a quantum system can affect the outcome of measurements runs counter to the idea of an independent observer. In this section we give an introduction to some of the more “spooky” aspects of the subject.

This reading is based on [McIntyre et al. 2012, Chapter 4]

Einstein-Podolsky-Rosen Paradox

The probabilistic nature of quantum mechanics and the way that role that measurement plays and integral role never sat well will Albert Einstein. He felt that there should be an independent objective reality separate from observers. Say we measure a particle to be spin up in the z -direction. In quantum mechanics we can not be certain about what its spin was before the measurement was made. Worse than that, we can not be certain about the other components of spin since, as we have seen, measurements of one component affects our knowledge of the other components. We say that the particle does not have a really have a definite spin until we measure it. Einstein was unhappy with this view point and thought that quantum mechanics was not the whole story since we can not use it, even in principle, to predict all the components of a particle’s spin given known initial conditions.

Einstein, Boris Podolsky, and Nathan Rosen (EPR) came up with a thought experiment to demonstrate what they viewed as the profoundly problematic features of quantum mechanics. This thought experiment is referred to as the **EPR Paradox**. What they argued is that if you believe measurements on two particles, which are far away from each other, can not instantaneously affect each other, you must conclude that the particles *really* have definite components of spin *independent* of our measurements.

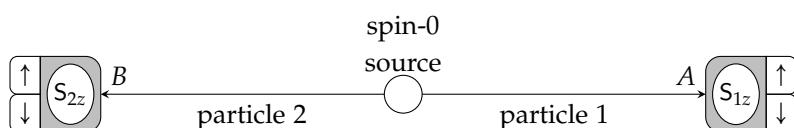


Figure 17.1: EPR thought experiment

A variation on the EPR experiment¹ is shown in fig. 17.1. A spin-0 particle decays into two spin- $\frac{1}{2}$ particles². The conservation of linear momentum tells us that they fly off in opposite directions. Two observers, *A* and *B*, on opposite sides of the source, each measure the spin of one of the particles produced by the decay. Since the total spin of the two particles is zero, the conservation of angular momentum tells us that the particle must have opposite spin. So if both observers measure spin components along the same direction, if one observer measures spin up, the other observer must measure spin down. The state of the two particle system is given by

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|+\rangle_1 |-\rangle_2 - |-\rangle_1 |+\rangle_2) . \quad (17.1)$$

The subscripts 1 and 2 label the particles – the notation³ $|+\rangle_1 |-\rangle_2$ means that particle 1 is in the state $|+\rangle$ and particle 2 is in the state $|-\rangle$. Equation (17.1) is a superposition of two states:

- particle 1 is spin up and particle 2 is spin down
- particle 2 is spin up and particle 1 is spin down

The minus sign in eq. (17.1) ensures that the overall spin of the two particle system is zero – you will hopefully learn about how that works in another course⁴.

The kets and operators for each particle are independent of one another. For example, \hat{S}_{1z} acts only on particle 1 kets:

$$\hat{S}_{1z} |+\rangle_1 |-\rangle_2 = (\hat{S}_{1z} |+\rangle_1) |-\rangle_2 = +\frac{\hbar}{2} |+\rangle_1 |-\rangle_2 . \quad (17.2)$$

Discussion 17.1

[McIntyre et al. 2012, Problem 4.1] Show that the quantum state vector of a two-particle system must be a product $|\psi\rangle_1 |\phi\rangle_2$ of two single-particle state vectors rather than a sum $|\psi\rangle_1 + |\phi\rangle_2$. Hint: consider the action of a single-particle state operator on the two-particle state vector.

Discussion 17.2

Write $|+\rangle_{x1} |-\rangle_{x2}$ out in the S_z -basis. How does it differ from the state in eq. (17.1)? Is

$$|\phi\rangle = \frac{1}{\sqrt{2}}(|+\rangle_{x1} |-\rangle_{x2} - |-\rangle_{x1} |+\rangle_{x2}) , \quad (17.3)$$

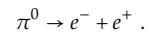
different from the state in eq. (17.1)?

When taking inner products we only combine bras and kets for the same particle :

$$({}_1(+| {}_2(-|)(|+\rangle_1 |-\rangle_2) = ({}_1(+| +\rangle_1)({}_2(-| -\rangle_2) = 1 \quad (17.4)$$

¹ David Bohm came up with this version of the experiment which was updated by N. David Mermin.

² An example of such a process could be a neutral π meson at rest decaying into an electron and a positron:



³ Mathematicians refer to this way of combining kets as the tensor product of the two Hilbert spaces.

⁴ see [McIntyre et al. 2012, Chapter 11] for details

💬 Discussion 17.3

Show that $|\psi\rangle$ in eq. (17.1) is normalised.

Figure 17.1 shows observer A measures the spin component of particle 1 and observer B measures the spin component 1.

💬 Discussion 17.4

Explain how we can see from the form of eq. (17.1) that the probability for observer A to measure spin up is 50% and 50% for spin down. Show that there is the same 50 – 50 split for observer B .

From Discussion 17.4 we see that if they observe a series of decays each observer will obtain a random sequence of spin up and spin down measurements and for a large sample, there will be a 50 – 50 ratio. There will however be a definite correlation between the observations of A and B . For example, if A measures spin up, we know with 100% certainty that B will measure spin down. We say that the two sets of results are perfectly anti-correlated. The state eq. (17.1) is known as an **entangled state**. Although the spin of a particular particle is uncertain, the spins of the two particles have a definite relationship (ie. an entanglement) with each other.

💬 Discussion 17.5

Suppose observer A makes the following sequence of measurements: + + + - - + - +. What *must* B have measured (assuming not particles are “lost”)?

Now suppose that our two observers are separated by some distance, with B slightly closer to the source than A . If B measures $S_{2z} = +\hbar/2$ we know that A will measure $S_{1z} = -\hbar/2$. This means that the wave function has instantaneously collapsed into the state $|-\rangle_1 |+\rangle_2$ when B makes the measurement. Einstein called this “spooky action at a distance”. Although the outcome of B ’s measurement is random its relationship to what A will measure is not⁵.

💬 Discussion 17.6

Can you see why, even though the collapse of the wave function is apparently instantaneous, B can not use it to transmit information to A ?

Now, the EPR argument is that since A and B are far apart, the spin of particle 1 must be independent of what B does, which means it must have had a definite spin all along. This is in contrast to the idea that the spin only became definite after a measurement was made. A essential assumption of the EPR argument is that

⁵ The story gets even more mind-boggling when you throw special relativity into the mix. Can you show that there can be an observer, travelling relative to A and B , who would claim that it was actually A who made a measurement before B ? This should give you a deeper appreciation why Einstein really didn’t like “spookiness”.

physically separated particle can not instantaneously affect - this is called **locality**.

As long as both observers choose to measure spin in the same direction, the anti-correlation of the results will be the same. For example if B measures spin up in the x -direction ($S_{2x} = \hbar/2$), then we can be 100% certain that A will measure $S_{1x} = -\hbar/2$. This leads EPR to claim that all the components of spin have some definite values before the measurement is made. On the other hand, using quantum mechanics, we have seen that we can only be certain about one component of the spin. According to EPR, the only what out of this quagmire is to conclude that quantum mechanics is not the whole story – there must be some other theory that gives a complete description of spin and quantum mechanics in a incomplete description.

If EPR are right, then there is some other theory, involving unknown or hidden variables, that we do not know. One can suppose that there is some **local hidden variable theory** which a set of rules or **instruction sets** that determines the results of measurements. Locality means that there is no “spooky action at a distance” – the result of any measurement is determined by the properties of a particle at the instant and place it is measured. The rules of this hidden variable theory must be such that it produces the same experimentally confirmed results as quantum mechanics. For many years, most physicists just assumed that this was an academic argument – if quantum mechanics and this mysterious hidden variable theory make the same predictions we can not experimentally distinguish between the two. If we can not experimentally determine which theory is correct, then the question is really beyond what physicists can do or “metaphysical”.

Fortunately, in 1964, John Bell demonstrated that there *are* observations that can be used to distinguish between a local hidden variable theory and quantum mechanics. Bell came up with a rather general inequality but we will look at a specific case to make things a little easier.

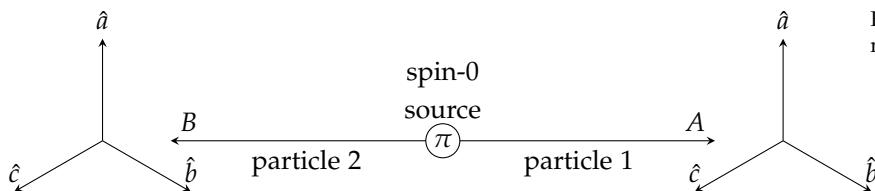


Figure 17.2: A Bell thought experiment

The clever trick Bell came up with is to consider a generalisation of the EPR experiment in which observers align their detectors along different axes. As shown in fig. 17.2, in the experiment we are looking at each observer can align their detector along one of three axes \hat{a} , \hat{b} or \hat{c} . Each axis makes an angle of 120° with the other two. The observers, who randomly choose a particular axis for each measurement, will either observe a particle to be spin up or spin

down along whatever axis they have aligned their detector. We will denote the results as + for spin up and – for spin down. We can record both observers measurements using a pair – for example, +- means observer A measured + and B measured -. There are four possible results ++, +- , -+, and --. We can also simply note whether observations are the *same* (++) or *different* (+-) or (-+).

A local hidden variable theory must have a set of rules that specifies ahead of time what the result of an observer's measurement along a particular axis will be. For example, $(\hat{a}+, \hat{b}+, \hat{c}+)$ corresponds to a measurement of spin up along all of the axes. For the state, eq. (17.1), measurements of observers along the same direction *must* be either +- or -+ (why?). To ensure that this happens, our local hidden variable theory must have rules for the eight possibilities shown in table 17.1.

Discussion 17.7

Can you explain why table 17.1 covers all the possible pairings for eq. (17.1)?

We now need to see if we can set up the proportions of N_1 to N_8 in table 17.1 so that our local hidden variable theory makes the same predictions as quantum mechanics. Let's see how these numbers relate to the results of spin component measurements. There are nine different ways the two observers can align their detectors $\hat{a}\hat{a}$, $\hat{a}\hat{b}$, $\hat{a}\hat{c}$, $\hat{b}\hat{a}$, $\hat{b}\hat{b}$, $\hat{b}\hat{c}$, $\hat{c}\hat{a}$, $\hat{c}\hat{b}$ and $\hat{c}\hat{c}$. For particles in set 1, no matter how the observers align their axes, the particles are always aligned differently. This means that in this case, the chance of observing opposite spin alignments $\mathcal{P}_{\text{opp}} = 1$ and the chance of measuring the same spin alignments is $\mathcal{P}_{\text{same}} = 0$. Similarly for particles in set 8, $\mathcal{P}_{\text{opp}} = 1$ and $\mathcal{P}_{\text{same}} = 0$. Now, for particles in set 2, i.e. $(\hat{a}+, \hat{b}+, \hat{c}-)$ and $(\hat{a}-, \hat{b}+, \hat{c}+)$, the possible results are +-, +-, ++, +- , -+, ++, -- and -+ (check that you agree with this). This means that there are 4 possible ways for the measurements to be the same and 5 possible ways for measurements to be different. So that in this case $\mathcal{P}_{\text{same}} = 4/9$ and $\mathcal{P}_{\text{opp}} = 5/9$. You can check that we get the same pattern for sets 3 to 7 so that in summary,

$$\begin{aligned} \mathcal{P}_{\text{opp}} &= 1 \\ \mathcal{P}_{\text{same}} &= 0 \end{aligned} \quad \left. \right\} \text{set 1 and 8}$$

$$\begin{aligned} \mathcal{P}_{\text{opp}} &= 5/9 \\ \mathcal{P}_{\text{same}} &= 4/9 \end{aligned} \quad \left. \right\} \text{sets 2 to 7} \quad (17.5)$$

Discussion 17.8

Check that you agree with the probabilities in eq. (17.5).

To work out the overall probabilities, we need to sum up all the possibilities. The probability of a particular set, for example, set 1

Population	Particle 1	Particle 2
N_1	$(\hat{a}+, \hat{b}+, \hat{c}+)$	$(\hat{a}-, \hat{b}-, \hat{c}-)$
N_2	$(\hat{a}+, \hat{b}+, \hat{c}-)$	$(\hat{a}-, \hat{b}-, \hat{c}+)$
N_3	$(\hat{a}+, \hat{b}-, \hat{c}+)$	$(\hat{a}-, \hat{b}+, \hat{c}-)$
N_4	$(\hat{a}+, \hat{b}-, \hat{c}-)$	$(\hat{a}-, \hat{b}+, \hat{c}+)$
N_5	$(\hat{a}-, \hat{b}+, \hat{c}+)$	$(\hat{a}+, \hat{b}-, \hat{c}-)$
N_6	$(\hat{a}-, \hat{b}+, \hat{c}-)$	$(\hat{a}+, \hat{b}-, \hat{c}+)$
N_7	$(\hat{a}-, \hat{b}-, \hat{c}+)$	$(\hat{a}+, \hat{b}+, \hat{c}-)$
N_8	$(\hat{a}-, \hat{b}-, \hat{c}-)$	$(\hat{a}+, \hat{b}+, \hat{c}-)$
,		

Table 17.1: Instruction sets (Hidden variables.)

occurring a particular measurement is $N_1 / \sum N_i$. This means that the using eq. (17.5) the probabilities are we have that

$$\begin{aligned}\mathcal{P}_{\text{same}} &= \frac{1}{\sum N_i} \frac{4}{9} (N_2 + N_3 + N_4 + N_5 + N_6 + N_7) \\ \mathcal{P}_{\text{opp}} &= \frac{1}{\sum N_i} \left(N_1 + N_8 + \frac{5}{9} (N_2 + N_3 + N_4 + N_5 + N_6 + N_7) \right)\end{aligned}\quad (17.6)$$

💬 Discussion 17.9

Can you explain the equalities in eq. (17.6)?

💬 Discussion 17.10

How can we conclude from eq. (17.6) that

$$\begin{aligned}\mathcal{P}_{\text{same}} &\leq \frac{4}{9} \\ \mathcal{P}_{\text{opp}} &\geq \frac{5}{9}\end{aligned}\quad (17.7)$$

Equation (17.7) must be satisfied no matter what the values of N_1-N_8 are, or in other words, no matter what local hidden variable theory we have. Such inequalities, which are satisfied by local hidden variable theories, are called **Bell inequalities**.

Let's see what quantum mechanics predicts for $\mathcal{P}_{\text{same}}$ and \mathcal{P}_{opp} using the tools we have learnt. Supposed observer A has measured $+$ along some axis. Let this axis be the z -axis. Observer B will make a measurement along some axis \hat{n} which makes an angle θ with the z -axis.

💬 Discussion 17.11

Explain why the probability that both observers measure $+$ will be given by

$$\mathcal{P}_{++} = \left| \left({}_1\langle + | {}_{2\hat{n}}\langle + | \right) |\psi\rangle \right|^2 \quad (17.8)$$

Substituting the state $|\psi\rangle$ and $|+\rangle_{\hat{n}}$ into eq. (17.8) leads to

$$\begin{aligned}\mathcal{P}_{++} &= \left| {}_1\langle + | \left(\cos \frac{\theta}{2} {}_2\langle + | + e^{-i\phi} \sin \frac{\theta}{2} {}_2\langle - | \right) \frac{1}{\sqrt{2}} (|+\rangle_1 |-\rangle_2 - |-\rangle_1 |+\rangle_2) \right|^2 \\ &= \left| \left(\cos \frac{\theta}{2} {}_2\langle + | + e^{-i\phi} \sin \frac{\theta}{2} {}_2\langle - | \right) \frac{1}{\sqrt{2}} (|-\rangle_2) \right|^2 \\ &= \frac{1}{2} \sin^2 \frac{\theta}{2}\end{aligned}\quad (17.9)$$

💬 Discussion 17.12

Explain the steps in eq. (17.9).

❖ Discussion 17.13

Show that, for this particular measurement we also get,
 $\mathcal{P}_{--} = \frac{1}{2} \sin^2 \frac{\theta}{2}$, so that

$$\mathcal{P}_{\text{same}} = \mathcal{P}_{++} + \mathcal{P}_{--} = \sin^2 \frac{\theta}{2} \quad (17.10)$$

The chance that B observes “–” along \hat{n} when A measures “+” is

$$\begin{aligned} \mathcal{P}_{+-} &= \left| \left({}_1\langle + | {}_{2\hat{n}}\langle - \right) |\psi\rangle \right|^2 \\ &= \left| {}_1\langle + | \left(\sin \frac{\theta}{2} {}_2\langle + | - e^{-i\phi} \cos \frac{\theta}{2} {}_2\langle - \right) \frac{1}{\sqrt{2}} (|+\rangle_1 |-\rangle_2 - |-\rangle_1 |+\rangle_2) \right|^2 \\ &= \left| \left(\sin \frac{\theta}{2} {}_2\langle + | - e^{-i\phi} \cos \frac{\theta}{2} {}_2\langle - \right) \frac{1}{\sqrt{2}} (|-\rangle_2) \right|^2 \\ &= \frac{1}{2} \cos^2 \frac{\theta}{2} \end{aligned} \quad (17.11)$$

❖ Discussion 17.14

Show that, for this particular measurement we also get,
 $\mathcal{P}_{-+} = \frac{1}{2} \cos^2 \frac{\theta}{2}$, so that

$$\mathcal{P}_{\text{opp}} = \mathcal{P}_{+-} + \mathcal{P}_{-+} = \cos^2 \frac{\theta}{2}. \quad (17.12)$$

Now overall, the angle θ between the measurement axes of A and B is 0° for $1/3$ of measurements and 120° for $2/3$ of the measurements. This means that the average probabilities are

$$\begin{aligned} \mathcal{P}_{\text{same}} &= \frac{1}{3} \times \sin^2 \frac{0^\circ}{2} + \frac{2}{3} \times \sin^2 \frac{120^\circ}{2} = \frac{1}{2} \\ \mathcal{P}_{\text{opp}} &= \frac{1}{3} \times \cos^2 \frac{0^\circ}{2} + \frac{2}{3} \times \cos^2 \frac{120^\circ}{2} = \frac{1}{2} \end{aligned} \quad (17.13)$$

❖ Discussion 17.15

Check the calculations in eq. (17.13).

We see that the predictions of quantum mechanics, eq. (17.13) are inconsistent with what we found for local hidden variable theories, eq. (17.7). As these probabilities can be measured, we can test whether local hidden variable theories are possible. The experiments preformed so far have always agreed with quantum mechanics meaning we can not have local hidden variable theories⁶.

Entangled states like eq. (17.1) seem to have genuine non-local behaviour at variance with a classical intuition. The good news is that researchers are learning how to exploit the properties of entangled states to engineer new quantum technologies like quantum computers and quantum cryptography.

⁶ There is a loop hole called superdeterminism. In our discussion of Bell's inequality, we assumed that each observer is free to randomly select how their analyser is oriented for each measurement. If this is not the case and there are additional hidden variables affecting how observers choose their axes, locality can be preserved.

Schrödinger Cat Paradox

The Schrödinger cat paradox is a thought experiment⁷, developed by Schrödinger to dramatically demonstrate some of the problems of quantum measurement. It brings into sharp focus the question of where the quantum world ends and the classical world begins. The set-up of the experiment is shown in fig. 17.3. The apparatus for the experiment includes a radioactive nucleus, a Geiger counter, a hammer, a bottle of cyanide, a box and of course a cat. Things are set up such that the nucleus has a 50% chance of decaying over an hour after which we will open the box. If the Geiger counter detects a decay, it causes the hammer to fall, breaking the bottle of cyanide and killing the cat. This means that there is a 50% chance for the cat to die.

Just before we open the box, the nucleus will be in a superposition of decayed and undecayed states:

$$|\psi_{\text{nucleus}}\rangle = \frac{1}{\sqrt{2}}(|\psi_{\text{decayed}}\rangle + |\psi_{\text{undecayed}}\rangle) \quad (17.14)$$

The essential idea of the experiment is that we have set up a relationship between a microscopic quantum superposition eq. (17.14) and a macroscopic object – the cat. If the nucleus decayed, the cat will have died and if the nucleus didn't decay, the cat will be alive. Since the cat is made up of atoms, which are described by quantum mechanics, the cat should be describable by a (very complicated) quantum state. Furthermore, since the cat's state is dependent on eq. (17.14), just before we open the box, we also expect the cat to be in a superposition state:

$$|\psi_{\text{cat}}\rangle = \frac{1}{\sqrt{2}}(|\psi_{\text{dead}}\rangle + |\psi_{\text{alive}}\rangle) \quad (17.15)$$

Both quantum mechanically, and classically, we expect a 50/50 chance for the cat to be alive in the end. The key difference, is that quantum mechanics suggests that the cat is not dead or alive until we open the box and its wave function collapses to either $|\psi_{\text{dead}}\rangle$ or $|\psi_{\text{alive}}\rangle$. This goes against our classical intuition which tells us that the cat is either dead or alive – we are just ignorant of its state until we open the box.

What is so paradoxical about a cat in a superposition state? We have been talking about electrons in superposition states for quite a while now – why should cats be different? We can get a clue from experiment 4 in chapter 2 – remember that we saw an interference effect relying on a particular phase relationship between two components of a superposition state for a spin- $\frac{1}{2}$ particle. Interference effects for cats which consist of more than 10^{23} atoms is a very different proposition.

The important questions raised by Schrödinger's thought experiment are

1. Can macroscopic objects, like cats, be described quantum mechanically?

⁷ No cats are actually harmed.

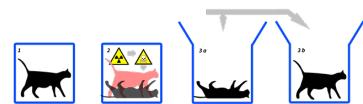


Figure 17.3: Schrödinger cat thought experiment

Image Attribution: Master of the Universe 322, CC BY-SA 4.0, via Wikimedia Commons. [Image Source](#).

2. What is the cause of wave function collapse?

In the Copenhagen interpretation, there is some boundary between the quantum and classical systems. Microscopic systems are governed by quantum mechanics and macroscopic systems follow classical rules. Exactly where the transition between classical and quantum occurs is not spelt out. In the Copenhagen interpretation, the act of measurement is what causes a quantum state to collapse *but* the actual mechanism for how collapse occurs is not described. It has been suggested that human consciousness plays a role in wave function collapse or even that the wave function doesn't collapse but instead the universe splits into alternate versions for each outcome. As they are apparently untestable experimentally, such ideas are more metaphysical and philosophical than the purview of physicists.

While the interpretation of quantum mechanics is confusing from our classical perspective (is the cat either dead or alive?), the mathematical formalism of quantum mechanics is not – we can do rigorous and precise calculations which agree with experimental observations (if we repeat the experiment many times, there will be a 50/50 of finding the cat dead or alive when we open the box).

While the nature of measurement in quantum mechanics remains mysterious, we are starting to learn more about the boundary between the quantum and classical descriptions through experiment. A key issue in the Schrödinger cat paradox is whether nucleus-cat system is described by the entangled state,

$$|\psi_{\text{system}}\rangle = \frac{1}{\sqrt{2}}(|\psi_{\text{dead}}\rangle|\psi_{\text{decayed}}\rangle + |\psi_{\text{alive}}\rangle|\psi_{\text{undecayed}}\rangle). \quad (17.16)$$

In other words do we have a coherent superposition like eq. (17.16) or is there really a 50/50 statistical mixed state of the two possibilities, $|\psi_{\text{dead}}\rangle|\psi_{\text{decayed}}\rangle$, and, $|\psi_{\text{alive}}\rangle|\psi_{\text{undecayed}}\rangle$. We could differentiate between the two cases by seeing if there are inference effects between the two states.

We can view the cat as a macroscopic measuring device which tells us whether the nucleus has decayed. Experimentalists model a Schrödinger cat-like experiment using an atom and an electric field in a cavity. The atom effectively has two states, a ground state, $|g\rangle$, and an excited state $|e\rangle$, analogous to the decayed or undecayed state of the nucleus. The electric field allows us to measure the state of the atom playing the role of a classical measuring device. The electric field in the cavity is engineered to be in a coherent superposition of photons described by the state $|\alpha\rangle$. The magnitude of the complex number, α , is equal to the square root of the average number of photons in the cavity. For $|\alpha|$ large, the state is equivalent to a classical electromagnetic field whereas for $|\alpha|$ small, the field is more quantum. The interesting feature of this experiment is that we can vary α to try study the transition between quantum and classical descriptions.

Figure 17.4 gives an overview of the Schrödinger cat-like exper-

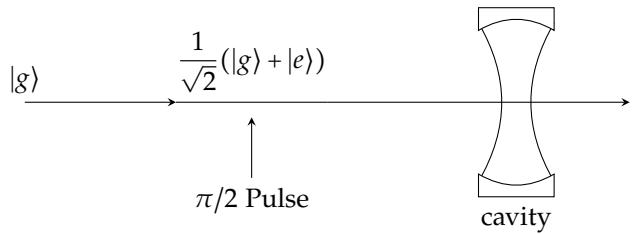


Figure 17.4: A Schrödinger cat-like experiment with an atom in a cavity

iment. Atoms are sent through a cavity and interact with the cavity's electromagnetic field. Before they enter the cavity, the atoms are subjected to a $\pi/2$ -pulse which puts them in a superposition state

$$|\psi_{\text{atom}}\rangle = \frac{1}{\sqrt{2}}(|e\rangle + |g\rangle) . \quad (17.17)$$

As atoms pass through the cavity they produce a phase shift in the electric field. The experiment is engineered so that, $|g\rangle$, and, $|e\rangle$, generate opposite phase shifts $\pm\phi$. This means that after the atom passes through the cavity, the atom-cavity system is in the entangled state

$$|\psi_{\text{atom-cavity}}\rangle = \frac{1}{\sqrt{2}}(|e\rangle |\alpha e^{+i\phi}\rangle + |g\rangle |\alpha e^{-i\phi}\rangle) , \quad (17.18)$$

which is analogous to eq. (17.16). We can then measure the state of the cavity field by sending in a second atom and looking for interference effects. As we repeat the experiment with increasing numbers of photons in the cavity, the interference effects disappear. This demonstrates that the fixed phase relationship between the two parts of the entangled state, eq. (17.18), has disappeared. We say that **decoherence** has occurred. Decoherence is due to interaction of the system with random environmental factors. The experiment shows how quantum coherence is lost when a state becomes sufficiently complicated so that it is effectively classical. Details of the experiment can be found in [Brune et al. 1996].

18

Spectroscopy & the Energy Eigenvalue equation

18.1 Spectroscopy

This reading is based on [McIntyre et al. 2012, Section 5.1-2]

So far we've used spin systems to learn about basic features of quantum mechanics. Now we want to understand how to explain the microscopic world, namely nuclei, atoms, molecules, and the properties of solids that **can not** be explained with classical physics. For example, why are sodium lamps yellow? How do lasers work? Why is uranium radioactive? A key to understanding the microscopic quantum system is to determine the energy states the system is allowed to have. A quantum system has a unique set of allowed energies – its energy “fingerprint” which quantum mechanics allows us to predict. The experimental techniques of measuring these fingerprints or spectra is called **spectroscopy**. Historically the spectrum of Hydrogen played an important role in development of quantum mechanics.

Our goal for now is to learn how to predict these energy spectra. In this chapter we will study a **very simple** model which is not realistic but has most of the important features shared by all microscopic systems.

Quantized energies commonly arise when particles interact in such a way that their spacial movement is limited forming a composite system. Such a system is called a **bound state**, examples of which are nuclei, atoms, molecules and solids. Such **bound states** are characterised by distinct spectral lines as a result of quantized energy states. These states correspond to eigenstates of the Hamiltonian with discrete energy eigenvalues. For example, the hydrogen atom energy levels are shown in Figure 18.1.

Spectral lines appear when electrons make transitions between energy levels. The downward transitions produce the **emission spectrum** while the upward transitions produce the **absorption spectrum**. For each pair of energies, E_i and E_j , there is a possible spectral line with photon energy $|E_i - E_j|$. The frequency of this photon is $f_{ij} = \frac{|E_i - E_j|}{h}$, and its wave-length is $\lambda_{ij} = c/f_{ij}$.

Hydrogen spectrum

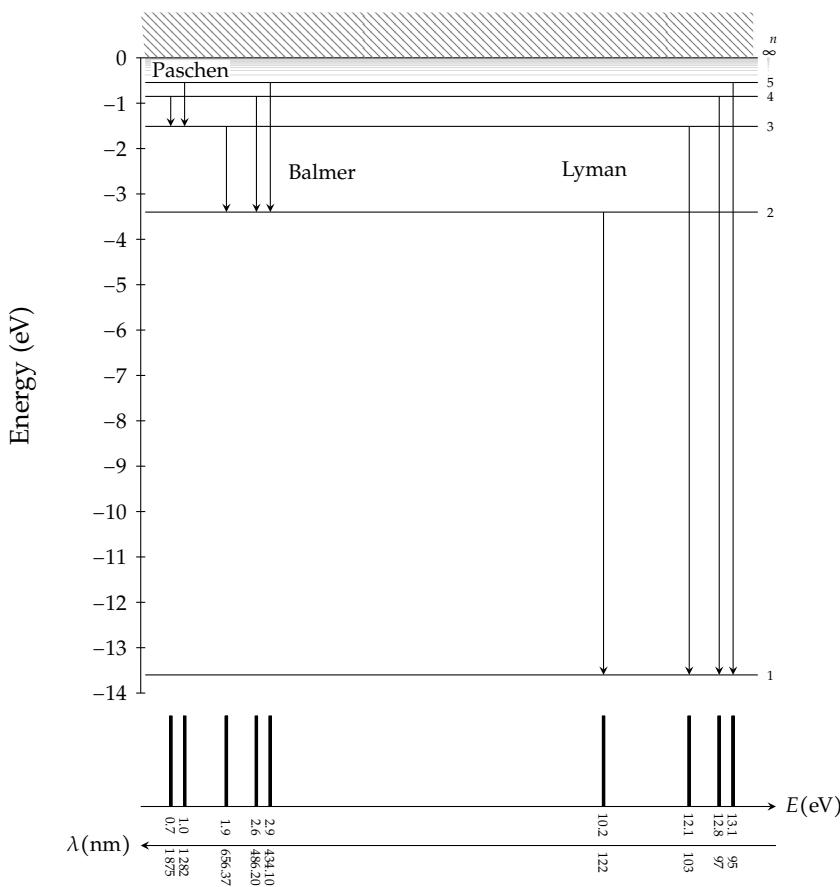


Figure 18.1: Hydrogen energy levels and the corresponding optical spectrum as a function of energy and wavelength (the wavelength scale is not a linear scale).

The spectral lines of hydrogen atoms sharing a common lowest level form a series as shown in Figure 18.1 and Table 5.1. Each series is named after its discoverer. For example, the spectral lines of the Balmer series are shown in Figure 18.2.



Figure 18.2: Visible emission spectrum lines in the Balmer series. Image by Adrignola: source

The lowest energy state is called the **ground state** and the levels above are called **excited states**. Please note that the word spectrum can refer to observed optical lines or a set of quantised energy states. A set of quantised energy states is also called an **energy spectrum**.

Spectroscopy is the measurement of the energy of a quantum state as schematically illustrated in Figure 18.3. In a real experiment we measure differences between energy levels and it takes a bit of work to determine the energy levels themselves. We'll assume this work has been done.

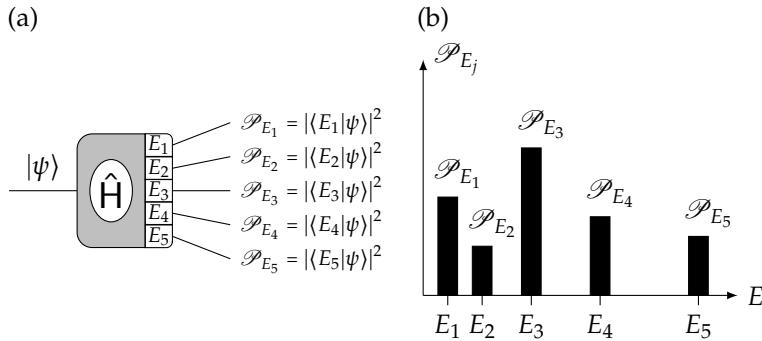


Figure 18.3: (a) Energy measurement and (b) histogram of results.

Energy Eigenvalue equation

For a given Hamiltonian operator, \hat{H} , the energies E_i and eigenstates $|E_i\rangle$ are solutions to the **eigenvalue equation**¹

$$\hat{H}|E_i\rangle = E_i|E_i\rangle. \quad (18.1)$$

¹ Also called, for reasons we will see later, the Time-independent Schrödinger equation.

For now, we will look at predicting the allowed energies, E_i , given some Hamiltonian, H . We will start looking at simple one dimensional systems. Now, the mechanical energy of classical particle in one dimension is

$$E = \text{Kinetic energy} + \text{Potential energy} = \frac{p_x^2}{2m} + V(x). \quad (18.2)$$

Discussion 18.1

Show $\frac{p_x^2}{2m} = \frac{1}{2}mv_x^2$.

From (18.2), we can write down the Quantum Mechanical Hamiltonian operator for a one dimensional particle as

$$\hat{H} = \frac{\hat{p}_x^2}{2m} + V(\hat{x}). \quad (18.3)$$

To solve the energy eigenvalue equation, (18.1), for this Hamiltonian operator, (18.3), we need to know what these new operators \hat{p}_x and \hat{x} are. It turns out that \hat{p}_x and \hat{x} are quite different from spin operators and we need to develop some more mathematical machinery before we can get any further. For spin we either used, abstract kets, for example $|+\rangle$ or $|-\rangle_x$, or we represented our kets as column vectors in particular basis, for example, in the $|\pm\rangle$ basis,

$$|+\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |-\rangle_x \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (18.4)$$

Now, most quantum mechanical problems are solved using a particular representation. For the problems we will be looking at now, it is useful to represent states with functions of position. These functions are called **wave functions**.

Wave functions represent quantum states. We will use, $\psi = \psi(x)$, to represent wave functions. Since wave functions correspond to a

particular representation, we write

$$|\psi\rangle \doteq \psi(x) . \quad (18.5)$$

This is called the **position representation**.

Notation: energy eigenstates

We will use ψ for general states and other Greek letters for specific eigenstates. For example, we will use φ for energy eigenstates:

$$|E_i\rangle \doteq \varphi_{E_i}(x) . \quad (18.6)$$

Using this notation, the energy eigenvalue equation in position representation is

$$\hat{H}\varphi_{E_i}(x) = E_i\varphi_{E_i}(x) . \quad (18.7)$$

To solve, (18.7), we need to know how to represent \hat{H} in the position representation.

It turns out that, in the position representation, \hat{x} is represented by multiplication by the variable x and \hat{p} is represented by a derivative with respect to position

$$\hat{x} \doteq x \quad (18.8)$$

$$\hat{p} \doteq -i\hbar \frac{d}{dx} \quad (18.9)$$

We will take (18.8) and (18.9) as *postulates*. For now we will also use \hat{p} as short hand for \hat{p}_x . You may wonder about the presence of $i\hbar$ in (18.9). We need the \hbar to get the dimensions right and the i to ensure the eigenvalues of \hat{p} are real.

💬 Discussion 18.2

Show that, $-i\hbar \frac{d}{dx}$, has the same dimensions as momentum.

💬 Discussion 18.3

Show that $e^{i\lambda x}$ is an eigenfunction of \hat{p} . What is its eigenvalue?

We can now write down the energy eigenvalue equation in the position basis:

$$\begin{aligned} \hat{H}\varphi_{E_i} &= E_i\varphi_{E_i} \\ \implies \left(\frac{\hat{p}^2}{2m} + V(\hat{x}) \right) \varphi_{E_i} &= E_i\varphi_{E_i} \\ \implies \left(\frac{1}{2m} \left(-i\hbar \frac{d}{dx} \right)^2 + V(x) \right) \varphi_{E_i} &= E_i\varphi_{E_i} , \end{aligned} \quad (18.10)$$

which leads to

Position space 1D energy eigenvalue equation

$$\left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right) \varphi_{E_i} = E_i \varphi_{E_i} \quad (18.11)$$

In the position basis, operator equations often become differential equations and, in this case, we need to solve differential equations to find allowed energies. We will solve (18.11) for several potentials, $V(x)$.

 **Discussion 18.4**

Show that $\sin(\alpha x)$ is an eigenfunction of \hat{H} when $V(x) = 0$. What is its eigenvalue? What dimensions does α have?

 **Discussion 18.5**

Write a summary of the important points and formulae in this chapter.

19

The wave function

How does the wave function, $\psi(x)$, introduced in the last section, relate to the state $|\psi\rangle$? To answer this question let's review spin- $\frac{1}{2}$ system. It was useful to represent a state $|\psi\rangle$ as a column. So for example, in the S_z representation:

$$|\psi\rangle = \begin{pmatrix} \langle +|\psi \rangle \\ \langle -|\psi \rangle \end{pmatrix} \quad \begin{array}{l} \leftarrow S_z = \hbar/2 \\ \leftarrow S_z = -\hbar/2 \end{array} \quad (19.1)$$

The (complex) numbers, $\langle \pm |\psi \rangle$, are the projections of $|\psi\rangle$ onto the S_z eigenstates. If we measure S_z , as shown in Figure 19.1, we can use, $\langle \pm |\psi \rangle$, to find probabilities:

$$\mathcal{P}_\pm = |\langle \pm |\psi \rangle|^2 \quad (19.2)$$

This reading is based on [McIntyre et al. 2012, Section 5.3]

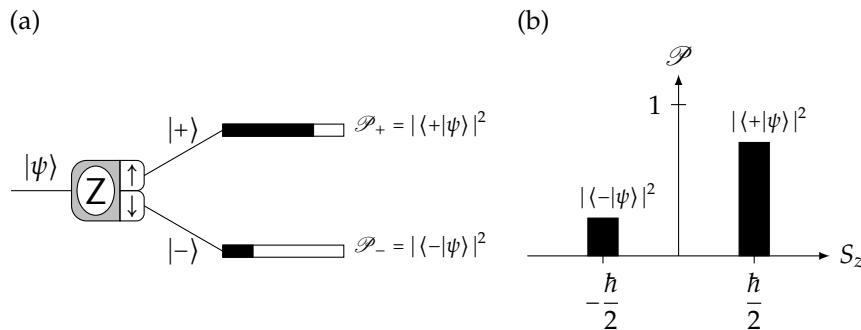


Figure 19.1: (a) Spin measurement and (b) probability histogram.

If we measure the energy of some quantum system, as shown in Figure 19.2, then a good basis to use is the energy eigenstate basis:

$$|\psi\rangle = \begin{pmatrix} \langle E_1|\psi \rangle \\ \langle E_2|\psi \rangle \\ \langle E_3|\psi \rangle \\ \vdots \end{pmatrix} \quad \begin{array}{l} \leftarrow E = E_1 \\ \leftarrow E = E_2 \\ \leftarrow E = E_3 \\ \vdots \end{array} \quad (19.3)$$

The probability for measuring a particular energy, E_i , is given by

$$\mathcal{P}_{E_i} = |\langle E_i|\psi \rangle|^2 \quad (19.4)$$

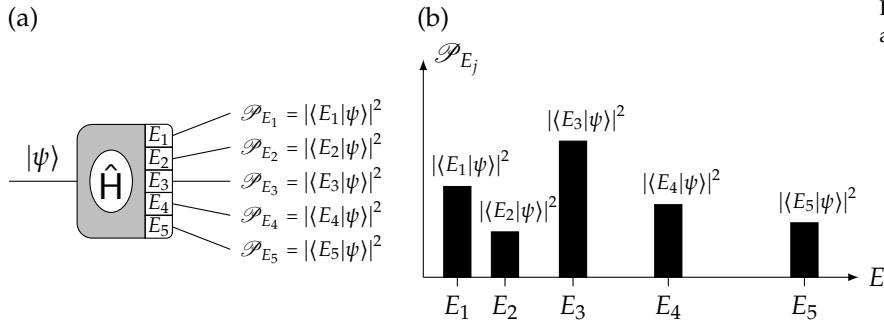


Figure 19.2: (a) Energy measurement and (b) histogram of results.

Discussion 19.1

Consider a stationary electron in uniform magnetic field $\vec{B} = B_0 \hat{i}$. Suppose the electron is in the state $|\psi\rangle = |+\rangle$. What does $|\psi\rangle$ look like in the energy eigenstate basis? What does \hat{S}_x look like expressed in the energy eigenstate basis?

By analogy, we could try to represent our particle using the eigenstates of the position operator, \hat{x} . Let, $|x_i\rangle$, denote an eigenstate¹ of the position operator with eigenvalue x_i . Then the generalisation of (19.1) and (19.3) would be:

$$|\psi\rangle \doteq \begin{pmatrix} \langle x_1 | \psi \rangle \\ \langle x_2 | \psi \rangle \\ \langle x_3 | \psi \rangle \\ \vdots \end{pmatrix} \quad \leftarrow x_1 \\ \leftarrow x_2 \\ \leftarrow x_3 \\ \vdots \quad , \quad (19.5)$$

where, the components $\langle x_i | \psi \rangle$, are the amplitudes for the probability for finding a particle at the point, x_i , given by $\mathcal{P}_{x_i} = |\langle x_i | \psi \rangle|^2$. A graphical representation of (19.5) is shown in Figure 19.3 (a).

¹ On the assumption that such an eigenstate exists.

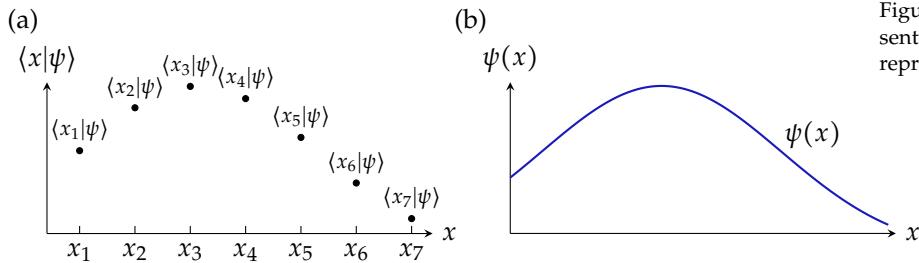


Figure 19.3: (a) Discrete basis representation and (b) continuous basis representation.

There is one major flaw with the approach presented in (19.5): experiment tells us that the physical observable x is *not* quantised. It is found that all values of x are possible. This is in contrast to say, spin- $\frac{1}{2}$ systems, where only two results are possible for the components of spin. We say the spectrum of spin- $\frac{1}{2}$ eigenvalues is **discrete**. All values of position are allowed and we say that the spectrum of eigenvalues of position is **continuous**. This means that (19.5) is not convenient representation for position – we can not write down an infinite number of components. Even with “only” 100 components, the column representation would be cumbersome.

Instead we might use a graph, as in Figure 19.3 (a), to record these 100 numbers as points. As we increase the number of points, these points would approach a curve. Indeed, if we have an infinite continuous set of position amplitudes, $\langle x|\psi \rangle$, we can represent them as a curve as in Figure 19.3 (b). This curve is what we call the wave function $\psi(x)$. We can write the wave function as $\psi(x)$ or $\langle x|\psi \rangle$:

Wavefunction

$$\psi(x) = \langle x|\psi \rangle \quad (19.6)$$

The wave function, $\psi(x)$, is the **probability amplitude** for the quantum state, $|\psi\rangle$, to be measured in the position eigenstate, $|x\rangle$, basis. We will assume $\psi(x)$ is continuous. Later, we will look more closely at position eigenstates.

Now, comparing the wave function, with spin amplitudes, say $\langle \pm|\psi \rangle$, suggests that the probability of measuring a particular value of position is given by taking the modulus squared of $\psi(x) = \langle x|\psi \rangle$. The modulus squared of $\psi(x)$, namely $\mathcal{P}(x) = |\psi(x)|^2$, is a probability *function*. This is as opposed to, say a set of discrete probabilities, $\{\mathcal{P}_{S_z=+\hbar/2}, \mathcal{P}_{S_z=-\hbar/2}\}$. Actually $\mathcal{P}(x)$ is more accurately referred to as a probability *density* but more about that later. In any case given a wave function $\psi(x)$, we can calculate the probability function $\mathcal{P}(x)$:

Probability function

$$\mathcal{P}(x) = |\psi(x)|^2 \quad (19.7)$$

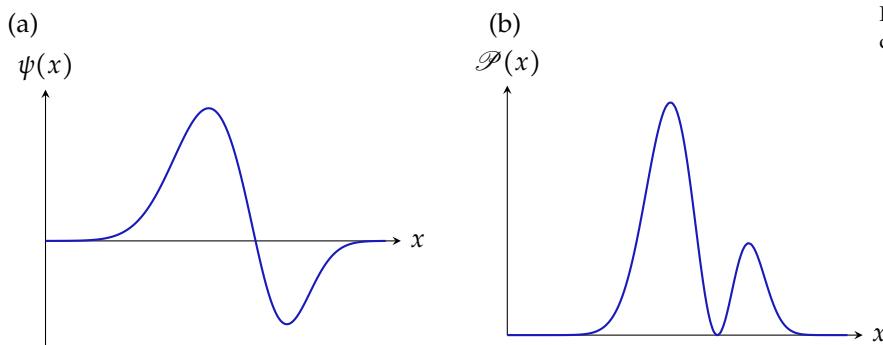


Figure 19.4: (a) Wave function and (b) corresponding probability density.

$\mathcal{P}(x)$, as for example plotted in Figure 19.4, is analogous to the histograms of discrete probability we've seen before (as in Figure 19.2 (b)).

In quantum mechanics, the sum of all probabilities must be 1 or in other words the state vector must be normalised. For the discrete spin- $1/2$ case, we require:

$$\sum_{\pm} \mathcal{P}_{\pm} = \sum_{\pm} |(\pm|\psi)\|^2 = 1 , \quad (19.8)$$

and similarly, if position was discrete we would need

$$\sum_n \mathcal{P}_{x_n} = \sum_n |\langle x_n | \psi \rangle|^2 = 1 \quad (19.9)$$

However, since the spectrum of position eigenvalues is continuous, we need an integral rather than a sum. This means that in one-dimension we require:

$$\int_{-\infty}^{\infty} \mathcal{P}(x) dx = \int_{-\infty}^{\infty} |\psi(x)|^2 dx = 1 \quad (19.10)$$

Example 19.1:

Problem Normalise the wavefunction

$$\psi(x) = Ne^{-(x-x_0)^2/2\alpha^2}. \quad (19.11)$$

You may find the integral $\int_{-\infty}^{\infty} e^{-u^2} du = \sqrt{\pi}$ useful.

Solution From (19.10), we require

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} |\psi(x)|^2 dx \\ &= \int_{-\infty}^{\infty} |N|^2 e^{-(x-x_0)^2/\alpha^2} dx \quad \left(\text{let } u = \frac{x-x_0}{\alpha} \implies dx = \alpha du \right) \\ &= \alpha |N|^2 \int_{-\infty}^{\infty} e^{-u^2} du \\ &= \alpha |N|^2 \sqrt{\pi} \end{aligned} \quad (19.12)$$

We take N to be real and positive so that

$$N = \left(\frac{1}{\alpha^2 \pi} \right)^{1/4} \quad (19.13)$$

Discussion 19.2

Explain the steps in (19.12). Explain why we can take N real and positive.

Notice that we've had to add a differential, dx , in the integral, (19.10). The differential has dimensions of length but the total integrated probability is dimensionless which implies $\mathcal{P}(x)$ must have dimensions of $(\text{length})^{-1}$. This means that $\mathcal{P}(x)$ is a probability density and we interpret, $\mathcal{P}(x)dx$, as the infinitesimal probability that a particle will be found between x and $x + dx$. This is shown in Figure 19.5 (a). To calculate the probability that a particle is found in a finite interval $a < x < b$, we need to add up all the probabilities in that interval:

$\mathcal{P}_{a < x < b}$

$$\mathcal{P}_{a < x < b} = \int_a^b |\psi(x)|^2 dx \quad (19.14)$$

Equation (19.14) is illustrated in Figure 19.5 (b).

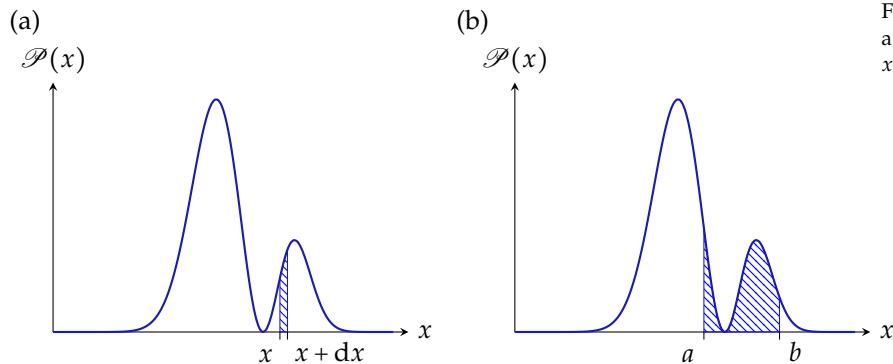


Figure 19.5: Probability for measuring a particle to be in the position range (a) x to $x + dx$, and (b) a to b .

Discussion 19.3

What are the dimensions of $\psi(x)$?

Discussion 19.4

Why do you think we use the word **density** when we refer to $\mathcal{P}(x)$ as a probability density?

Discussion 19.5

We say that a function, $f(x)$, is even if $f(x) = f(-x)$. Show that for an even function

$$\int_{-a}^a f(x) dx = 2 \int_0^a f(x) dx . \quad (19.15)$$

Example 19.2:

Problem Given the wave function, (19.11) with N as in (19.13), what is the probability the particle will be found in the region $x > x_0$?

Solution

$$\begin{aligned}
 \mathcal{P}_{x>x_0} &= \int_{x_0}^{\infty} |\psi(x)|^2 dx \\
 &= \frac{1}{\alpha\sqrt{\pi}} \int_{x_0}^{\infty} e^{-(x-x_0)^2/\alpha^2} dx \quad \left(\text{let } u = \frac{x-x_0}{\alpha}\right) \\
 &= \frac{1}{\sqrt{\pi}} \int_0^{\infty} e^{-u^2} du \\
 &= \frac{1}{\sqrt{\pi}} \frac{1}{2} \int_{-\infty}^{\infty} e^{-u^2} du \\
 &= \frac{1}{2}
 \end{aligned} \tag{19.16}$$

Looking at Figure 19.6, can you explain this result?

💬 Discussion 19.6

Can you explain the steps in (19.16)?

Translating wave functions to kets

We now need to translate bra-ket rules to wave function rules to make predictions when we are working in the position basis.

Comparing

$$\langle \psi | \psi \rangle = 1 \tag{19.17}$$

and

$$\int_{-\infty}^{\infty} |\psi(x)|^2 dx = \int_{-\infty}^{\infty} \psi^*(x)\psi(x) dx = 1 \tag{19.18}$$

leads us to postulate the following dictionary between bra-kets and wave functions:

Words	Symbols
1. Replace ket with wave function	$ \psi\rangle \rightarrow \psi(x)$
2. Replace bra with wave function conjugate	$\langle \psi \rightarrow \psi^*(x)$
3. Replace braket with wave integral over all space	$\langle \cdot \cdot \rangle \rightarrow \int_{-\infty}^{\infty} \cdot \cdot dx$
4. Replace operator with position representation	$\hat{A}(\hat{p}, \hat{x}) \rightarrow A\left(-i\hbar \frac{d}{dx}, x\right)$

We have added an additional rule 4. The reason for rule 4 will become clear soon.

Notation

The notation $\langle \cdot | \cdot \rangle$ represents the inner product abstractly. The “dots” indicate space for states we can input into the inner product. Similarly, the notation, $\int_{-\infty}^{\infty} \cdot \cdot dx$, means we can input two functions into the integral.

❖ Discussion 19.7

Explain how we can read off rules 1-3 in our bra-ket wavefunction dictionary when comparing (19.17) and (19.18).

Probability amplitudes

Once we have normalised our wave function, we can use it to calculate probability amplitudes. Using our translation rules, a probability amplitude, in terms of wave functions is

$$\langle \phi | \psi \rangle = \int_{-\infty}^{\infty} \phi^*(x) \psi(x) dx \quad (19.19)$$

The square of the amplitude, $\langle \phi | \psi \rangle$, is then the probability the state $\psi(x)$ is found to be in the state $\phi(x)$:

$$\mathcal{P}_{\psi \rightarrow \phi} = |\langle \phi | \psi \rangle|^2 = \left| \int_{-\infty}^{\infty} \phi^*(x) \psi(x) dx \right|^2 \quad (19.20)$$

This is the probability that a system prepared in a state $|\psi\rangle$ is measured to have the observable for which $|\phi\rangle$ is an eigenstate. In practise we measure observables *not* states. For example, the probability of measuring the energy to be E_n is

$$\mathcal{P}_{E_n} = |\langle E_n | \psi \rangle|^2 = \left| \int_{-\infty}^{\infty} \varphi_n^*(x) \psi(x) dx \right|^2, \quad (19.21)$$

where $\varphi_n(x)$ is the energy eigenstate with energy² E_n .

² For now, we are assuming that there is only one state with energy E_n .

Compare & contrast

We have looked at two different ways of calculating different types of probabilities. Firstly,

$$\mathcal{P}_{a < x < b} = \int_a^b |\psi(x)|^2 dx, \quad (19.22)$$

sums up the probabilities of measuring a particle at many different positions in the range (a, b) . On the other hand

$$\mathcal{P}_{E_n} = |\langle E_n | \psi \rangle|^2 = \left| \int_{-\infty}^{\infty} \varphi_n^*(x) \psi(x) dx \right|^2, \quad (19.23)$$

gives us the probability of measuring a particular energy eigenvalue, E_n , corresponding to the eigenstate φ_n .

Expectation values

To calculate expectation values, in the position representation, we need to deal with operators. To this end, we rewrite an expectation value, $\langle \hat{A} \rangle = \langle \psi | \hat{A} | \psi \rangle$, as

$$\langle \hat{A} \rangle = \langle \psi | (\hat{A} | \psi \rangle). \quad (19.24)$$

Equation (19.24) tells us that we can view an expectation value as the inner product between two states, $|\psi\rangle$, and, $|\phi\rangle = (\hat{A} | \psi \rangle)$. This in

turn tells us that, when using position space representation for $|\psi\rangle$, we should use position space representation for \hat{A} , which is the reason for rule 4 in our bra-ket wavefunction dictionary.

Discussion 19.8

We say a function $f(x)$ is odd if $f(-x) = -f(x)$. Show that for an odd function,

$$\int_{-a}^a f(x) dx = 0. \quad (19.25)$$

Example 19.3: Expectation value

Problem Given the wave function, (19.11) with N as in (19.13), find $\langle \hat{x} \rangle$.

Solution

$$\begin{aligned} \langle \hat{x} \rangle &= \langle \psi | \hat{x} | \psi \rangle \\ &= \int_{-\infty}^{\infty} \left(\left(\frac{1}{\alpha^2 \pi} \right)^{1/4} e^{-(x-x_0)^2/2\alpha^2} \right)^* x \left(\left(\frac{1}{\alpha^2 \pi} \right)^{1/4} e^{-(x-x_0)^2/2\alpha^2} \right) dx, \\ &= \frac{1}{\alpha \sqrt{\pi}} \int_{-\infty}^{\infty} x e^{-(x-x_0)^2/\alpha^2} dx, \quad (\text{let } u = \frac{x-x_0}{\alpha} \implies x = \alpha u + x_0) \\ &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (\alpha u + x_0) e^{-u^2} du, \\ &= \frac{1}{\sqrt{\pi}} \left(x_0 \int_{-\infty}^{\infty} e^{-u^2} du + \alpha \int_{-\infty}^{\infty} u e^{-u^2} du \right) \\ &= x_0 \end{aligned} \quad (19.26)$$

This result for $\langle \hat{x} \rangle$ is not surprising if we look at a plot of $\psi(x)$ as shown in Figure 19.6 (a).

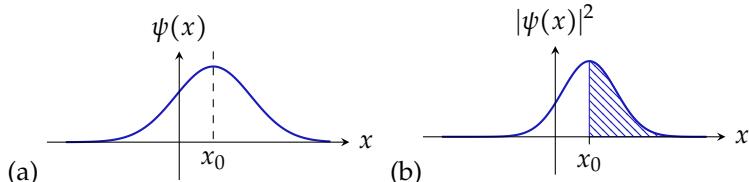


Figure 19.6: (a) Wave function and (b) corresponding probability density. The area of the hatched region (which actually extends to infinity) in (b) represents the probability for the particle to be measured in the region $x > x_0$.

Discussion 19.9

Explain how we can use the rules of our bra-ket wave function dictionary to get the second line of (19.26).

Discussion 19.10

Justify the further steps in (19.26).

Discussion 19.11

Explain why the result for $\langle \hat{x} \rangle$, in (19.26), is not surprising if we look at a plot of $\psi(x)$ as shown in Figure 19.6 (a).

Example 19.4: Uncertainty

Problem Given the wave function, (19.11) with N as in (19.13), find $\langle \hat{x}^2 \rangle$ and determine Δx . You may find the integral, $\int_{-\infty}^{\infty} u^2 e^{-u^2} du = \sqrt{\pi}/2$, useful.

Solution

$$\begin{aligned}
 \langle \hat{x}^2 \rangle &= \langle \psi | \hat{x}^2 | \psi \rangle \\
 &= \int_{-\infty}^{\infty} \left(\left(\frac{1}{\alpha^2 \pi} \right)^{1/4} e^{-(x-x_0)^2/2\alpha^2} \right)^* x^2 \left(\left(\frac{1}{\alpha^2 \pi} \right)^{1/4} e^{-(x-x_0)^2/2\alpha^2} \right) dx , \\
 &= \frac{1}{\alpha \sqrt{\pi}} \int_{-\infty}^{\infty} x^2 e^{-(x-x_0)^2/\alpha^2} dx , \quad (\text{let } u = \frac{x-x_0}{\alpha} \implies x = \alpha u + x_0) \\
 &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (\alpha u + x_0)^2 e^{-u^2} du , \\
 &= \frac{1}{\sqrt{\pi}} \left(x_0^2 \int_{-\infty}^{\infty} e^{-u^2} du + 2\alpha x_0 \int_{-\infty}^{\infty} ue^{-u^2} du + \alpha^2 \int_{-\infty}^{\infty} u^2 e^{-u^2} du \right) \\
 &= x_0^2 + \alpha^2/2
 \end{aligned} \tag{19.27}$$

The using the result form the previous example, we find

$$\Delta x = \sqrt{\langle \hat{x}^2 \rangle - \langle \hat{x} \rangle^2} = \sqrt{x_0^2 + \alpha^2/2 - x_0^2} = \alpha/\sqrt{2} . \tag{19.28}$$

From (19.28) we see that $\alpha = \sqrt{2}\Delta x$.

Discussion 19.12

As shown in Figure 19.7, the larger α is, the “wider” the probability density $|\psi|^2$ is. Explain why increasing the width of $|\psi|^2$ will increase the uncertainty Δx .

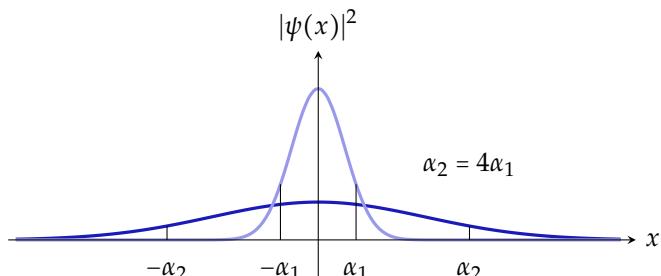


Figure 19.7: A plot of two Gaussian wave functions with different widths.

Position and momentum operators are Hermitian

Discussion 19.13

Check that $\hat{x}^\dagger = \hat{x}$, or in other words, show that $\langle \phi | \hat{x} \psi \rangle = \langle \hat{x} \phi | \psi \rangle$.

Now showing that \hat{p} is Hermitian requires a little more work but at least we'll see why we need that factor of i in the position space representation of \hat{p} . We first need to establish a basic fact about wave functions, namely that, for normalisable wave functions, we require that

$$\psi(x) \rightarrow 0 \text{ as } x \rightarrow \pm\infty. \quad (19.29)$$

To see why we need (19.29), let's suppose it is not true. Think about a plot of $|\psi|^2$ like Figure 19.8. The integral, $\int |\psi|^2 dx$, corresponds to the area under the curve. If $|\psi|^2$ does not go to zero, as we do the integral, we'll just keep adding more and more area and the integral will grow indefinitely which is inconsistent with the wave function being normalised as in (19.18). We conclude that we need (19.29) for a normalisable wave function.

Now let's show that \hat{p} is Hermitian,

$$\begin{aligned} \langle \phi | \hat{p} \psi \rangle &= \int_{-\infty}^{\infty} \phi^*(x) \left(-i\hbar \frac{d(\psi(x))}{dx} \right) dx \\ &= \underbrace{[\phi^*(x)\psi(x)]_{-\infty}^{\infty}}_{(19.29)=0} - \int_{-\infty}^{\infty} \left(-i\hbar \frac{d(\phi^*(x))}{dx} \right) \psi(x) dx \\ &= \int_{-\infty}^{\infty} \left(-i\hbar \frac{d\phi(x)}{dx} \right)^* \psi(x) dx \\ &= \langle \hat{p}\phi | \psi \rangle \end{aligned} \quad (19.30)$$

Discussion 19.14

Can you explain *all* the steps in (19.30)? Make sure you understand how the overall sign works out.

From (19.30) we conclude that $\hat{p} = \hat{p}^\dagger$.

Discussion 19.15

Write a summary of the important points and formulae in this chapter.

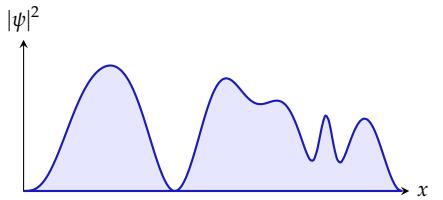


Figure 19.8: The modulus of a wave function.

20

Stationary states

What are we going to do in this section and why?

How can we know what the wave function might be? We will start learning about how to solve the Schrödinger equation in position space. We will re-derive the time-dependence of states in a slightly different way and revisit stationary states.

This reading is based on [Griffiths 2016, Section 2.1] with some changes in notation to make it more consistent with [McIntyre et al. 2012].

We've looked at how to make statistical predictions given a particular wave function but how do we get the wave function in the first place? For a given potential $V(x, t)$, one "just" needs to solve the Schrödinger equation, which in position space becomes

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + V(x, t)\psi \quad (20.1)$$

💬 Discussion 20.1

Explain how we obtain eq. (20.1) from eq. (12.1).

💬 Discussion 20.2

Suppose ψ_1 and ψ_2 are solutions to the Schrödinger equation eq. (20.1). Show that $c_1\psi_1 + c_2\psi_2$ is also a solution where c_1 and c_2 are constants.

For now we will only consider time-independent potentials $V = V(x)$ and use the technique of separation of variables to solve eq. (20.1). Let's consider solutions which can be written as a simple product of two functions which only depend on space and time respectively:

$$\psi(x, t) = \varphi(x)f(t) . \quad (20.2)$$

Solutions of the form eq. (20.2) are called **separable solutions** since we can separate out the space and time dependence.

Take note!

Equation eq. (20.2) is a strong restriction **but** such solutions turn out to be very important since we'll be able to build general solutions out of separable ones. For this to work we need our set of separable solutions

Substituting eq. (20.2) into eq. (20.1), and dividing by φf , we obtain

$$i\hbar \frac{1}{f} \frac{df}{dt} = -\frac{\hbar^2}{2m} \frac{1}{\varphi} \frac{d^2\varphi}{dx^2} + V(x) \quad (20.3)$$

Discussion 20.3

Show how we obtain eq. (20.3) after substituting eq. (20.2) into eq. (20.1). Why have we gone from partial to ordinary derivatives?

Now, the left hand side of eq. (20.3) is a function only of t while the right hand side of eq. (20.3) is a function only of x . This is only possible if both sides of the equation are constant. We give this constant a special name – the separation constant. For reasons that will be clear later, we will call the separation constant in this case E . We can now write eq. (20.3) as two separate equations

$$i\hbar \frac{1}{f} \frac{df}{dt} = E \quad (20.4)$$

$$-\frac{\hbar^2}{2m} \frac{d^2\varphi}{dx^2} + V\varphi = E\varphi \quad (20.5)$$

Discussion 20.4

Show how we can obtain eq. (20.4) and eq. (20.5) from eq. (20.3).

Multiplying eq. (20.4) by $\frac{-i}{\hbar} dt$ and integrating we obtain

$$\ln f = \frac{-i}{\hbar} Et + A, \quad (20.6)$$

where A is some constant. We can then simplify eq. (20.6) to

$$f = A' e^{-\frac{iEt}{\hbar}}, \quad (20.7)$$

where A' is a constant related to A .

Discussion 20.5

Show how eq. (20.7) follows from eq. (20.6) and determine the relationship between A and A' .

We can absorb the constant A' into $\varphi(x)$ and take

$$f(t) = e^{-\frac{iEt}{\hbar}}. \quad (20.8)$$

Discussion 20.6

Explain the statement that we can absorb A' in eq. (20.7) into $\varphi(x)$.

So, we can rewrite the ansatz eq. (20.2) as

$$\psi(x, t) = \varphi(x) e^{-\frac{iEt}{\hbar}} . \quad (20.9)$$

To find the position dependant part, $\varphi(x)$, we need to solve eq. (20.5) what you should recognise and the energy eigenvalue equation. Equation eq. (20.5) is also called the time independent Schrödinger equation(TISE). For the next few chapters we will concentrate on solving it for various potentials.

Solutions of the form eq. (20.9), have several properties which makes them very useful, which are enumerated below.

1. They are **stationary states** which means that expectation values of observables for these states are time-independent.

While the wave function eq. (20.9) depends on time the probability density associated with it, $|\psi(x, t)|^2$, does not

$$|\psi(x, t)|^2 = \varphi(x) e^{-\frac{iEt}{\hbar}} \varphi^*(x) e^{+\frac{iEt}{\hbar}} = |\varphi(x)|^2 , \quad (20.10)$$

and, as advertised above, the expectation values of any dynamical variable¹, $Q(x, p)$, is constant in time:

$$\begin{aligned} \langle Q(x, p) \rangle &= \int_{-\infty}^{\infty} \psi^*(x, t) Q\left(x, \frac{\hbar}{i} \frac{\partial}{\partial x}\right) \psi(x, t) dx \\ &= \int_{-\infty}^{\infty} \varphi^*(x) Q\left(x, \frac{\hbar}{i} \frac{d}{dx}\right) \varphi(x) dx , \end{aligned} \quad (20.11)$$

so that for these states $\frac{d}{dt} \langle Q(x, p) \rangle = 0$.

¹ Dynamical variables are quantities which depend on x and p .

Discussion 20.7

Explain how we go from the first to the second line of eq. (20.11) and how we conclude that $\langle Q \rangle$ does not depend on time *for a stationary state*.

2. These are states with a definite energy – or in other words we will always measure the same energy for a particle in this state. Recall that the Hamiltonian operator in position space is given by,

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(\hat{x}) . \quad (20.12)$$

Following our recipe for calculating expectation values eq. (10.9), to the determine the expectation value of the total energy, we should sandwich \hat{H} between Ψ^* and Ψ and integrate

$$\langle \hat{H} \rangle = \int_{-\infty}^{\infty} \psi^*(x, t) \hat{H} \psi(x, t) dx = \int_{-\infty}^{\infty} \varphi^*(x) \hat{H} \varphi(x) dx . \quad (20.13)$$

From the form of eq. (20.12) we see that the time independent Schrödinger equation eq. (20.5) can be written:

$$\hat{H}\varphi(x) = E\varphi(x) . \quad (20.14)$$

Discussion 20.8

Show that eq. (20.5) can be written as eq. (20.14).

Notice that eq. (20.5) is nothing but the energy eigenvalue equation, and we can now see why people also refer to eq. (20.14) as the time-independent Schrödinger equation.

Using eq. (20.14), eq. (20.13) can be written

$$\begin{aligned} \langle \hat{H} \rangle &= \int_{-\infty}^{\infty} \varphi^* \hat{H} \varphi dx \\ &\stackrel{\text{eq. (20.14)}}{=} \int_{-\infty}^{\infty} \varphi^* E \varphi dx \\ &= E \int_{-\infty}^{\infty} \varphi^* \varphi dx \stackrel{\text{eq. (20.10)}}{=} E \underbrace{\int_{-\infty}^{\infty} |\psi|^2 dx}_1 \\ &= E \end{aligned} \quad (20.15)$$

In addition,

$$\begin{aligned} \hat{H}^2 \varphi &= \hat{H} (\underbrace{\hat{H} \varphi}_{E\varphi}) \\ &= E \hat{H} \varphi = E^2 \varphi , \end{aligned} \quad (20.16)$$

so

$$\langle \hat{H}^2 \rangle = \int_{-\infty}^{\infty} \varphi^* \hat{H}^2 \varphi dx = E^2 \int_{-\infty}^{\infty} |\psi|^2 dx = E^2 , \quad (20.17)$$

so

$$(\Delta E)^2 = \langle \hat{H}^2 \rangle - \langle \hat{H} \rangle^2 = 0 . \quad (20.18)$$

Recall the $\Delta E = 0$ implies that there is no spread in the measurements of total energy. This means that a measurement of the total energy is certain to return a value of E .

Discussion 20.9

Explain why $\Delta E \neq 0$ if ψ is a superposition of energy eigenstates with different energies.

3. The general solution of the Schrödinger can be written as a linear combination of separable solutions.

A linear combination of a set of functions, $\psi_1, \psi_2, \psi_3, \dots$, is of the form

$$\psi = c_1 \psi_1 + c_2 \psi_2 + c_3 \psi_3 + \dots , \quad (20.19)$$

where c_1, c_2, c_3, \dots , are constants.

We'll see that solutions to the time independent Schrödinger equation give an infinite number of solutions, $\{\varphi_1, \varphi_2, \varphi_3, \dots\}$.

Each solution is associated with some value of the separation constant $\{E_1, E_2, E_3, \dots\}$, respectively so there are different wave functions for each allowed energy.

$$\begin{aligned}\psi_1(x, t) &= \varphi_1(x)e^{-iE_1 t/\hbar} \\ \psi_2(x, t) &= \varphi_2(x)e^{-iE_2 t/\hbar} \\ \psi_3(x, t) &= \varphi_3(x)e^{-iE_3 t/\hbar} \\ &\text{etc...}\end{aligned}\tag{20.20}$$

Any linear combination of solutions to the Schrödinger equation eq. (20.1), is also a solution. Once we have a separable solution, we can construct a much more general solution:

$$\psi(x, t) = \sum_{n=1}^{\infty} c_n \varphi_n(x)e^{-iE_n t/\hbar}.\tag{20.21}$$

It turns out the eq. (20.21) is the most general form of the solution to the Schrödinger equation.

Discussion 20.10

Show that a linear combination of solutions to the time-dependent Schrödinger equation, as in eq. (20.19), is also a solution to the time-dependent Schrödinger equation. Show that a linear combination of solutions to the time-*independent* Schrödinger equation, with *different* energies, is *not* a solution to the time-independent Schrödinger equation. What goes wrong?

Summary

We now review our recipe for finding general time-dependent solutions to the Schrödinger equation in the language of position space.

Given $V(x)$ and starting wave function $\psi(x, 0)$ our job will be to find $\psi(x, t)$ for later times.

To achieve this goal, we follow these steps:

- Solve the time-independent Schrödinger equation which leads to an infinite set $\{\varphi_1, \varphi_2, \dots\}$.
- To fit $\psi(x, 0)$ we need to find $\{c_1, c_2, \dots\}$ so that

$$\psi(x, 0) = \sum_{n=1}^{\infty} c_n \varphi_n(x).\tag{20.22}$$

We will see later how we can do this.

- To get $\psi(x, t)$ we need to tack on $\exp(-iE_n t/\hbar)$ to the n -th term.

$$\psi(x, t) = \sum_{n=1}^{\infty} c_n \varphi_n(x)e^{-iE_n t/\hbar}.\tag{20.23}$$

Take note!

While separable solutions, $\psi_n(x, t) = \varphi_n(x)e^{-iE_nt/\hbar}$, are stationary, the general solution, eq. (20.23), may **not** be.

Recall that for stationary states all expectation values are constant.

Energy eigenfunction reality

It turns out that we can always take our energy eigenfunctions, $\varphi_E(x)$, which satisfy, eq. (20.5), to be real. This doesn't mean that every solution to the time-independent Schrödinger equation is real; what it says is that if you've got one that is not, it can always be expressed as a linear combination of solutions (with the same energy) that are. So you might as well stick to φ 's that are real².

To see how this works, we take the complex conjugate of eq. (20.5) which leads to

$$-\frac{\hbar^2}{2m} \frac{d^2\varphi^*}{dx^2} + V\varphi^* = E\varphi^* \quad (20.24)$$

² Sometimes, we will see that it is actually convenient to use complex φ .

❖ Discussion 20.11

Explain how taking the complex conjugate of eq. (20.5) leads to eq. (20.24).

Equation eq. (20.24) tells us that φ^* also satisfies the energy eigenvalue equation with the same eigenvalue as φ . Since eq. (20.5) is a second order linear differential equation it has two linearly independent solutions. Let's call them φ_a and φ_b . The general solution is then of the form $\varphi = A\varphi_a + B\varphi_b$. Now, using the fact that linear combinations of solutions to eq. (20.5) are also solutions (with the same energy eigenvalue), in particular, the real linear combinations $\varphi + \varphi^*$ and $i(\varphi - \varphi^*)$ are also solutions. This tells us that given a complex solution, we can obtain real ones.

❖ Discussion 20.12

Show that given two solutions φ_a and φ_b which solve eq. (20.5) with energy eigenvalue, E , the linear combination $c_a\varphi_a + c_b\varphi_b$ is also a solution with the same energy eigenvalue, E . What goes wrong if they have *different* energy eigenvalues?

❖ Discussion 20.13

Show that given a complex φ , $\varphi + \varphi^*$ and $i(\varphi - \varphi^*)$ are real.

❖ Discussion 20.14

In your notes, try to write as many equations in this chapter using Dirac notation rather than in position space.

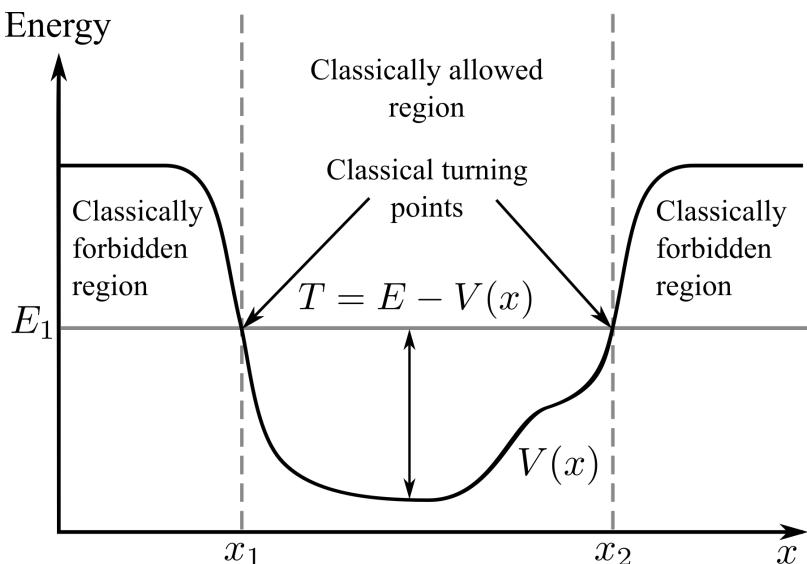
21

Infinite Square well

We will now investigate solving the energy eigenvalue equation:

$$\left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x)\right) \varphi_E(x) = E \varphi_E(x), \quad (21.1)$$

whose solutions depend on the form of the potential energy function $V(x)$.



Most potentials we will study resemble the one in Figure 21.1. Firstly, note that, it has a minimum – such a potential is called a **potential well**. Classically, for such a system, the total energy is conserved. This implies that we can write the kinetic energy of our particle, at the point x , as $T(x) = E - V(x)$. So that, in Figure 21.1, the length of the vertical arrow (\uparrow) between the fixed energy E_1 and $V(x)$ is a measure of the kinetic energy, T . Classically, the kinetic energy is positive: $T \geq 0$. This means that a particle with energy E_1 subject to the potential, $V(x)$, shown in Figure 21.1, can only move between the points x_1 & x_2 . These points are called **classical turning points** and the region between them is called the **classically allowed region**. Note that, as shown in Figure 21.2, for many potentials, the size of allowed region depends on

This reading is based on [McIntyre et al. 2012, Section 5.4]

Figure 21.1: A generic potential energy well.

Modification of this image
Attribution: Benjamin D. Esham (bdesham), Public domain, via Wikimedia Commons

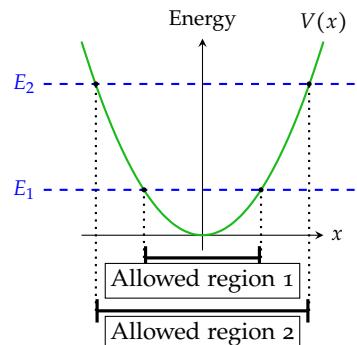


Figure 21.2: As we increase the total energy, $E = T + V$, the size of the allowed region gets larger.

¹ An exception to this will be the infinite square well discussed below.

a particle's energy¹. The regions beyond the allowed region are called **classically forbidden regions**. We refer to particles trapped in a potential as **bound states** and particles with enough energy to escape the well as **unbound states**.

Particle in a box

We will be looking at many potentials but will start with a very simple model of a confined particle. Classically, the system we will study corresponds to a ball bouncing between two perfectly elastic walls. We'll call this a **particle in a box**.

The ball flies freely between walls where there is no force on the ball. The ball is reflected perfectly and instantaneously at each bounce. This means that there is an infinite force at the walls. Furthermore, the ball stays in the box no matter its energy which means that there is infinite potential energy outside the box. Putting this all together, our model corresponds to the potential:

$$V(x) = \begin{cases} 0 & 0 < x < L \\ \infty & \text{otherwise} \end{cases} \quad (21.2)$$

Obviously the infinite square well is not realistic, but it is easier to solve than other potentials and has many features of more realistic potentials. This makes it a good "toy" model to study and learn about quantum mechanics.

Discussion 21.1

Explain why, classically, a discontinuous jump in a potential corresponds to an infinite force.

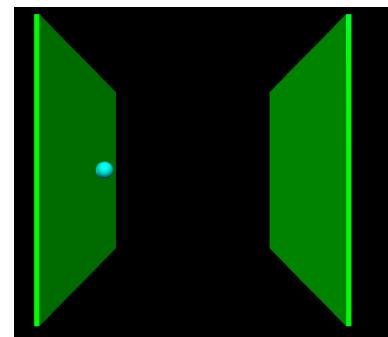


Figure 21.3: Particle in a box. You can view an animation [here](#).

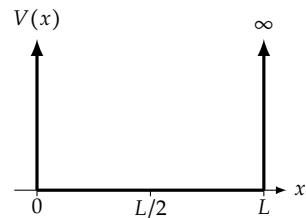


Figure 21.4: Infinte square well

Discussion 21.2

Why would we classically need an infinite force for the direction of a particle to change instantaneously?

Finding energies & eigenstates

Having decided on a potential, to find the allowed quantum energies and energy eigenstates of the system, we need to solve the energy eigenvalue equation (21.1). Thinking outside the box, when $|x| > L$, using (21.2), (21.1) becomes,

$$\left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \infty \right) \varphi_E(x) = E \varphi_E(x) . \quad (21.3)$$

Now if we are looking for solutions with finite energy, E , the right hand side of (21.3) must be finite. This means that we must have, $\varphi_E = 0$, for the left hand side of (21.3) to be finite.

Then looking, inside the box, using (21.2), (21.1) becomes

$$\left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + 0\right) \varphi_E(x) = E \varphi_E(x) , \quad (21.4)$$

which can be rewritten as

$$\frac{d^2 \varphi_E(x)}{dx^2} = -k^2 \varphi_E(x) , \quad (21.5)$$

where we have defined,

$$k^2 = \frac{2mE}{\hbar^2} \quad (21.6)$$

The energy is positive in this case which means that k must be real². The parameter, k , is called the **wave number**. Solutions to (21.5) can either be written in terms of trigonometric functions

$$\varphi_E(x) = A \sin kx + B \cos kx , \quad (21.7)$$

or complex exponentials

$$\varphi_E(x) = C e^{ikx} + D e^{-ikx} . \quad (21.8)$$

As we saw in the last chapter, bound state energy eigenstate wave functions can always be taken to be real. This means it will be more convenient to work with (21.7). In summary we have found, thinking both inside and outside the box:

$$\varphi_E(x) = \begin{cases} 0, & x < 0 , \\ A \sin kx + B \cos kx, & 0 < x < L , \\ 0, & x > L . \end{cases} \quad (21.9)$$

Our solution has 3 unknowns: A , B and k . Two of the unknowns, A and B come from the fact that we have solved a second order homogeneous linear equation, whereas the parameter k is related to the energy eigenvalue. To complete our discussion we will need 3 more pieces of information. Two of these come from **boundary conditions**.

For our mathematical solution, (21.9), to represent a physical solution, the wave function must be continuous across different regions of space as we do not expect probability density to suddenly jump. Requiring continuity at, $x = 0$, and, $x = L$, and using (21.9), tells us that

$$\begin{aligned} \varphi_E(0) : \quad & A \sin(0) + B \cos(0) = 0 \\ \varphi_E(L) : \quad & A \sin(kL) + B \cos(kL) = 0 \end{aligned} \quad (21.10)$$

Evaluating the boundary condition at $x = 0$ we get $B = 0$. Then, using $B = 0$ and looking at $x = L$ we get,

$$A \sin(kL) = 0 \implies kL = n\pi \quad \text{where } n \in \mathbb{Z} . \quad (21.11)$$

So by imposing the boundary conditions we can find the possible values for k :

$$k_n = n \frac{\pi}{L} \quad (21.12)$$

² Intuitively we expect the energy to be larger than (or equal to) the minimum of the potential, $V_{\min} = 0$. Mathematically, if we consider solutions to (21.5) with $E < 0$, we will get exponential functions which will not allow us to satisfy the boundary conditions discussed below – see Discussion 21.3.

where, $n = 1, 2, 3, \dots$. Notice that we have excluded $n = 0$, since it would lead to $\varphi = 0$. This is not an interesting case since it would mean the particle is nowhere. Negative n has also been excluded since that would give an equivalent state to the one with the corresponding positive n (remember an overall phase $(-1 = e^{i\pi})$ does not change the state).

Discussion 21.3

Show that if we consider solutions with $E < V_{\min}$, we get exponential solutions. Show that non-trivial³ exponential solutions can not satisfy both boundary conditions:

$\varphi_E(0) = \varphi_E(L) = 0$. We will look at why we expect $E > V_{\min}$ in general later in the course.

³ The trivial solution is $\varphi = 0$ everywhere.

Looking at (21.12) we see that only discrete k_n are allowed. This is called a **quantisation condition** and the index n called a **quantum number**. Now, substituting (21.12) into (21.6) allows us to find the quantised energy eigenvalues

$$E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{\hbar^2 n^2 \pi^2}{2mL^2} \quad (21.13)$$

with the corresponding energy eigenstates (inside the well):

$$\varphi_E(x) = A \sin \frac{n\pi x}{L}. \quad (21.14)$$

A plot of the first few energy eigenvalues is shown in Figure 21.5.

One final unknown in (21.14), A , remains. We can determine A by normalising our wave functions:

$$\begin{aligned} 1 = \langle E_n | E_n \rangle &= \int_{-\infty}^{\infty} |\varphi_E(x)|^2 dx \\ &= \int_0^L |A|^2 \sin^2 \left(\frac{n\pi x}{L} \right) dx \\ &= |A|^2 \frac{L}{2}. \end{aligned} \quad (21.15)$$

Discussion 21.4

Show that $\int_0^L \sin^2(k_n x) dx = L/2$.

Hint: $\sin^2(x) = \frac{1}{2}(1 - \cos(2x))$.

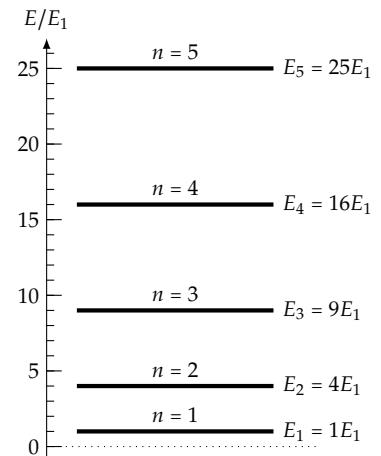


Figure 21.5: Energy spectrum of the infinite square potential energy well.

Taking A as real and positive, we find that $A = \sqrt{2/L}$, so that our energy eigenstates are

$$\varphi_E(x) = \begin{cases} 0, & x \leq 0 \text{ and } x \geq L, \\ \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right), & 0 \leq x \leq L, \end{cases} \quad \text{where } n = 1, 2, 3, \dots \quad (21.16)$$

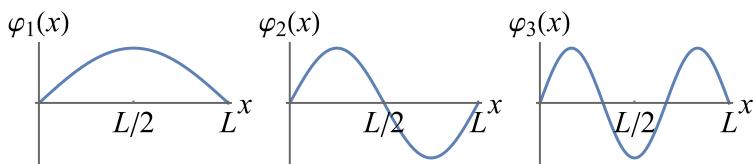


Figure 21.6: Wave functions of the first three energy eigenstates of the infinite square potential energy well.

The first few energy eigenstates are shown in Figure 21.6. From these plots it should be obvious why we refer to $\psi(x)$ as a wave function – they look like modes on a guitar string. Notice that they also fit inside the well. Indeed, if we relate the wave number, k , to a wave length λ ,

$$\varphi_E(x) = \sin(kx) = \sin\left(\frac{2\pi x}{\lambda}\right) \implies k = \frac{2\pi}{\lambda} , \quad (21.17)$$

we can rewrite the quantisation condition (21.12) as

$$\frac{2\pi}{\lambda_n} = k_n = \frac{n\pi}{L} \implies L = \frac{n}{2}\lambda_n . \quad (21.18)$$

This tells us that an integer number of half wavelengths fit inside the well. While analogous, classical and quantum waves are not the same. For example, the energy of a vibrating guitar string, depends on its amplitude and it can have continuous range of energies. On the other hand quantum standing waves have discrete energies. In classical physics, waves and particles are distinct. On the other hand, in quantum mechanics, depending on the measurements we make, a system can have particle-like or wave-like properties. This is called **wave-particle duality**.

Example 21.1: Atom sized box

To get a feel for these solutions let's look at what the energy of a particle in an atom sized box is. Suppose we have an electron, with, $m_e = 511 \text{ keV}/c^2$, confined to a roughly atom size box of width 0.2 nm. The, $n = 1$, or ground state energy is

$$\begin{aligned} E_1 &= \frac{\hbar^2 \pi^2}{2mL^2} \\ &= \frac{\pi^2 (6.58 \times 10^{-16} \text{ eVs})}{2(511 \times 10^3 \text{ eV}/c^2)(0.2 \times 10^{-9} \text{ m})^2} \\ &= 9.4 \text{ eV} , \end{aligned} \quad (21.19)$$

which is the right order of magnitude associated with atomic electron ionization energies as shown in Figure 21.7. Calculating the exact ionisation energy for each atom will require much more realistic models – in this course we will only look at the simplest case, namely Hydrogen.

Having found the wavefunctions, we can study probability densities and expectation values. The probability density for the n^{th} state is

$$\mathcal{P}_n(x) = |\varphi_n(x)|^2 = \frac{2}{L} \sin^2\left(\frac{n\pi x}{L}\right) . \quad (21.20)$$

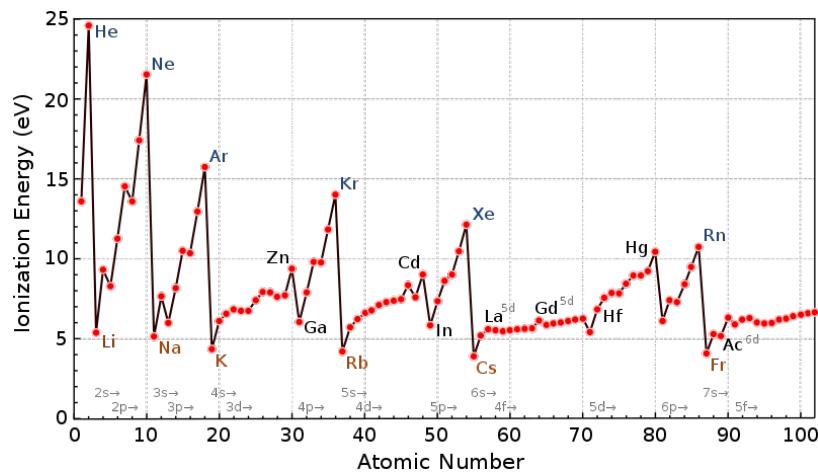


Figure 21.7: Atomic electron ionisation energies

[Link to image source](#)

Attribution: Ponor, CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0/>, via Wikimedia Commons

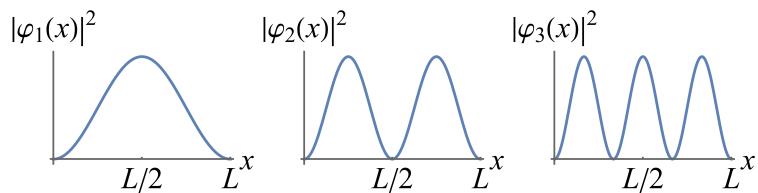


Figure 21.8: Probability densities of the first three energy eigenstates of the infinite square potential energy well.

The probability densities for the first three states are shown in Figure 21.8. As expected the probability for finding the particle outside the well is zero which is what we had in the classical case. We see however that, for $n > 1$, there are places in the well where we have zero probability of finding the particle – this is not what we expect classically. These points correspond to nodes of the wave function demonstrating a wave-like property of our quantum particle.

Example 21.2: Middling Expectations

Problem What is $\langle \hat{x} \rangle$ for the n^{th} state of infinite square well?

Solution

$$\begin{aligned}
 \langle \hat{x} \rangle &= \langle E_n | \hat{x} | E_n \rangle \\
 &= \int_{-\infty}^{\infty} \varphi_n^*(x) x \varphi_n(x) dx = \int_{-\infty}^{\infty} x |\varphi_n(x)|^2 dx \\
 &= \frac{2}{L} \int_0^L x \sin^2\left(\frac{n\pi x}{L}\right) dx = \frac{2}{L} \left(\frac{L}{n\pi}\right)^2 \int_0^{n\pi} u \sin^2(u) du \\
 &= \frac{2}{L} \left(\frac{L}{n\pi}\right)^2 \left[\frac{u^2}{4} - \frac{u \sin(2u)}{4} - \frac{\cos(2u)}{8} \right]_0^{n\pi} \\
 &= \frac{2}{L} \left(\frac{L}{n\pi}\right)^2 \left(\frac{(\pi n)^2}{4} - \frac{n\pi \sin(2n\pi)}{4} - \frac{\cos(2n\pi)}{8} + \frac{1}{8} \right) \\
 &= \frac{2}{L} \left(\frac{L}{n\pi}\right)^2 \left(\frac{(\pi n)^2}{4} \right) \\
 &= \frac{L}{2}
 \end{aligned} \tag{21.21}$$

This is the answer we expect from the symmetry of the problem – given that there is no preference to either side of the well, we take the average of our measurements to give $L/2$.

💬 Discussion 21.5

Show, using trigonometric identities and integration by parts that

$$\int u \sin^2(u) du = \frac{1}{4}u^2 - \frac{1}{4}u \sin(2u) - \frac{1}{8} \cos(2u).$$

💬 Discussion 21.6

$$\text{Show that } \left[\frac{u^2}{4} - \frac{1}{4}u \sin(2u) - \frac{1}{8} \cos(2u) \right]_0^{n\pi} = \frac{(n\pi)^2}{4}.$$

💬 Discussion 21.7

For the state $\varphi_2(x)$, the probability of finding the particle at $x = L/2$ is zero. On the other hand, we found that $\langle \hat{x} \rangle = L/2$. How do you reconcile these two facts?

💬 Discussion 21.8

Before we apply the boundary conditions, the solutions to the infinite square well involves 3 unknowns A , B , and k :

$\varphi_E(x) = A \cos(kx) + B \sin(kx)$. Usually when we solve a second order linear differential equation there are 2 constants of integration. Where does the third constant come from?

The results of this section are graphically summarised in Figure

21.9.

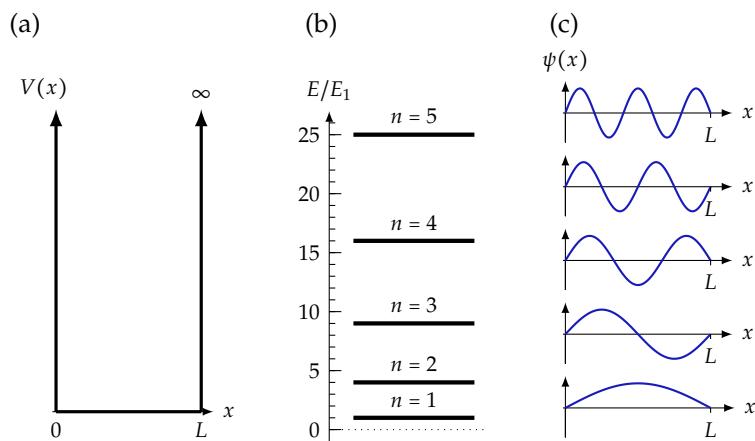


Figure 21.9: (a) Infinite square potential energy well, (b) spectrum of allowed energies, and (c) energy eigenstate wave functions.

The three figures shown in Figure 21.9 are often combined into a single figure as in Figure 21.10.

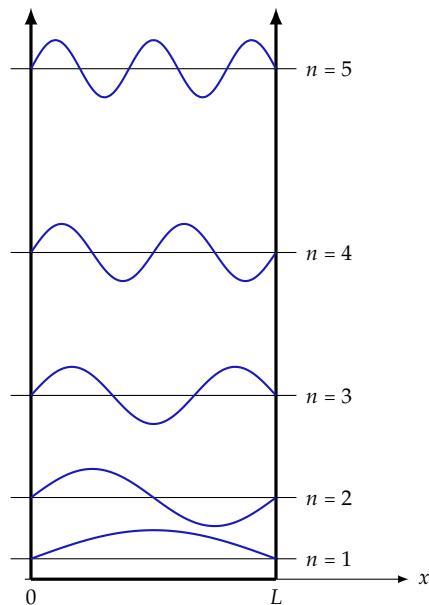


Figure 21.10: Unified schematic diagram of infinite square well problem and solution. Note that two vertical scales are implied. For the potential energy well and the energy spectrum, the vertical scale is energy with the origin at the bottom of the well. For the wave functions, the vertical scale is probability amplitude ($1/\sqrt{\text{length}}$) with the $\psi = 0$ origin for each state centred on the energy of that state.

Discussion 21.9

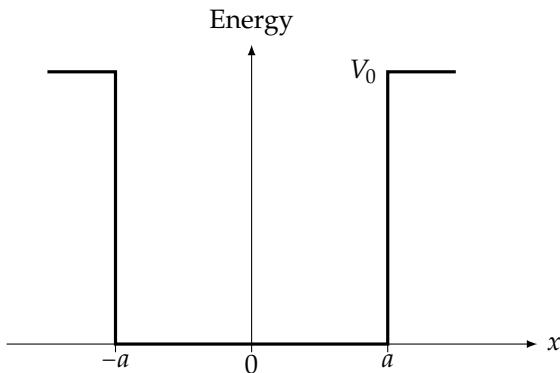
Write a summary of the important points and formulae in this chapter.

22

Finite square well

Following on from the infinite square well, we now look at a *slightly* more realistic model by allowing energy outside the box to be finite. This new potential will allow us to investigate both bound and scattering states¹. We still take a square well, only this time centring it on $x = 0$. More explicitly we have

$$V(x) = \begin{cases} V_0, & x < -a \\ 0, & -a < x < a \\ V_0, & x > a . \end{cases} \quad (22.1)$$



This reading is based on [McIntyre et al. 2012, Section 5.5]

¹ Do you remember what bound and scattering states are?

Figure 22.1: Finite square well potential

Equation (22.1) is illustrated in Figure 22.1.

Symmetric potentials

Before studying the potential (22.1) further we will discuss a useful result about symmetric potentials.

Suppose we have a potential which satisfies $V(x) = V(-x)$. We say such a potential is symmetric. Solutions to symmetric potentials can be taken to either even, that is $\varphi(x) = \varphi(-x)$ or odd, that is $\varphi(x) = -\varphi(-x)$. Suppose φ is a solution to the energy eigenvalue equation (20.5). Then substituting $x \rightarrow -x$ into (20.5) gives

$$-\frac{\hbar^2}{2m} \frac{d^2\varphi}{d(-x)^2} + V(-x)\varphi(-x) = E\varphi(-x) . \quad (22.2)$$

Then using the symmetry of the potential $V(x) = V(-x)$ and the

fact that, $\frac{d^2\varphi}{d(-x)^2} = \frac{d^2\varphi}{dx^2}$, we conclude that

$$-\frac{\hbar^2}{2m} \frac{d^2\varphi}{dx^2} + V(x)\varphi(-x) = E\varphi(-x). \quad (22.3)$$

This in turn tells us that $\varphi(-x)$ satisfies the energy eigenvalue equation with the same energy eigenvalue as $\varphi(x)$. It now follows that, $\varphi_+(x) = \varphi(x) + \varphi(-x)$ and $\varphi_-(x) = \varphi(x) - \varphi(-x)$ are also solutions to (22.2). The conclusion is that, when the potential is symmetric, we can take the solutions to (22.2) to be even or odd.

💬 Discussion 22.1

Show that $\varphi_+(x) = \varphi(x) + \varphi(-x)$ and $\varphi_-(x) = \varphi(x) - \varphi(-x)$ are even and odd respectively.

💬 Discussion 22.2

Explain why, if for a given energy eigenvalue E , we have a symmetric potential with *only one* linearly independent solution, $\varphi(x)$ then either $\varphi = \varphi_+$ (with $\varphi_- = 0$) or $\varphi = \varphi_-$ (with $\varphi_+ = 0$). **Hint:** Start off by explaining why in this case we must have $\varphi(x) = \lambda\varphi(-x)$ for some constant λ .

Solution

Now, with the potential (22.1), the energy eigenvalue equation both inside and outside the box becomes

$$\left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + 0 \right) \varphi_E(x) = E\varphi_E(x), \quad \text{inside the box} \quad (22.4)$$

$$\left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V_0 \right) \varphi_E(x) = E\varphi_E(x), \quad \text{outside the box} \quad (22.5)$$

For now we look at **bound states** – in other words states with energies less than V_0 : $E < V_0$. In what follows, it will be useful to define

$$k = \sqrt{\frac{2mE}{\hbar^2}}, \quad (22.6)$$

and

$$q = \sqrt{\frac{2m}{\hbar^2}(V_0 - E)}. \quad (22.7)$$

Note that since $V_0 > E$, q is a real number. Using (22.6) and (22.7), we can rewrite (22.4) and (22.5) as

$$\frac{d^2\varphi_E(x)}{dx^2} = -k^2\varphi_E(x), \quad \text{inside the box} \quad (22.8)$$

$$\frac{d^2\varphi_E(x)}{dx^2} = +q^2\varphi_E(x), \quad \text{outside the box} \quad (22.9)$$

💬 Discussion 22.3

Show that (22.5) can be written as in (22.9) where q is given by (22.7).

Note that (22.8) is the same equation we had inside the box when looking at the infinite square well so we have the same solutions

$$\varphi_E(x) = C \sin kx + D \cos kx, \quad -a < x < a . \quad (22.10)$$

On the other hand, (22.9) describes the behaviour of the wave function outside the box even though this is a classically forbidden region, 🤯. Indeed (22.9) has the solution:

$$\varphi_E(x) = Ae^{qx} + Be^{-qx} . \quad (22.11)$$

💬 Discussion 22.4

Show that $f(x) = Ae^{qx} + Be^{-qx}$ satisfies $\frac{d^2f}{dx^2} = q^2 f$.

💬 Discussion 22.5

Explain why it is surprising that the wave function is non-zero in the classically forbidden region.

We see that, in the classically forbidden region, the solution is exponentially growing or decaying. The characteristic length scale associated with this decay or growth length is $1/q$.

💬 Discussion 22.6

Explain, in your own words the claim that e^{-qx} has a decay length of $1/q$.

Bringing the solutions in all regions together we have

$$\varphi_E(x) = \begin{cases} Ae^{qx} + Be^{-qx} , & x < -a \\ C \sin kx + D \cos kx, & -a < x < a \\ Fe^{qx} + Ge^{-qx} , & x > a . \end{cases} \quad (22.12)$$

To proceed, we now need to apply boundary conditions. As before, we require the wave function to be continuous across boundaries so that there are no sudden jumps in probability. A new condition is that we also require that the slope of the wave function is also continuous across a boundary. This is because a discontinuous slope would imply infinite kinetic energy – can you see why? We did not require continuity of $\psi'(x)$ for the infinite square well since in that case $V(\text{boundary}) = \infty$.

In summary, we generally have the following boundary conditions for our wave function:

Wave function boundary conditions

- 1) $\varphi_E(x)$ is continuous
 - 2) $\frac{d\varphi_E(x)}{dx}$ is continuous unless $V = \infty$
- (22.13)

In addition, we also have boundary conditions at $x = \pm\infty$. Outside the well, the wave function must be decaying since a growing exponential term would prevent normalisation of wave function. This means that for the wave function to be normalisable, we must have $B = F = 0$. With this simplification, (22.12) becomes continuous

$$\varphi_E(x) = \begin{cases} Ae^{qx}, & x < -a \\ C \sin kx + D \cos kx, & -a < x < a \\ Ge^{-qx}, & x > a . \end{cases} \quad (22.14)$$

The potential (22.1) is symmetric about $x = 0$, or in other words, $V(x) = V(-x)$. We have seen that this means that, solutions can be taken to be either even (i.e. $\psi(x) = \psi(-x)$) or odd (i.e. $\psi(x) = -\psi(-x)$).

solutions are either even (i.e. $\psi(x) = \psi(-x)$) or odd (i.e. $\psi(x) = -\psi(-x)$). For now, we will assume that this is true (or you can grind through the algebra without making the assumption and end up with the same result). We will look at this issue again in relation to the **parity** operator.

Assuming the wave function is even or, in other words $\varphi_E(x) = \varphi_E(-x)$, we find that, from (22.14), this will be true if, $C = 0$ and $A = G$, leading to

$$\varphi_{\text{even}}(x) = \begin{cases} Ae^{qx}, & x < -a \\ D \cos kx, & -a < x < a \\ Ae^{-qx}, & x > a . \end{cases} \quad (22.15)$$

On the other hand, assuming the wave function is odd or, in other words $\varphi_E(x) = -\varphi_E(-x)$, we see that, from (22.14), this will be true if, $D = 0$ and $A = -G$, leading to

$$\varphi_{\text{odd}}(x) = \begin{cases} Ae^{qx}, & x < -a \\ C \sin kx, & -a < x < a \\ -Ae^{-qx}, & x > a . \end{cases} \quad (22.16)$$

Discussion 22.7

Given (22.14), show that assuming φ_E is odd, leads to (22.16).

Now looking at the even solutions (22.15), and applying the boundary conditions at $x = a$ gives us

$$\text{Continuity of } \varphi: \quad \varphi_{\text{even}}(a) : D \cos(ka) = Ae^{-qa}, \quad (22.17)$$

and

$$\text{Continuity of } \frac{d\varphi}{dx}: \quad \left. \frac{d\varphi_{\text{even}}}{dx} \right|_{x=a} : -kD \sin(ka) = -qAe^{-qa}. \quad (22.18)$$

Discussion 22.8

Given (22.15), show requiring φ and $\frac{d\varphi}{dx}$ are continuous at $x = -a$ leads to (22.17) and (22.18).

Evaluating the boundary conditions at $x = -a$, lead to the same equations. This means that although we have two boundary conditions, which allow us to eliminate two unknowns, we have three unknowns, A , D and E (remember that E is hidden in k and q).

Dividing (22.18) by (22.17), we get

$$k \tan(ka) = q , \quad (22.19)$$

and since, k and q depend on E , (22.19), allows us to find the allowed energies.

Applying the same line of reasoning to the odd solutions (22.16), leads to

$$-k \cot(ka) = q . \quad (22.20)$$

Discussion 22.9

Given (22.16), show how we get (22.20) from the boundary condition at $x = a$.

One way to solve (22.19) and (22.20) involves defining new dimensionless parameters:

$$\begin{aligned} z &= ka = \sqrt{\frac{2mEa^2}{\hbar^2}} , \\ z_0 &= \sqrt{\frac{2mV_0a^2}{\hbar^2}} . \end{aligned} \quad (22.21)$$

Notice that, $z \propto \sqrt{E}$, so it parametrises the energy. On the other hand, $z_0 \propto \sqrt{V_0}$ which means it parametrises the depth of the well. On the other hand, $z_0 \propto a$, and it also parametrises the width of the well².

Together with, $qa = \sqrt{\frac{2m(V_0 - E)a^2}{\hbar^2}}$, (22.21), implies the useful relations

$$z_0^2 = (ka)^2 + (qa)^2 \implies (qa)^2 = z_0^2 - (ka)^2 = z_0^2 - z^2 . \quad (22.22)$$

Discussion 22.10

Given (22.21) and $qa = \sqrt{\frac{2m(V_0 - E)a^2}{\hbar^2}}$, show that $(ka)^2 + (qa)^2 = z_0^2$.

² You may also have noticed that we also have, $z \propto a$, but we will be solving for z given a particular z_0 to find the allowed energies. This means we will just be treating z as a dependant variable that we not free to vary.

Then using (22.21) and (22.22) means we can rewrite (22.19) and

(22.20) as

$$\begin{aligned} ka \tan(ka) &= qa \quad \rightarrow \quad z \tan(z) = \sqrt{z_0^2 - z^2} \\ -ka \cot(ka) &= qa \quad \rightarrow \quad -z \cot(z) = \sqrt{z_0^2 - z^2} \end{aligned}$$

Our task is now to solve two transcendental equations

$$z \tan(z) = \sqrt{z_0^2 - z^2} \quad \text{and} \quad -z \cot(z) = \sqrt{z_0^2 - z^2}. \quad (22.23)$$

The left hand side of both equations involve trig functions while the right hand side corresponds to a semi-circle with radius z_0 . We can solve them graphically or numerically for a given z_0 .

To find the allowed energies graphically, for a given z_0 , we plot both sides of the equations in (22.23). The solutions then correspond to intersections of the functions. Such a plot is shown in Figure 22.2 for the case $z_0 = 6$.

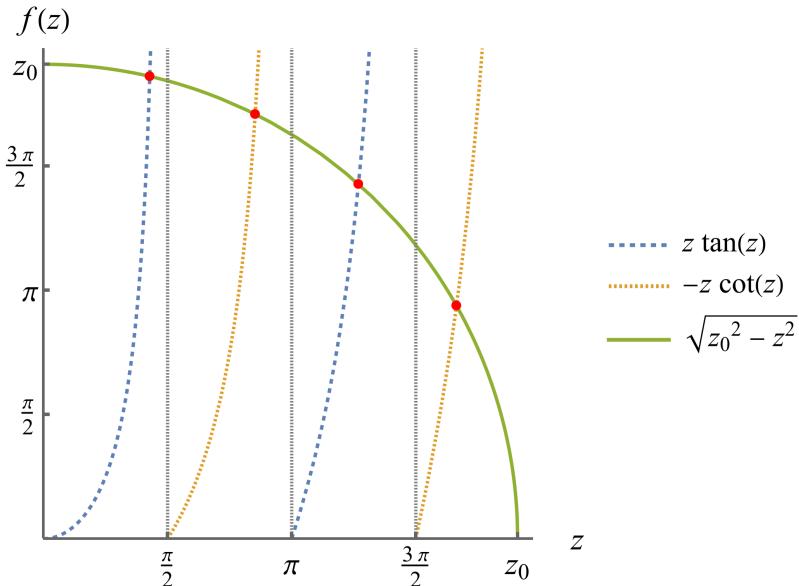


Figure 22.2: Plot of $z \tan z$ (blue dashed line), $-z \cot z$ (orange dotted line) and $\sqrt{z_0^2 - z^2}$ (green solid line) with $z_0 = 6$. Points of intersection correspond to solutions to (22.23). Red dots show points of intersection.

Notice from 22.2 that increasing z_0 makes the circle bigger. This means that we will get more solutions which implies that deeper, wider wells have more allowed bound states. With $z_0 = 6$ we have 4 intersection points which implies that there are 4 bound states. The intersections and allowed energies are (accurate to 2 decimal places)

$$\begin{aligned} z_1 &= 1.34 \rightarrow E_1 = 1.81 \frac{\hbar^2}{2ma^2} \\ z_2 &= 2.68 \rightarrow E_2 = 7.18 \frac{\hbar^2}{2ma^2} \\ z_3 &= 3.99 \rightarrow E_3 = 15.89 \frac{\hbar^2}{2ma^2} \\ z_4 &= 5.23 \rightarrow E_4 = 27.31 \frac{\hbar^2}{2ma^2} \end{aligned} \quad (22.24)$$

Once we have the energies, we just need to normalise the states. This is a straightforward but tedious processs. The results for $z_0 = 6$ are summarised in Figure 22.3.

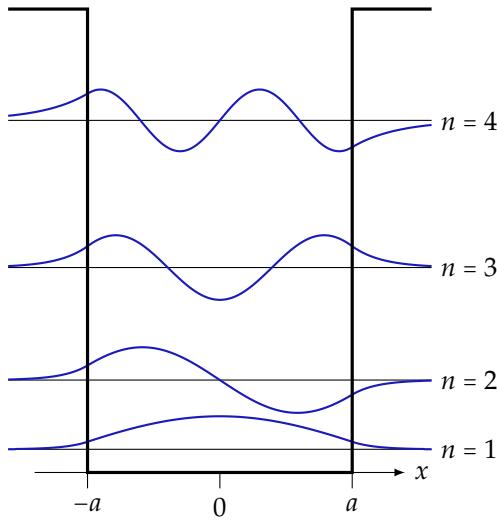


Figure 22.3: Unified schematic diagram of the finite potential energy well and the bound state solutions, showing the well, the allowed energies, and the energy eigenstate wave functions.

We can see that the results are similar to infinite square well with one **major exception**. For the finite square well solutions, the wave functions is non-zero in the classically forbidden region. This means it has a finite probability of being found where classical particles can not be! This is a purely quantum effect called **barrier penetration**. Barrier penetration leads to the phenomenon of quantum tunnelling which plays a role in many processes, for example radioactive decay.

Connecting the finite & infinite square well

A deep well corresponds to the limit in which $z_0 \rightarrow \infty$. As the circle in Figure 22.2 moves up, the intersections approach the asymptotes of the trigonometric functions. This means that the intersections are approximately at, $k_n a = z_n \approx \frac{n\pi}{2}$ which implies that $k_n \approx n\pi/2a$. Then using (22.6) we find

$$E_n = \frac{\hbar^2 k_n^2}{2m} \approx \frac{n^2 \pi^2 \hbar^2}{2m(2a)^2}, \quad (22.25)$$

which is the same result as the infinite square well,

$$E_n = \frac{n^2 \pi^2 \hbar^2}{2mL^2} \quad (22.26)$$

since our finite well has width $L = 2a$.

Inside the very deep well, we get the solutions

$$\begin{aligned} \varphi_n(x) &= \sqrt{\frac{2}{2a}} \cos\left(\frac{n\pi x}{2a}\right), \quad n = 1, 3, 5, \dots \\ \varphi_n(x) &= \sqrt{\frac{2}{2a}} \sin\left(\frac{n\pi x}{2a}\right), \quad n = 2, 4, 6, \dots \end{aligned} \quad (22.27)$$

which are the same solutions as the infinite square well –look different because the origin of the well shifted. Outside the well, in the limit $z_0 \rightarrow \infty$ the decay length ($1/q$) becomes zero which means the exponentials fall off instantaneously and the wave function is zero outside the well as expected.

 **Discussion 22.11**

Show that the solutions (22.27) are the same as the ones we had for the infinite square well once we shift our x -coordinate.

 **Discussion 22.12**

Write a summary of the important points and formulae in this chapter.

Review & Superposition states

Compare & Contrast

This reading is based on [McIntyre et al. 2012, Section 5.6-7]

We will now compare the energy eigenfunctions of the finite & infinite square well potentials, the first few of which are shown in Figure 23.1. We will also learn about some general features of energy eigenfunctions.

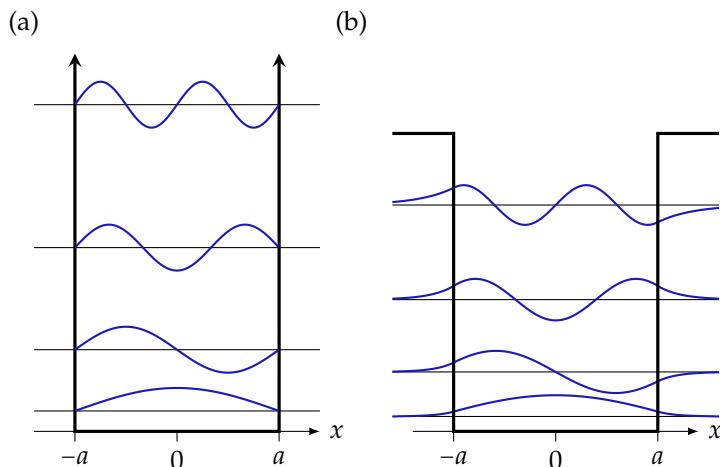


Figure 23.1: (a) Infinite and (b) finite well energy eigenstates.

Firstly note that both wave functions oscillate inside the well. They correspond to sinusoidal functions $\sin(kx)$ or $\cos(kx)$. Outside the well, we have exponential decay, e^{qx} or e^{-qx} , or disappearance of the wave function. For the ∞ -square well, we can think of the wave function as being an exponential that instantly decays with $\frac{1}{q} = 0$.

Interlude on f'' , concavity and curvature.

Before we proceed, we will briefly look at the relationship between the second derivative and concavity.

Consider the animation of an inflection point on [wikipedia](#). A “screenshot” of the animation is shown in Figure 23.2. From the animation we see that when $f'' < 0$ (green tangent), which means f' is decreasing and the function is curving down. We say the function has negative curvature or is concave down: \smile . On the other hand when $f'' > 0$ (blue tangent), f' is increasing and the

function is curving up. We say the function has positive curvature or is concave up:  . The simplest example to think of is a parabola, $f(x) = ax^2$: when a is positive, the function has positive curvature (concave up) and when a is negative, the function has negative curvature (concave down).



Wave Function Concavity

We would now like to look at the concavity of the wave function. We can rewrite energy eigenvalue equation in the following way:

$$\frac{d^2\varphi_E}{dx^2} = -\frac{2m}{\hbar^2}(E - V(x))\varphi_E \quad (23.1)$$

Note that, in the allowed region, $E > V$, while in the forbidden region $E < V$. From (23.1) we see that concavity is related to the signs of $E - V(x)$ and φ . Recall that $E - V$ is positive in allowed regions and negative in forbidden regions. The result of putting these two factors together, is summarised in Figure 23.3. In the allowed region, the curvature has opposite sign to the wave function, whereas in the forbidden region, the curvature has the same sign as the wave function.

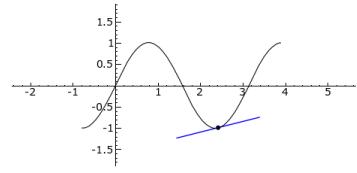


Figure 23.2: You can click on the image to get an animated version.

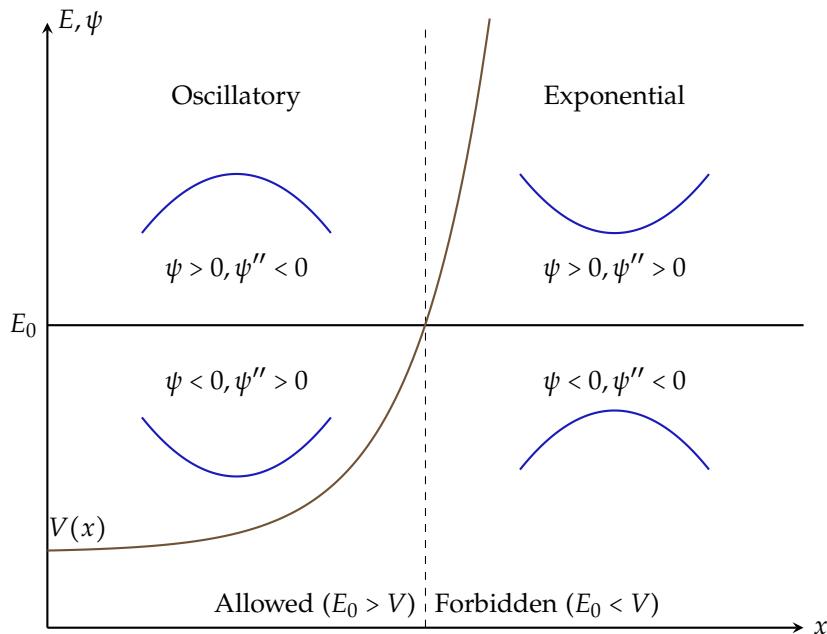


Figure 23.3: Curvature of the energy eigenstate wave functions in the allowed and forbidden regions.

Zooming in on general potentials

Take a general potential $V(x)$ and zoom in on a small enough region so that $V \approx \text{constant}$. Now we will look at the behaviour in the allowed and forbidden regions.

In the allowed region, with a constant potential we have the

energy eigenvalue equation:

$$\frac{d^2\varphi_E}{dx^2} = -k^2 \varphi_E \quad k = \frac{\sqrt{2m(E-V)}}{\hbar}. \quad (23.2)$$

Which leads to an oscillating wave function inside the well. Roughly speaking, since the wavelength of this wave is

$$\lambda = \frac{2\pi}{k} = \frac{h}{\sqrt{2m(E-V)}} \propto \frac{1}{\sqrt{\text{Kinetic Energy}}} \quad (23.3)$$

From (23.3) we see that the larger the difference between E and V is, or in other words the more kinetic energy there is, the smaller the wavelength. So the higher the kinetic energy the more wavy the wave function is¹.

Discussion 23.1

Explain the steps in eq. (23.3).

In the forbidden region, with a constant potential we have the energy eigenvalue equation:

$$\frac{d^2\varphi_E}{dx^2} = q^2 \varphi_E \quad q = \frac{\sqrt{2m(V-E)}}{\hbar}, \quad (23.4)$$

which tells us that we have exponentials in the forbidden region. Further more q gets smaller as E approaches V . If q is smaller, the decay length is larger and a larger decay length means that the wave function penetrates further into forbidden region. In other words, the more energetic states penetrate further into the forbidden region. This effect is visible in Figure 23.1 (b).

Looking at Figure 23.1 we see that each higher state has one more node, and the n^{th} energy level has $(n - 1)$ nodes. We do not count points where $V = \infty$ as nodes.

¹ Note that, in most cases we can not be certain about a particle's kinetic energy until we measure it. If the potential was constant *and* the particle was in an energy eigenstate, then we could be certain about its kinetic energy. Remember we are looking in at energy eigenstates and have zoomed in on our potential so that we can treat it as effectively constant.

States

The infinite well has an infinite number of bound states whereas the finite well has a finite number of bound states. We will see that the finite square well also has an infinite number of unbound states. It turns out that potential wells always have at least 1 bound state.

Parity

For both wells, the states are either symmetric or anti-symmetric relative to the centre of the well. Both Hamiltonians are symmetric with respect to the centre of the well. We say the Hamiltonian is invariant under the **parity** operation $x \rightarrow -x$. Since \hat{H} not affected by the **parity** operation, \hat{P} , \hat{H} commutes with parity, i.e. $[\hat{P}, \hat{H}] = 0$. This means that energy eigenstates are also eigenstates of parity

Symmetric states, that is those which satisfy $\hat{P}\varphi_n(x) = \varphi_n(-x) = +\varphi_n(x)$, have a parity eigenvalue of +1, and are called **even parity** states.

Antisymmetric states, i.e. those which satisfy, $\hat{\mathcal{P}}\varphi_n(x) = \varphi_n(-x) = -\varphi_n(x)$, have a parity eigenvalue of -1 , and are called **odd parity** states.

❖ Discussion 23.2

The parity operator, $\hat{\mathcal{P}}$, sends $x \rightarrow -x$. Which of the following functions below are parity eigenfunctions? If they are parity eigenfunctions, what is their eigenvalue?

- (a) e^{2x}
- (b) $\tan(x)$
- (c) $3x^2 + 7x^4$
- (d) $7 + x$

❖ Discussion 23.3

Show that the infinite square well potential below is parity invariant

$$V(x) = \begin{cases} \infty, & x < -a \\ 0, & -a < x < a \\ \infty, & x > a \end{cases}$$

How does this relate to (21.2)?

❖ Discussion 23.4

how that the kinetic energy operator, written in the position basis, $\hat{T} \doteq -\frac{\hbar^2}{2m} \frac{d^2}{dx^2}$, commutes with the parity operator $\hat{\mathcal{P}}$.

Hint: Show that for some general function $f(x)$, which may or may not be a parity eigenstate, that $\hat{\mathcal{P}}\hat{T}f(x) = \hat{T}\hat{\mathcal{P}}f(x)$.

Show that the momentum operator, $\hat{p} \doteq \frac{\hbar}{i} \frac{d}{dx}$, does not commute with the parity operator.

Orthonormality

Energy eigenstates form an orthonormal set. Normalisation is something we put in so that total probability is 1. On the other hand, orthogonality is a property of eigenstates of Hermitian operators. In Dirac notation, orthogonality corresponds to, $\langle E_n | E_m \rangle = \delta_{nm}$, where as in position space this becomes $\int_{-\infty}^{\infty} \varphi_n^*(x) \varphi_m(x) dx = \delta_{nm}$. It is relatively easy to show that the infinite square well states are orthogonal but more tedious to show for the finite well.

Completeness

The energy eigenstates form a complete basis – this is a property of eigenstates of Hermitian operators. This means we can use these basis functions to construct all possible solutions to the Schrödinger equation. In other words, the wave function of a general solution to the infinite square well can be written as a linear combination of the energy eigenfunctions:

$$\psi(x) = \sum_n c_n \varphi_n(x). \quad (23.5)$$

Each eigenstate satisfies, $H\varphi_n(x) = E_n \varphi_n$, **but** $\psi(x)$ may not. The completeness relation, eq. (23.5) is exact for infinite square well. For the finite well, we need to include unbound states.

The completeness relation, eq. (23.5), can be expressed as

$$\sum_n |E_n\rangle \langle E_n| = \hat{1}, \quad (23.6)$$

where $\hat{1}$ is the identity operator.

Fitting Energy eigenstates by eye

Using what we have learnt in this chapter, we can make rough sketches of energy eigenstate solutions for other potential well problems. Take home lessons to use are:

1. Oscillatory wave solution in allowed region

- Wavelength proportional to $1/\sqrt{E - V(x)}$
- $\psi > 0 \implies \psi'' < 0 : \sim$
- $\psi < 0 \implies \psi'' > 0 : \sim$

2. Exponential solution in forbidden region

- Decay length proportional to $1/\sqrt{V(x) - E}$
- $\psi > 0 \implies \psi'' > 0 : \sim$
- $\psi < 0 \implies \psi'' < 0 : \sim$

3. The n^{th} state has $(n - 1)$ -nodes²

4. Amplitude inside well related to wavelength – the larger the wavelength, the larger the amplitude.

- Hand waving argument to justify this rule: A long wavelength → lower kinetic energy in that part → particle spends more time there → more likely to find particle there → amplitude higher

5. The wave function, $\varphi_E(x)$, and its first derivative, $\frac{d\varphi_E(x)}{dx}$, must be continuous at boundaries where the nature of the potential changes³.

² This is a consequence of the Sturm-Liouville theorem which will not have time to discuss in this course. In counting nodes, we do not include points where an infinite potential results in the wave-function being zero. For example, we say that the ground state of the infinite square well has 0 nodes and not 2. Although usually the case when we have a single well, a state with $n - 1$ nodes does not necessarily have n -anti-nodes.

³ Unless the potential makes an infinite jump in which case $\frac{d\varphi_E(x)}{dx}$ is not continuous.

Example 23.1: Curve sketching

We can demonstrate how to use these rules by looking at how one sketches the 15th excited state of the potential in Figure 23.4. Firstly since its the 15th excited state, there will be 15 nodes. Then going from left to right, starting in the region $x < 0$, since the potential is infinite, the wave function must be zero. Continuity of the wave function then implies that $\psi = 0$ at $x = 0$. Note that the dotted line also indicates the energy of the state. This means that the region, $0 < x < L$, is an allowed region, and we have an oscillating wave function where as in the region $x > L$ the wave function will be exponentially decaying. Now in the region, $0 < x < L/2$, the potential is constant which means the amplitude and wavelength of the wave function should be constant. At the point $x = L/2$, ψ and $\frac{d\psi}{dx}$ are continuous. In the region, $L/2 < x < L$, the potential is increasing which means the amplitude and wavelength of the wave function must be increasing. Then at $x = L$, ψ and $\frac{d\psi}{dx}$ are continuous again. Finally in the region, $x > L$, we get exponential decay.

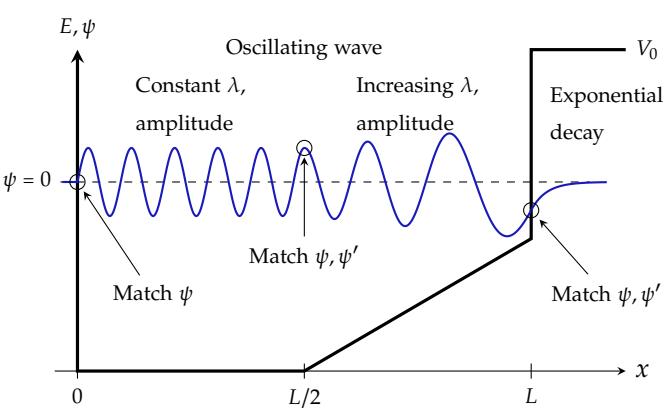


Figure 23.4: Drawing approximate energy eigenstate solutions

$E < V_{min}$ is forbidden even in QM

Using these curve sketching techniques, we can see why a state with energy less than the minimum of the potential, as shown in Figure 23.5, will not be normalisable. Now $E < V_{min}$, means that *everywhere* is classically forbidden. Suppose, we have $\psi(x) > 0$, then the wave function is concave up as shown in the figure.

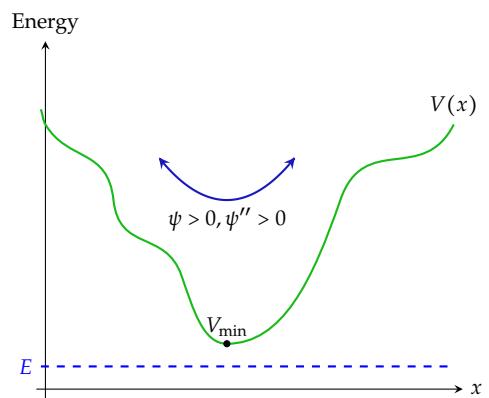


Figure 23.5: Energy less than the minimum of the potential.

Discussion 23.5

Explain why the wave function shown in Figure 23.5 is not normalisable. Why is this not physically acceptable?

Discussion 23.6

Explain why, if we also have a non-normalisable state if $E < V_{\min}$ and we start off with $\psi < 0$.

The result of Discussions 23.5 and 23.6 is that, for physically acceptable energy eigenstates, we can not have the energy less than the minimum of the potential.

Superposition states & time dependence

To predict future probabilities, we need to solve Schrödinger equation

$$\hat{H}|\psi\rangle = i\hbar \frac{d}{dt}|\psi\rangle . \quad (23.7)$$

We have already found the general solution

$$|\psi(t)\rangle = \sum_n c_n e^{-iE_n t/\hbar} |E_n\rangle \quad (23.8)$$

In other words, the energy eigenstates are a good basis in which to express time dependence. The phase factor of each component state depends on its energy. Note that we must use the energy basis to express things as in (23.8). This is one of the reasons for devoting a lot of time to finding energy eigenstates. There is one snag – to use (23.8) we need to know the coefficients c_n .

At $t = 0$ our state is

$$|\psi(0)\rangle = \sum_n c_n |E_n\rangle , \quad (23.9)$$

and the c_n are determined by the initial state of the system. The c_n are the probability amplitude for $|\psi(0)\rangle$ to be in the state $|E_n\rangle$. They can be calculated using

$$c_n = \langle E_n | \psi(0) \rangle , \quad (23.10)$$

since,

$$\begin{aligned} \langle E_n | \psi(0) \rangle &= \langle E_n | \sum_m c_m |E_m\rangle \\ &= \sum_m \langle E_n | c_m |E_m\rangle \\ &= \sum_m c_m \langle E_n | E_m\rangle \\ &= \sum_m c_m \delta_{nm} \\ &= c_n . \end{aligned}$$

Then at time $t = 0$,

$$\mathcal{P}_{E_n} = |\langle E_n | \psi(0) \rangle|^2 = |c_n|^2 \quad (23.11)$$

Although we've seen it before, we'll show again that the probabilities are time independent:

$$\begin{aligned}
 \mathcal{P}_{E_k} \text{ at time } t &= |\langle E_n | \psi(t) \rangle|^2 \\
 &= \left| \langle E_k | \sum_n c_n e^{-iE_n t / \hbar} | E_n \rangle \right|^2 \\
 &= \left| \sum_n c_n e^{-iE_n t / \hbar} \langle E_k | E_n \rangle \right|^2 \\
 &= \left| \sum_n c_n e^{-iE_n t / \hbar} \delta_{kn} \right|^2 = \left| c_k e^{-iE_k t / \hbar} \right|^2 \\
 &= |c_k|^2
 \end{aligned}$$

Since the probability of measuring a particular energy is time independent, the expectation value of the Hamiltonian is also time independent:

$$\langle \hat{H} \rangle = \sum_k \mathcal{P}_{E_k} E_k = \sum_k |c_k|^2 E_k \quad (23.12)$$

We can also show (23.12) directly using the time-dependent states:

$$\begin{aligned}
 \langle \hat{H} \rangle &= \langle \psi(t) | \hat{H} | \psi(t) \rangle \\
 &= \sum_m c_m^* \langle E_m | e^{iE_m t / \hbar} \hat{H} \sum_n c_n e^{-iE_n t / \hbar} | E_n \rangle \\
 &= \sum_{m,n} c_m^* c_n e^{iE_m t / \hbar} e^{-iE_n t / \hbar} \langle E_m | \hat{H} | E_n \rangle \\
 &= \sum_{m,n} c_m^* c_n e^{i(E_m - E_n)t / \hbar} E_n \langle E_m | E_n \rangle \\
 &= \sum_{m,n} c_m^* c_n e^{i(E_m - E_n)t / \hbar} E_n \delta_{mn} \\
 &= \sum_n c_n^* c_n E_n = \sum_n |c_n|^2 E_n
 \end{aligned}$$

Discussion 23.7

Write a summary of the important points and formulae in this chapter.

24

Superposition states

Interlude: plotting wave functions

This reading is based on [McIntyre et al. 2012, Section 5.7-10]

Consider the infinite square well ground state:

$$\psi(x, t) = \sqrt{\frac{2}{L}} \sin\left(\pi \frac{x}{L}\right) e^{-iE_1 t/\hbar} \quad (24.1)$$

This is a complex function so how can we plot it? One way is to plot the real and imaginary parts. Taking the real and imaginary parts of (24.1) we get

$$\text{Re}(\psi(x, t)) = \sqrt{\frac{2}{L}} \sin\left(\pi \frac{x}{L}\right) \cos(-E_1 t/\hbar) \quad (24.2)$$

and

$$\text{Im}(\psi(x, t)) = \sqrt{\frac{2}{L}} \sin\left(\pi \frac{x}{L}\right) \sin(-E_1 t/\hbar) \quad (24.3)$$

A plot of (24.2) and (24.3), using units of, $\frac{x}{L}$, on the horizontal axis, is shown in Figure 24.1.

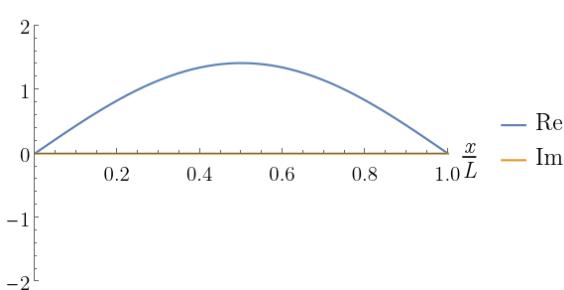


Figure 24.1: Real and Imaginary parts of the infinite square well ground state (24.1) at $t = 0$. You can click on the image to see an animated version, showing the time dependence.

As another example, a plot of the first excited state is shown in Figure 24.2.

Complex numbers can also be written as a modulus and a phase

$$\psi(x, t) = |\psi(x, t)| e^{i\theta}, \quad (24.4)$$

so for example, we can write the infinite square well ground state as:

$$\psi(x, t) = |\psi(x, t)| e^{i\theta} = \left| \sqrt{\frac{2}{L}} \sin\left(\pi \frac{x}{L}\right) \right| e^{-iE_1 t/\hbar} \quad (24.5)$$

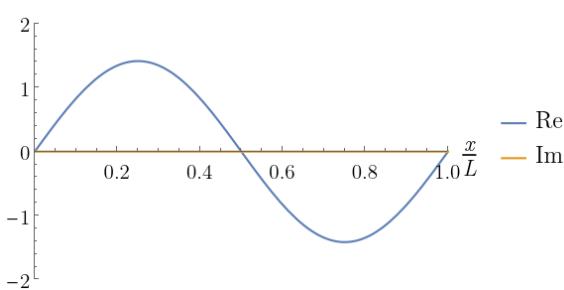


Figure 24.2: Real and Imaginary parts of the first excited infinite square well state at $t = 0$. You can click on the image to see an animated version, showing the time dependence.

This suggests another way to plot the wave function: as shown in Figure 24.3, we use height to represent the modulus and we use colour coding to show the phase. The key in the figure can be used to read off the phase – so for example in Figure 24.3, the red colour indicates a phase of zero, ie. $e^{i\pi 0}$ at $t = 0$.

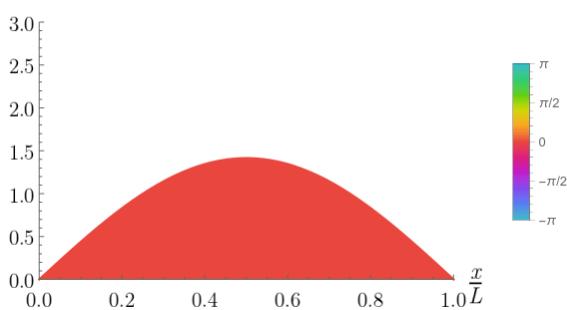


Figure 24.3: Modulus-phase plot the infinite square well ground state (24.5) at $t = 0$. You can click on the image to see an animated version, showing the time dependence.

The ground state is a stationary since $|\psi(x, t)|$ is constant and the phase does not depend on position. As you can see from plots of the first and second excited states in Figures 24.4-24.5, the phase changes by $i\pi$ when we cross a node. At $t = 0$, the phase change of $e^{i\pi} = -1$, corresponds to a sign flip.

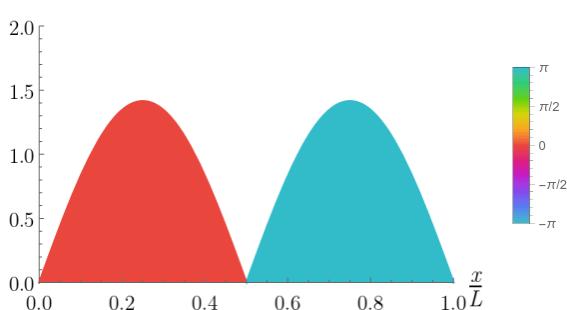


Figure 24.4: Modulus-phase plot the infinite square well first excited state at $t = 0$. You can click on the image to see an animated version, showing the time dependence.

Now, since the phase is proportional to, Et , the higher the energy, the faster the phase changes. You can view a [comparison of the ground and first excited state here](#).

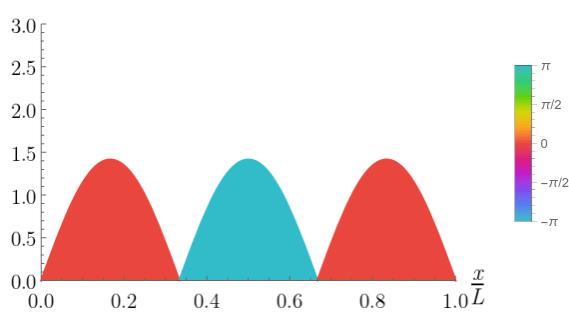


Figure 24.5: Modulus-phase plot the infinite square well second excited state at $t = 0$. You can click on the image to see an animated version, showing the time dependence, on the course site.

General solution

The general solution to the infinite square well potential can be written

$$|\psi(t)\rangle = \sum_n c_n e^{-iE_n t/\hbar} |E_n\rangle , \quad (24.6)$$

where we found that, $c_n = \langle E_n | \psi(0) \rangle$, which in position space becomes

$$c_n = \int_{-\infty}^{\infty} \varphi_n^* \psi(x, 0) dx \quad (24.7)$$

Let's show (24.7) directly in position space. Starting with

$$\begin{aligned} \int_{-\infty}^{\infty} \varphi_n^* \psi(x, 0) dx &= \int_{-\infty}^{\infty} \varphi_n^* \sum_m c_m \varphi_m dx \\ &= \sum_m c_m \int_{-\infty}^{\infty} \varphi_n^* \varphi_m dx \\ &= \sum_m c_m \delta_{nm} \\ &= c_n \end{aligned} \quad (24.8)$$

Example 24.1: Finding $\psi(x, t)$ given $\psi(x, 0)$

Problem A particle in an infinite square well, at $t = 0$, has the wave function

$$\begin{aligned} \psi(x, 0)_{\text{inside}} &= A \left(\frac{x}{L} \right)^2 \left(1 - \frac{x}{L} \right) \\ \psi(x, 0)_{\text{outside}} &= 0 \end{aligned} \quad (24.9)$$

1. What is the wave function at a later time?
2. What is the probability of energy measurements ?
3. What is $\langle \hat{H} \rangle$?

Solution The first thing we need to do is to normalise the wave function:

$$\begin{aligned}\langle \psi | \psi \rangle &= 1 = \int_0^L |\psi(x, 0)|^2 dx \\ &= |A|^2 \int_0^L \left(\frac{x}{L}\right)^4 \left(1 - \frac{x}{L}\right)^2 dx \\ &= |A|^2 \frac{L}{105}.\end{aligned}\quad (24.10)$$

💬 Discussion 24.1

Fill in the missing details in (24.10).

Choosing, A , to be real and positive, we find the initial wave function inside the well

$$\psi(x, 0) = \sqrt{\frac{105}{L}} \left(\frac{x}{L}\right)^2 \left(1 - \frac{x}{L}\right)\quad (24.11)$$

To find the c_n , we need to evaluate

$$\begin{aligned}c_n &= \langle E_n | \psi(0) \rangle = \int_{-\infty}^{\infty} \varphi_n^* \psi(x, 0) dx \\ &= \int_0^L \sqrt{\frac{2}{L}} \sin\left(n\pi \frac{x}{L}\right) \sqrt{\frac{105}{L}} \left(\frac{x}{L}\right)^2 \left(1 - \frac{x}{L}\right) dx.\end{aligned}\quad (24.12)$$

Doing the integral we find,

$$\begin{aligned}c_n &= \sqrt{210} \left[\left(\frac{(2(x/L) - 3(x/L)^2)}{(n\pi)^2} + \frac{6}{(n\pi)^4} \right) \sin\left(\frac{n\pi x}{L}\right) \right. \\ &\quad \left. + \left(\frac{(x/L)^3 - (x/L)^2}{n\pi} + \frac{2 - 6(x/L)}{(n\pi)^3} \right) \cos\left(\frac{n\pi x}{L}\right) \right]_0^L\end{aligned}\quad (24.13)$$

💬 Discussion 24.2

Given (24.12), do the integral & show (24.13).

You may use the integrals in [McIntyre

et al. 2012, Appendix F] and the fact that

$$\int u^3 \sin(au) du = \frac{3(a^2 u^2 - 2) \sin(au)}{a^4} - \frac{u(a^2 u^2 - 6) \cos(au)}{a^3} + C$$

Once we evaluate the limits of the integrals, the expression for the c_n 's can be simplified to

$$\begin{aligned}c_n &= \frac{(4(-1)^{n+1} - 2)\sqrt{210}}{(n\pi)^3} \\ &= \begin{cases} \frac{2\sqrt{210}}{(n\pi)^3}, & \text{if } n \text{ is odd} \\ -\frac{6\sqrt{210}}{(n\pi)^3}, & \text{if } n \text{ is even} \end{cases}\end{aligned}\quad (24.14)$$

Discussion 24.3

Given (24.13), evaluate the limits and simplify showing (24.14).

Evaluating the first few coefficients from (24.14), one finds, to three decimal places, $c_1 = 0.935$, $c_2 = -0.351$, $c_3 = 0.035$ and $c_4 = -0.044$. From this list we see that $\psi(x, 0)$ is composed mostly of ground state. Indeed in Figure 24.6 (a), we see that $c_1\varphi_1$ roughly approximates $\psi(x, 0)$. Adding more coefficients gets us closer to $\psi(x, 0)$ as we can see from Figure 24.6 (b) which shows $c_1\varphi_1 + c_2\varphi_2$. Once we have 8 terms, as in Figure 24.6 (d), the two are virtually indistinguishable by eye.

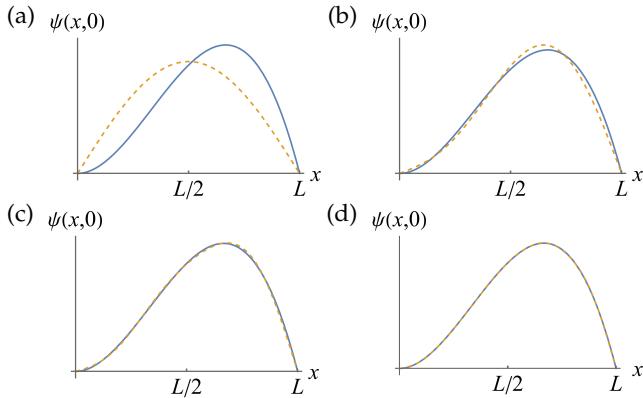


Figure 24.6: Plots of $\psi(x, 0)$ (solid blue line) vs $\sum_{n=1}^{n_{max}} c_n \varphi_n$ (orange dashed lines). In (a), $n_{max} = 1$, in (b) $n_{max} = 2$, in (c) $n_{max} = 4$, and in (d) $n_{max} = 8$.

Having expressed, $\psi(x, 0)$ in terms of the energy eigenfunctions, φ_n , we can find the wave function at later times by folding in the time dependence:

$$\begin{aligned} \psi(x, t) &= \sum_{n=1}^{\infty} c_n \varphi_n(x) e^{-iE_n t/\hbar} = \sum_{n=1}^{\infty} c_n \sqrt{\frac{2}{L}} \sin\left(n\pi \frac{x}{L}\right) e^{-in^2\pi^2\hbar t/2mL^2} \\ &= 2\sqrt{\frac{105}{L}} \sum_{n=1}^{\infty} \frac{4(-1)^{n+1} - 2}{(n\pi)^3} \sin\left(n\pi \frac{x}{L}\right) e^{-in^2\pi^2\hbar t/2mL^2} \end{aligned} \quad (24.15)$$

A plot of the real and imaginary parts of (24.15) (actually only 10 terms) is shown in Figure 24.7.

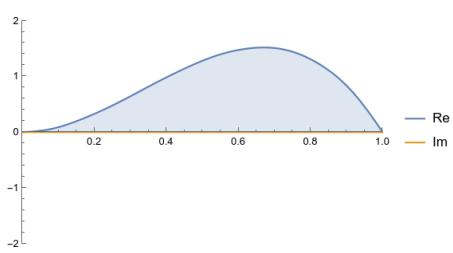


Figure 24.7: Real and Imaginary parts of (24.15) at $t = 0$. You can click on it to see an animated version.

Alternatively we can plot the modulus and phase as shown in Figure 24.8.

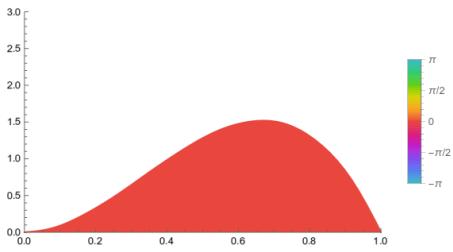


Figure 24.8: Modulus phase plot of (24.15) at $t = 0$. You can click on it to see an animated version.

Now, the expectation value of the total energy or Hamiltonian is

$$\begin{aligned}
 \langle \hat{H} \rangle &= \int \psi^* \hat{H} \psi \, dx \\
 &= \int_0^L \sqrt{\frac{105}{L}} \left(\frac{x}{L}\right)^2 \left(1 - \frac{x}{L}\right) \left(\frac{-\hbar^2}{2m} \frac{d^2}{dx^2}\right) \sqrt{\frac{105}{L}} \left(\frac{x}{L}\right)^2 \left(1 - \frac{x}{L}\right) \, dx \\
 &= \frac{105}{L} \left(\frac{-\hbar^2}{2m}\right) \int_0^L \left(\frac{x}{L}\right)^2 \left(1 - \frac{x}{L}\right) \left(\frac{2}{L^2}\right) \left(1 - \frac{3x}{L}\right) \, dx \\
 &= \frac{210}{L^2} \left(\frac{-\hbar^2}{2m}\right) \int_0^1 u^2 (1-u)(1-3u) \, du \\
 &= \left(\frac{\hbar^2}{2m}\right) \left(\frac{14}{L^2}\right) = \frac{14}{\pi^2} E_1 \approx 1.42 E_1 .
 \end{aligned} \tag{24.16}$$

💬 Discussion 24.4

Explain how we get the 3rd, 4th and 5th equalities in (24.16).

The probability of measuring a particular energy, given by

square of the expansion coefficient:

$$\begin{aligned}\mathcal{P}_{E_n} &= |\langle E_n | \psi(t) \rangle|^2 = |c_n|^2 \\ &= \frac{210}{(n\pi)^6} (4(-1)^{n+1} - 2)^2 \\ &= \frac{840}{(n\pi)^6} (5 + 4(-1)^n)\end{aligned}\quad (24.17)$$

As a check, we can also use (24.17) to find the expectation value of the energy.

$$\begin{aligned}\langle \hat{H} \rangle &= \sum_n \mathcal{P}_{E_n} E_n = \sum_n |c_n|^2 E_n \\ &= \sum_{n=1}^{\infty} \frac{840}{(n\pi)^6} (5 + 4(-1)^n) \left(\frac{n^2 \pi^2 \hbar^2}{2mL^2} \right) \\ &= \sum_{n=1,3,5,\dots}^{\infty} \frac{840}{(n\pi)^6} \left(\frac{n^2 \pi^2 \hbar^2}{2mL^2} \right) + \sum_{n=2,4,6,\dots}^{\infty} \frac{9(840)}{(n\pi)^6} \left(\frac{n^2 \pi^2 \hbar^2}{2mL^2} \right) \\ &= \frac{420\hbar^2}{\pi^4 mL^2} \left(\sum_{n=1,3,5,\dots}^{\infty} \frac{1}{n^4} + 9 \sum_{n=2,4,6,\dots}^{\infty} \frac{1}{n^4} \right) \\ &= \frac{420\hbar^2}{\pi^4 mL^2} \left(\frac{\pi^4}{96} + 9 \frac{\pi^4}{1440} \right) \\ &= 7 \frac{\hbar^2}{mL^2} = \frac{14}{\pi^2} E_1 \approx 1.42 E_1.\end{aligned}\quad (24.18)$$

The expectation value of energy we found in Example 24.1 is time-independent. On the other hand, $\langle \hat{x} \rangle$ and $\langle \hat{p} \rangle$, which are time-independent for energy eigenstates, are time dependent for a general superposition state. calculating $\langle \hat{x} \rangle$ and $\langle \hat{p} \rangle$ for the state in Example 24.1 is quite complicated so lets look at a simpler case, namely, superposition of two states in an infinite square well:

$$|\psi(0)\rangle = \frac{1}{\sqrt{2}}|E_1\rangle + \frac{1}{\sqrt{2}}|E_2\rangle \quad (24.19)$$

Using our usual prescription, at a time t , we have:

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}}|E_1\rangle e^{-iE_1 t/\hbar} + \frac{1}{\sqrt{2}}|E_2\rangle e^{-iE_2 t/\hbar}, \quad (24.20)$$

which in the wave function representation is

$$\begin{aligned}\psi(x, t) &= \frac{1}{\sqrt{2}}\varphi_1(x)e^{-iE_1 t/\hbar} + \frac{1}{\sqrt{2}}\varphi_2(x)e^{-iE_2 t/\hbar} \\ &= \sqrt{\frac{1}{L}} \left(\sin\left(\frac{\pi x}{L}\right) e^{-iE_1 t/\hbar} + \sin\left(\frac{2\pi x}{L}\right) e^{-iE_2 t/\hbar} \right)\end{aligned}\quad (24.21)$$

A plot of the imaginary and real part of (24.21) as well as an amplitude and phase plot are shown in Figures 24.9-24.10.

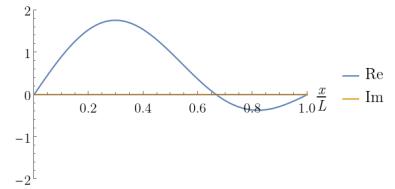


Figure 24.9: Real and imaginary parts of (24.21). Click on it for an animated version

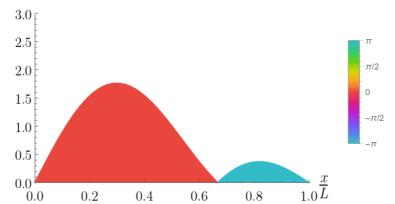


Figure 24.10: Modulus phase plot of (24.21). Click on it for an animated version

Now lets find

$$\begin{aligned}
 \langle \hat{x} \rangle &= \langle \psi(t) | \hat{x} | \psi(t) \rangle \\
 &= \left(\frac{1}{\sqrt{2}} \langle E_1 | e^{iE_1 t/\hbar} + \frac{1}{\sqrt{2}} \langle E_2 | e^{iE_2 t/\hbar} \right) \hat{x} \left(\frac{1}{\sqrt{2}} |E_1\rangle e^{-iE_1 t/\hbar} + \frac{1}{\sqrt{2}} |E_2\rangle e^{-iE_2 t/\hbar} \right) \\
 &= \frac{1}{2} \left(\langle E_1 | \hat{x} | E_1 \rangle + \langle E_2 | \hat{x} | E_2 \rangle + \langle E_1 | \hat{x} | E_2 \rangle e^{\frac{i(E_1 - E_2)t}{\hbar}} + \langle E_2 | \hat{x} | E_1 \rangle e^{\frac{-i(E_1 - E_2)t}{\hbar}} \right)
 \end{aligned} \tag{24.22}$$

Notice that we have been using Dirac notation which saves us some writing (as compared to expressing things in position space). To evaluate, (24.22) we are going to need certain matrix elements. Lets define:

$$\begin{aligned}
 \langle \hat{x} \rangle_n &= \langle E_n | \hat{x} | E_n \rangle = \int_0^L \varphi_n^*(x) x \varphi_n(x) dx = \int_0^L x |\varphi_n(x)|^2 dx \\
 \langle \hat{x} \rangle_{nk} &= \langle E_n | \hat{x} | E_k \rangle = \int_0^L \varphi_n^*(x) x \varphi_k(x) dx \quad (\text{where } n \neq k).
 \end{aligned} \tag{24.23}$$

We calculated, $\langle \hat{x} \rangle_n$, which is the expectation value of position before and found that it is $L/2$, the midpoint of the well. The second integral is

$$\begin{aligned}
 \langle \hat{x} \rangle_{nk} &= \int_0^L \varphi_n^*(x) x \varphi_k(x) dx \\
 &= \frac{2}{L} \int_0^L \sin\left(\frac{n\pi x}{L}\right) x \sin\left(\frac{k\pi x}{L}\right) dx \\
 &= \frac{2}{L} \left(\frac{L}{\pi}\right)^2 \int_0^\pi u \sin(nu) \sin(ku) du ,
 \end{aligned} \tag{24.24}$$

which can be simplified using a trigonometric identity,

$$\begin{aligned}
 \langle \hat{x} \rangle_{nk} &= \frac{2}{L} \left(\frac{L}{\pi}\right)^2 \int_0^\pi u \frac{1}{2} \{ \cos((n-k)u) - \cos((n+k)u) \} du \\
 &= \frac{1}{L} \left(\frac{L}{\pi}\right)^2 \left[\frac{\cos((n-k)\pi)}{(n-k)^2} + \frac{u \sin((n-k)u)}{(n-k)} - \frac{\cos((n+k)\pi)}{(n+k)^2} - \frac{u \sin((n+k)u)}{(n+k)} \right]_0^\pi \\
 &= \frac{1}{L} \left(\frac{L}{\pi}\right)^2 \left(\frac{\cos((n-k)\pi)}{(n-k)^2} - \frac{\cos((n+k)\pi)}{(n+k)^2} - \frac{1}{(n-k)^2} + \frac{1}{(n+k)^2} \right)
 \end{aligned} \tag{24.25}$$

which simplifies to

$$\langle \hat{x} \rangle_{nk} = \frac{-4Lnk}{\pi^2(n^2 - k^2)^2} \left(1 - (-1)^{n+k} \right). \tag{24.26}$$

⊕ Discussion 24.5

Show that the last line of (24.25) simplifies to (24.26) where n and k are positive integers.

Notice that (24.26) is zero for states where $n+k$ is even – i.e. states with the same parity – can you see why?

Specialising to our two state example, we have

$$\begin{aligned}\langle \hat{x} \rangle_1 &= \langle \hat{x} \rangle_2 = \frac{L}{2} \\ \langle \hat{x} \rangle_{12} &= \langle \hat{x} \rangle_{21} = -\frac{16L}{9\pi^2}.\end{aligned}\quad (24.27)$$

So finally we get

$$\begin{aligned}\langle \hat{x} \rangle &= \langle \psi(t) | \hat{x} | \psi(t) \rangle \\ &= \frac{1}{2} \left(\frac{L}{2} + \frac{L}{2} - \frac{16L}{9\pi^2} e^{i(E_1-E_2)t/\hbar} - \frac{16L}{9\pi^2} e^{-i(E_1-E_2)t/\hbar} \right) \\ &= \frac{L}{2} \left(1 - \frac{32}{9\pi^2} \cos \left(\frac{3\pi^2\hbar}{2mL^2} t \right) \right)\end{aligned}\quad (24.28)$$

Notice that (24.28) oscillates with the Bohr frequency $(E_2 - E_1)/\hbar$.

The probability density is also oscillating:

$$\begin{aligned}\mathcal{P}(x, t) &= |\langle x | \psi(t) \rangle|^2 = |\psi(x, t)|^2 \\ &= \left| \sqrt{\frac{1}{L}} \left(\sin \left(\frac{\pi x}{L} \right) e^{-iE_1 t/\hbar} + \sin \left(\frac{2\pi x}{L} \right) e^{-iE_2 t/\hbar} \right) \right|^2 \\ &= \frac{1}{L} \left(\sin^2 \left(\frac{\pi x}{L} \right) + \sin^2 \left(\frac{2\pi x}{L} \right) + 2 \sin^2 \left(\frac{\pi x}{L} \right) \sin^2 \left(\frac{2\pi x}{L} \right) \cos \left(\frac{E_2 - E_1}{\hbar} t \right) \right),\end{aligned}\quad (24.29)$$

which also oscillates with Bohr frequency $(E_2 - E_1)/\hbar$.

A similar calculation [McIntyre et al. 2012, Problem 5.27] gives

$$\begin{aligned}\langle \hat{p} \rangle &= \langle \psi(t) | \hat{p} | \psi(t) \rangle \\ &= \int_0^L \psi(x, t)^* \left(\frac{\hbar}{i} \frac{d}{dx} \right) \psi(x, t) dx \\ &= \frac{8}{3} \frac{\hbar}{L} \sin \left(\frac{3\pi^2\hbar}{2mL^2} t \right)\end{aligned}\quad (24.30)$$

Using (24.28) and (24.30), we can show an example of Ehrenfest's theorem

$$\langle \hat{p} \rangle = m \frac{d\langle \hat{x} \rangle}{dt}\quad (24.31)$$

Discussion 24.6

Show that (24.28) and (24.30) satisfy an example of Ehrenfest's theorem, (24.31).

Discussion 24.7

Write a summary of the important points and formulae in this chapter.

25

Harmonic Oscillator I

The classical harmonic oscillator consists of a mass, m , on a spring with a force constant k . Hooke's law together with Newton's second law result in the equation of motion

$$F = -kx = m \frac{d^2x}{dt^2}, \quad (25.1)$$

which has the solution

$$x(t) = A \sin(\omega t) + B \cos(\omega t), \quad (25.2)$$

where $\omega = \sqrt{k/m}$. It is not hard to see that taking a potential of the form

$$V = \frac{1}{2}kx^2, \quad (25.3)$$

leads to (25.1).

Discussion 25.1

Show that the potential (25.3) results in the equation of motion (25.1).

While there is no such thing as a perfect spring – if the extension of any spring is too large it will break – almost any potential near a minimum is approximately parabolic. To see this, consider the Taylor series of some potential $V(x)$ about a minimum at some point, x_0 :

$$\begin{aligned} V(x) &= V(x_0 + (x - x_0)) \\ &= V(x_0) + V'(x_0)(x - x_0) + \frac{1}{2}V''(x_0)(x - x_0)^2 + \dots \end{aligned} \quad (25.4)$$

Discussion 25.2

Show that shifting a potential by a constant, ie. $V \rightarrow V + c$, does not affect the force.

We can subtract $V(x_0)$ from our potential since shifting $V(x)$ doesn't change the force and since we are at a minimum $V'(x_0) = 0$. If we are near the minimum, $(x - x_0)$ will be small and we can neglect the higher order terms¹ in (25.4). Putting all these points

This reading is based on [Griffiths 2016, Section 2.3]

¹ More precisely we need

$$\begin{aligned} \frac{1}{2!}V''(x_0)(x - x_0)^2 &>> \frac{1}{3!}V'''(x_0)(x - x_0)^3 \\ \Rightarrow (x - x_0) &<< \frac{3V''(x_0)}{V'''(x_0)} \end{aligned}$$

and in general

$$(x - x_0) << \frac{(n+1)V_0^{(n)}(x_0)}{V_0^{(n+1)}(x_0)} \quad \forall n > 1$$

together, we can take our potential to be

$$V(x) \approx \frac{1}{2} V''(x_0)(x - x_0)^2. \quad (25.5)$$

The fact that small oscillations are almost always governed by the harmonic oscillator makes it a very important problem. To study the quantum harmonic oscillator, we consider Schrödinger's equation with the potential (25.3). It turns out to more convenient to write the potential in terms of the angular frequency ω rather than k :

$$V = \frac{1}{2}m\omega^2x^2. \quad (25.6)$$

Discussion 25.3

Show that (25.3) can be written as (25.6).

Now substituting (25.6) into the time-independent Schrödinger equations leads to

$$\frac{-\hbar^2}{2m} \frac{d^2\varphi}{dx^2} + \frac{1}{2}m\omega^2x^2\varphi = E\varphi. \quad (25.7)$$

There are two main methods to solve (25.7). The first method is a brute force power series approach. The advantage of this method, is that it can be applied to many other potentials but it can be tricky. The second method, called the algebraic method, is easier but unfortunately harder to generalise. We will study both methods.

25.1 Algebraic method

We can rewrite the time-independent Schrödinger equation, (25.7), as

$$\hat{H}|E\rangle = \frac{1}{2m} (\hat{p}^2 + (m\omega\hat{x})^2)|E\rangle = E|E\rangle, \quad (25.8)$$

where as usual \hat{H} is the Hamiltonian operator, $\hat{p} \doteq \frac{\hbar}{i} \frac{d}{dx}$ is the momentum operator and $\hat{x} \doteq x$ is the position operator.

Discussion 25.4

Show that (25.7) can be rewritten as (25.8).

The basic idea of the operator method is to try factorise \hat{H} which may simplify the problem. If we were just dealing with numbers this would be simple

$$u^2 + v^2 = (u + iv)(u - iv), \quad (25.9)$$

but \hat{x} and \hat{p} are not just numbers, they are operators. In particular, unlike numbers \hat{x} and \hat{p} do not commute, that is to say $\hat{x}\hat{p}$ is not the same as $\hat{p}\hat{x}$.

Even though \hat{x} and \hat{p} do not commute, we'd like to see how close we can get to writing (25.7) in the form (25.9). To this end let

$$\hat{a}_+ = \frac{1}{\sqrt{2\hbar m\omega}} (m\omega\hat{x} - i\hat{p}), \quad (25.10)$$

$$\hat{a}_- = \frac{1}{\sqrt{2\hbar m\omega}} (m\omega\hat{x} + i\hat{p}). \quad (25.11)$$

The overall factor in the definition of \hat{a}_\pm , namely $\frac{1}{\sqrt{2\hbar m\omega}}$, is chosen with the benefit of hindsight to make things look nicer in the end.

Now

$$\begin{aligned}\hat{a}_-\hat{a}_+ &= \frac{1}{2\hbar m\omega} (m\omega\hat{x} + i\hat{p})(m\omega\hat{x} - i\hat{p}) \\ &= \frac{1}{2\hbar m\omega} (\hat{p}^2 + (m\omega\hat{x})^2 - im\omega(\hat{x}\hat{p} - \hat{p}\hat{x})) \\ &= \frac{\hat{H}}{\hbar\omega} - \frac{i}{2\hbar} (\hat{x}\hat{p} - \hat{p}\hat{x}).\end{aligned}\quad (25.12)$$

Discussion 25.5

Show how we get from the second last to the last line of (25.12).

Notice that the extra term in (25.12), is proportional the commutator of \hat{x} and \hat{p} , and we can rewrite (25.12) more compactly as

$$\hat{a}_-\hat{a}_+ = \frac{\hat{H}}{\hbar\omega} - \frac{i}{2\hbar} [\hat{x}, \hat{p}]. \quad (25.13)$$

In a similar fashion, one also gets

$$\hat{a}_+\hat{a}_- = \frac{\hat{H}}{\hbar\omega} + \frac{i}{2\hbar} [\hat{x}, \hat{p}]. \quad (25.14)$$

Discussion 25.6

Obtain (25.14) from (25.11) and (25.10).

To calculate $[\hat{x}, \hat{p}]$, it is best to let it act on some function $f(x)$:

$$\begin{aligned}[\hat{x}, \hat{p}]f(x) &= x \frac{\hbar}{i} \frac{d}{dx} f - \frac{\hbar}{i} \frac{d}{dx} (xf) \\ &= \frac{\hbar}{i} \left(x \frac{df}{dx} - x \frac{df}{dx} - f \right) \\ &= i\hbar f(x).\end{aligned}\quad (25.15)$$

Since $f(x)$ was an arbitrary function, we conclude that, as operators

$$[\hat{x}, \hat{p}] = i\hbar. \quad (25.16)$$

so substituting (25.16) in (25.13) and (25.14) gives

$$\begin{aligned}\hat{a}_-\hat{a}_+ &= \frac{\hat{H}}{\hbar\omega} + \frac{1}{2}, \\ \hat{a}_+\hat{a}_- &= \frac{\hat{H}}{\hbar\omega} - \frac{1}{2}.\end{aligned}\quad (25.17)$$

From (25.17) we see that

$$[\hat{a}_-, \hat{a}_+] = 1 . \quad (25.18)$$

Discussion 25.7

Show how (25.18) follows from (25.17).

Then solving (25.17) for \hat{H} gives

$$\begin{aligned} \hat{H} &= \hbar\omega (\hat{a}_- \hat{a}_+ - \frac{1}{2}) \\ &= \hbar\omega (\hat{a}_+ \hat{a}_- + \frac{1}{2}) . \end{aligned} \quad (25.19)$$

Discussion 25.8

Show that if \hat{x} and \hat{p} **did** commute, (25.10) and (25.11) would imply that

$$\hat{H} = \hbar\omega \hat{a}_+ \hat{a}_- = \hbar\omega \hat{a}_- \hat{a}_+ .$$

Comparing (25.19) with the result of Discussion 25.8, we see that there is an extra factor of $\frac{1}{2}$ due to the non-commutativity of \hat{a}_\pm . Finally we find that (25.8) can be written

$$\begin{aligned} \hbar\omega (\hat{a}_- \hat{a}_+ - \frac{1}{2}) |E\rangle &= E |E\rangle \\ \hbar\omega (\hat{a}_+ \hat{a}_- + \frac{1}{2}) |E\rangle &= E |E\rangle , \end{aligned} \quad (25.20)$$

or more compactly as

$$\hbar\omega (\hat{a}_\pm \hat{a}_\mp \pm \frac{1}{2}) |E\rangle = E |E\rangle . \quad (25.21)$$

Now

$$\begin{aligned} (\hat{a}_+)^{\dagger} &= \left(\frac{1}{\sqrt{2\hbar m\omega}} (m\omega \hat{x} - i\hat{p}) \right)^{\dagger} \\ &= \frac{1}{\sqrt{2\hbar m\omega}} (m\omega \hat{x} + i\hat{p}) \\ &= \hat{a}_- \end{aligned} \quad (25.22)$$

Discussion 25.9

Explain the steps in (25.22) and show that $(\hat{a}_-)^{\dagger} = \hat{a}_+$.

We conclude that \hat{a}_+ and \hat{a}_- are hermitian conjugates of each other:

Hermitian conjugates

$$(\hat{a}_\pm)^{\dagger} = \hat{a}_\mp \quad (25.23)$$

In the next chapter, we'll see how we can use the properties of \hat{a}_\pm to solve (25.20) and to find the allowed values of E .

 **Discussion 25.10**

Write a summary of the important points and formulae in this chapter.

26

Harmonic Oscillator II

Algebraic method (continued)

This reading is based on [Griffiths 2016, Section 2.3]

In the last chapter we laid the ground work for solving the quantum harmonic oscillator using the algebraic method. Now we'll see that operators, \hat{a}_+ and \hat{a}_- , allow us to move from one energy eigenstate to another. Suppose we have a state, $|E\rangle$ which satisfies the energy equation for the harmonic oscillator, (25.7), with energy E so that:

$$\hat{H}|E\rangle = E|E\rangle . \quad (26.1)$$

Then using \hat{a}_+ we can generate a new energy eigenstate state, $|E'\rangle = \hat{a}_+|E\rangle$, with energy eigenvalue $E' = E + \hbar\omega$. For this reason we call \hat{a}_+ a **raising operator**. We can also use \hat{a}_- to produce a new energy eigenstate state, $\hat{a}_-|E\rangle$, with energy eigenvalue $E - \hbar\omega$ and we call \hat{a}_- a **lowering operator**. Before seeing how this works, we need to learn about some general properties of operators as well as special features of the operators \hat{a}_\pm discussed below.

Discussion 26.1

Show that

$$[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B} \quad (26.2)$$

Hint: just write out the LHS and RHS of eq. (26.2) explicitly.

Discussion 26.2

Using eq. (26.2) and

$$\hat{H} = \hbar\omega\left(\hat{a}_+\hat{a}_- + \frac{1}{2}\right) \quad (26.3)$$

show that

$$\begin{aligned} [\hat{H}, \hat{a}_+] &= +\hbar\omega\hat{a}_+ \\ [\hat{H}, \hat{a}_-] &= -\hbar\omega\hat{a}_- \end{aligned} \quad (26.4)$$

Discussion 26.3

Show that for two operators \hat{A} and \hat{B} ,

$$\hat{A}\hat{B} = [\hat{A}, \hat{B}] + \hat{B}\hat{A}$$

We finally have the tools to see how the raising and lower operators work. Given an energy eigenstate, $|E\rangle$, with $\hat{H}|E\rangle = E|E\rangle$, we find

$$\begin{aligned}\hat{H}(\hat{a}_+|E\rangle) &= \hat{H}\hat{a}_+|E\rangle \stackrel{(26.3)}{=} ([\hat{H}, \hat{a}_+] + \hat{a}_+\hat{H})|E\rangle \\ &= (\hbar\omega\hat{a}_+ + \hat{a}_+E)|E\rangle \\ &= (E + \hbar\omega)(\hat{a}_+|E\rangle)\end{aligned}\quad (26.5)$$

Discussion 26.4

Explain the steps in eq. (26.5) and how we can now conclude that $\hat{a}_+|E\rangle$ is an energy eigenstate with energy eigenvalue $E + \hbar\omega$.

Similarly, one can show that if $\hat{H}|E\rangle = E|E\rangle$, $\hat{a}_-|E\rangle$ has energy $(E - \hbar\omega)$:

$$\hat{H}(\hat{a}_-|E\rangle) = (E - \hbar\omega)(\hat{a}_-|E\rangle). \quad (26.6)$$

Discussion 26.5

Prove (26.6).

At this point we come to a key stage in the story. Using the lowering operator, one can keep on creating new solutions with lower and lower energy – iterating (26.6) gives

$$\hat{H}(\hat{a}_-^n|E\rangle) = (E - n\hbar\omega)(\hat{a}_-^n|E\rangle). \quad (26.7)$$

At some point we'll generate a solution with $E < V_{min}$ where V_{min} is the minimum value of our potential which in this case, (25.6), is zero. We previously showed normalisable solutions must have $E > V_{min}$. That means that at some point, after repeatedly applying \hat{a}_- , we reach a solution, $|\circlearrowleft\rangle = \hat{a}_-|E_0\rangle$, which is not normalizable. Now, $|\circlearrowleft\rangle$, can be non-normalizable in two ways – either it is zero or $\langle\circlearrowleft|\circlearrowleft\rangle = \infty$. It turns out that, in practise, $|\circlearrowleft\rangle = 0$, so that to find the state with the lowest energy, $|E_0\rangle$, (ie. the ground state¹) we need to solve

$$0 = \hat{a}_-|E_0\rangle \stackrel{(25.11)}{=} \frac{1}{\sqrt{2\hbar m\omega}}(m\omega\hat{x} + i\hat{p})|E_0\rangle, \quad (26.8)$$

which in position representation becomes

$$0 = \frac{1}{\sqrt{2\hbar m\omega}}\left(m\omega x + i\left(\frac{\hbar}{i}\right)\frac{d}{dx}\right)\varphi_0(x) \quad (26.9)$$

¹ For the harmonic oscillator, we will label the ground state energy as E_0 rather than E_1 .

leading to the differential equation

$$\frac{d\varphi_0}{dx} = -\frac{m\omega}{\hbar}x\varphi_0(x), \quad (26.10)$$

Discussion 26.6

Before looking at how we solve (26.10), try use the curve sketching techniques you have learnt to draw the ground and first excited state of the harmonic oscillator.

Now, from (26.10) which we obtain

$$\begin{aligned} \int \frac{d\varphi_0}{\varphi_0} &= -\frac{m\omega}{\hbar} \int x dx \\ \implies \ln \varphi_0 &= -\frac{m\omega x^2}{2\hbar} + C \\ \implies \varphi_0 &= Ae^{-\frac{m\omega x^2}{2\hbar}}, \end{aligned} \quad (26.11)$$

where $A = e^C$. Normalising φ_0 requires

$$\begin{aligned} 1 &= |A|^2 \int_{-\infty}^{\infty} e^{-\frac{m\omega x^2}{\hbar}} dx \quad \text{let } u = \sqrt{\frac{m\omega}{\hbar}} x \\ &= |A|^2 \sqrt{\frac{\hbar}{m\omega}} \underbrace{\int_{-\infty}^{\infty} e^{-u^2} du}_{\sqrt{\pi}} \\ &= |A|^2 \sqrt{\frac{\pi\hbar}{m\omega}}, \end{aligned} \quad (26.12)$$

$$\begin{aligned} u &= \sqrt{\frac{m\omega}{\hbar}} x \implies x = \sqrt{\frac{\hbar}{m\omega}} u \\ x &= \sqrt{\frac{\hbar}{m\omega}} u \implies dx = \sqrt{\frac{\hbar}{m\omega}} du \\ x &= \sqrt{\frac{\hbar}{m\omega}} u \implies x^2 = \left(\sqrt{\frac{\hbar}{m\omega}}\right)^2 u^2 \end{aligned}$$

which implies we can take

$$A = \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}} \quad (26.13)$$

which leads to

$$\varphi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}} e^{-\frac{m\omega x^2}{2\hbar}}. \quad (26.14)$$

Finally we would like to know the energy of the ground state. To this end, consider

$$\begin{aligned} \hat{H}|E_0\rangle &\stackrel{(25.19)}{=} \hbar\omega(\hat{a}_+ \hat{a}_- + \frac{1}{2})|E_0\rangle \\ &= \hbar\omega \hat{a}_+ (\underbrace{\hat{a}_-|E_0\rangle}_{=0(26.8)}) + \frac{1}{2}\hbar\omega|E_0\rangle \\ &= \underbrace{\frac{1}{2}\hbar\omega|E_0\rangle}_{E_0}. \end{aligned} \quad (26.15)$$

Discussion 26.7

Qualitatively sketch the ground state $\varphi_0(x)$ given by chapter 26. Does it roughly agree with what you found in Discussion 26.6. If not, go back and workout where you went wrong.

To get the energies of the excited states, we iteratively apply \hat{a}_+ to φ_0 to get

$$\hat{H}(\hat{a}_+^n |E_0\rangle) = \underbrace{\hbar\omega(n + \frac{1}{2})}_{E_n} (\hat{a}_+^n |E_0\rangle), \quad (26.16)$$

which implies that

$$E_n = \hbar\omega(n + \frac{1}{2}) \quad n = 0, 1, 2, 3, \dots \quad (26.17)$$

Equation (26.17) is illustrated fig. 26.1 with the raising operator, \hat{a}_+ , moving us up the “ladder” of states in steps of $\hbar\omega$, and the lowering operator, \hat{a}_- , moving us down.

Notation

We will adopt the convention of denoting (normalised) energy eigenstates of the harmonic oscillator using just the quantum number n

$$|n\rangle = |E_n\rangle. \quad (26.18)$$

Discussion 26.8

Show that if $|\psi\rangle$ satisfies an energy eigenvalue equation, $\hat{H}|\psi\rangle = E|\psi\rangle$, then $|\psi'\rangle = \lambda|\psi\rangle$, where λ is a constant, is also an energy eigenstate with the same eigenvalue. Explain why a ket, $|\psi\rangle$, which happens to be an energy eigenstate, may be normalised.

Discussion 26.8 tells us that, while we know that the ket, $\hat{a}_+|0\rangle$, has the energy eigenvalue $\frac{1}{2}\hbar\omega + \hbar\omega = \frac{3}{2}\hbar\omega$ it may not be normalised. We conclude that, using the raising operator, \hat{a}_+ , to generate the first excited state $|1\rangle$ from $|0\rangle$ using the raising operator gives :

$$|1\rangle = N_1 \hat{a}_+ |0\rangle, \quad (26.19)$$

Discussion 26.9

Explain why we have to include the constant N_1 in (26.19).

Writing (26.19) in position space gives

$$\begin{aligned} \varphi_1 &= \frac{N_1}{\sqrt{2\hbar m\omega}} \left(-\hbar \frac{d}{dx} + m\omega x \right) \varphi_0 \\ &\stackrel{(26.14)}{=} \frac{N_1}{\sqrt{2\hbar m\omega}} \left(-\hbar \frac{d}{dx} + m\omega x \right) \left(\frac{m\omega}{\pi\hbar} \right)^{\frac{1}{4}} e^{-\frac{m\omega x^2}{2\hbar}} \\ &= \frac{N_1}{\sqrt{2\hbar m\omega}} \left(\frac{m\omega}{\pi\hbar} \right)^{\frac{1}{4}} (2m\omega x) e^{-\frac{m\omega x^2}{2\hbar}} \\ &= N_1 \left(\frac{m\omega}{\pi\hbar} \right)^{\frac{1}{4}} \sqrt{\frac{2m\omega}{\hbar}} x e^{-\frac{m\omega x^2}{2\hbar}}. \end{aligned} \quad (26.20)$$

Then, normalising φ_1 , we find

$$1 = \int_{-\infty}^{\infty} |\varphi_1|^2 dx = |N_1|^2 \sqrt{\frac{m\omega}{\pi\hbar}} \left(\frac{2m\omega}{\hbar} \right) \int_{-\infty}^{\infty} x^2 e^{-\frac{m\omega x^2}{2\hbar}} dx$$

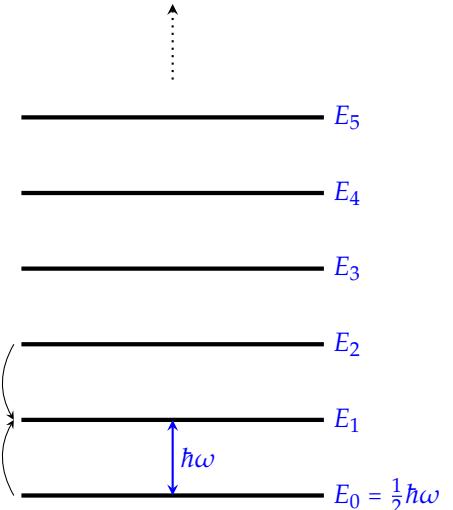


Figure 26.1: Allowed energies of the harmonic oscillator.

$$\begin{aligned}
&= |N_1|^2 \sqrt{\frac{m\omega}{\pi\hbar}} \left(\frac{2m\omega}{\hbar}\right) \left(\sqrt{\frac{\hbar}{m\omega}}\right)^3 \underbrace{\int_{-\infty}^{\infty} u^2 e^{-u^2} du}_{\sqrt{\pi}/2} \\
&= |N_1|^2 ,
\end{aligned} \tag{26.21}$$

so that $N_1 = 1$. Plots of the first three energy eigenstates are shown in Figure 26.2.

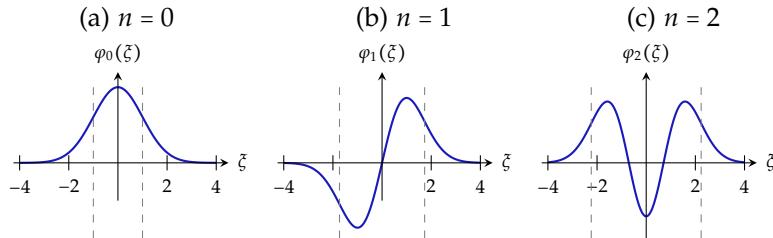


Figure 26.2: First three energy eigenstates of the Harmonic oscillator plotted as a function of the dimensionless variable $\xi = \sqrt{\frac{m\omega}{\hbar}} x$. The vertical dashed lines show the boundaries of the classically allowed region.

Discussion 26.10

What are the classically allowed regions for the first 3 states of the Harmonic oscillator? Check that your answers agree with the regions shown in Figure 26.2.

Discussion 26.11

Explain the features of the wave functions in Figure 26.2 using the curve sketching techniques you learnt in chapter 24.

While we had to work out ϕ_1 explicitly to get N_1 , it turns out we can get all the normalisation constants, N_n , without explicitly determining all the wave functions, ϕ_n , using a cunning trick. Since we know that

$$\begin{aligned}
\hat{a}_+ |n\rangle &\propto |n+1\rangle , \\
\hat{a}_- |n\rangle &\propto |n-1\rangle ,
\end{aligned} \tag{26.22}$$

define α_n and β_n as follows

$$\begin{aligned}
\hat{a}_+ |n\rangle &= \alpha_n |n+1\rangle , \\
\hat{a}_- |n\rangle &= \beta_n |n-1\rangle .
\end{aligned} \tag{26.23}$$

Discussion 26.12

Explain how (26.22) comes about. In other words, why can we just not assume that, $\hat{a}_+ |n\rangle = |n+1\rangle$, and, $\hat{a}_- |n\rangle = |n-1\rangle$?

Using (25.23), namely the fact that $(\hat{a}_\pm)^\dagger = \hat{a}_\mp$, we see that

$$\begin{aligned}
\langle \phi | \hat{a}_\pm \psi \rangle &= \langle (\hat{a}_\pm)^\dagger \phi | \psi \rangle \\
&= \langle \hat{a}_\mp \phi | \psi \rangle
\end{aligned} \tag{26.24}$$

❖ Discussion 26.13

Explain the steps in (26.24).

With (26.24) in hand we are ready to determine the N_n . Firstly,

$$\langle \hat{a}_{\pm} n | \hat{a}_{\pm} n \rangle \stackrel{(26.24)}{=} \langle \hat{a}_{\mp} \hat{a}_{\pm} n | n \rangle , \quad (26.25)$$

but combining (25.17) and (26.17) gives

$$\begin{aligned} \hat{a}_+ \hat{a}_- |n\rangle &\stackrel{(25.17)}{=} \left(\frac{\hat{H}}{\hbar\omega} - \frac{1}{2} \right) |n\rangle \\ &\stackrel{(26.17)}{=} \left(\frac{\hbar\omega(n + \frac{1}{2})}{\hbar\omega} - \frac{1}{2} \right) |n\rangle \\ &= n |n\rangle \end{aligned} \quad (26.26)$$

and similarly

$$\hat{a}_- \hat{a}_+ |n\rangle = (n + 1) |n\rangle \quad (26.27)$$

❖ Discussion 26.14

Explain the steps in (26.26) and show how we get (26.27).

Now, substituting (26.26) and (26.27) into (26.25) implies

$$\langle \hat{a}_+ n | \hat{a}_+ n \rangle = n + 1 , \quad (26.28)$$

$$\langle \hat{a}_- n | \hat{a}_- n \rangle = n . \quad (26.29)$$

❖ Discussion 26.15

Show how we can get (26.28) and (26.29) from (26.25), (26.26) and (26.27).

On the other hand from (26.23) we conclude that

$$\langle \hat{a}_+ n | \hat{a}_+ n \rangle = |\alpha_n|^2 , \quad (26.30)$$

$$\langle \hat{a}_- n | \hat{a}_- n \rangle = |\beta_n|^2 , \quad (26.31)$$

❖ Discussion 26.16

Use (26.23) to obtain (26.30) and (26.31).

Comparing (26.28)-(26.29) to (26.30)-(26.31), we can read off $\alpha_n = \sqrt{n+1}$ and $\beta_n = \sqrt{n}$ so that

$$\hat{a}_+ |n\rangle = \sqrt{n+1} |n+1\rangle \Rightarrow |n+1\rangle = \frac{1}{\sqrt{n+1}} \hat{a}_+ |n\rangle , \quad (26.32)$$

$$\hat{a}_- |n\rangle = \sqrt{n} |n-1\rangle . \quad (26.33)$$

Finally using (26.32) we get

$$|1\rangle = \hat{a}_+ |0\rangle = \frac{1}{\sqrt{1}} \hat{a}_+ |0\rangle$$

$$\begin{aligned} |2\rangle &= \frac{1}{\sqrt{2}} \hat{a}_+ |1\rangle = \frac{1}{\sqrt{2 \cdot 1}} (\hat{a}_+)^2 |0\rangle \\ |3\rangle &= \frac{1}{\sqrt{3}} \hat{a}_+ |2\rangle = \frac{1}{\sqrt{3 \cdot 2 \cdot 1}} (\hat{a}_+)^3 |0\rangle \\ |4\rangle &= \frac{1}{\sqrt{4}} \hat{a}_+ |3\rangle = \frac{1}{\sqrt{4 \cdot 3 \cdot 2 \cdot 1}} (\hat{a}_+)^4 |0\rangle \quad \text{etc...}, \end{aligned} \quad (26.34)$$

and one can show by induction that

$$|n\rangle = \frac{1}{\sqrt{n!}} (\hat{a}_+)^n |0\rangle , \quad (26.35)$$

or in other words that $N_n = 1/\sqrt{n!}$.

We can also now show that, just like the states for the infinite square well, the harmonic oscillator states are orthonormal

$$\langle m|n\rangle = \delta_{mn} . \quad (26.36)$$

Proof.

$$\langle m|\hat{a}_+\hat{a}_-n\rangle \stackrel{(26.26)}{=} n \langle m|n\rangle \quad (26.37)$$

but on the other hand

$$\begin{aligned} \langle m|\hat{a}_+\hat{a}_-n\rangle &\stackrel{(26.24)}{=} \langle \hat{a}_+\hat{a}_-m|n\rangle \\ &\stackrel{(26.26)}{=} m \langle m|n\rangle . \end{aligned} \quad (26.38)$$

Discussion 26.17

Explain how we use (26.24) twice to obtain the first line of (26.38).

Now comparing (26.37) and (26.38) we conclude that either $m = n$ or $\langle m|n\rangle = 0$ from which the result follows. \square

Discussion 26.18

Explain the last step of the proof of (26.36) above.

Having shown the orthonormality of the harmonic oscillator wave functions, we can use the same trick that we used for the infinite square well to express functions as a linear combination of the harmonic oscillator wave functions. Suppose we are given the wave function $\psi(x, 0)$ at $t = 0$, then using²

$$\psi(x, 0) = \sum_{n=0}^{\infty} c_n \varphi_n(x) , \quad (26.39)$$

then using (26.36) we can extract

$$c_m = \langle E_m|\psi\rangle = \int_{-\infty}^{\infty} \varphi_m^* \psi(x, 0) dx \quad (26.40)$$

and we can find $\psi(x, t)$ using the usual prescription

$$\psi(x, t) = \sum_{n=0}^{\infty} c_n \varphi_n(x) e^{-iE_n t/\hbar} . \quad (26.41)$$

² We will just assume that the solutions to the harmonic oscillator time independent Schrödinger equation are complete.

Once again $|c_n|^2$ will be the probability that a measurement of the energy returns the value E_n .

Finally with our new bag of tricks, we can calculate expectations values of x and p without even having explicit expressions for φ_n . The first step is to write \hat{x} and \hat{p} in terms of \hat{a}_\pm . Inverting (25.10) gives

$$\begin{aligned}\hat{x} &= \sqrt{\frac{\hbar}{2m\omega}}(\hat{a}_+ + \hat{a}_-) , \\ \hat{p} &= i\sqrt{\frac{\hbar m\omega}{2}}(\hat{a}_+ - \hat{a}_-) .\end{aligned}\quad (26.42)$$

💬 Discussion 26.19

Use (25.10) to derive (26.42).

Now, suppose we wish to calculate $\langle x^2 \rangle$ for the state $|n\rangle$, first note the

$$\hat{x}^2 = \frac{\hbar}{2m\omega} (\hat{a}_+^2 + \hat{a}_+ \hat{a}_- + \hat{a}_- \hat{a}_+ + \hat{a}_-^2) \quad (26.43)$$

so

$$\begin{aligned}\langle \hat{x}^2 \rangle &= \frac{\hbar}{2m\omega} \langle n | (\hat{a}_+^2 + \hat{a}_+ \hat{a}_- + \hat{a}_- \hat{a}_+ + \hat{a}_-^2) | n \rangle \\ &\stackrel{(26.26),(26.27),(26.32),(26.33)}{=} \frac{\hbar}{2m\omega} \langle n | \left(\sqrt{n+2} \sqrt{n+1} |n+2\rangle + n |n\rangle + (n+1) |n\rangle + \sqrt{n} \sqrt{n-1} |n-2\rangle \right) \\ &\stackrel{(26.36)}{=} \frac{\hbar}{m\omega} \left(n + \frac{1}{2} \right) .\end{aligned}\quad (26.44)$$

💬 Discussion 26.20

Fill in the missing details in (26.44).

Notice that as expected for the stationary state $|n\rangle$, $\langle \hat{x}^2 \rangle$, does not depend on time.

💬 Discussion 26.21

Write a summary of the important points and formulae in this chapter.

27

Harmonic Oscillator III

In this chapter we will solve the quantum harmonic oscillator using the series expansion method. Starting with

$$\frac{-\hbar^2}{2m} \frac{d^2\varphi}{dx^2} + \frac{1}{2}m\omega^2x^2\varphi = E\varphi , \quad (25.7')$$

it is convenient to clean things up by defining the dimensionless variables¹

$$\xi := \sqrt{\frac{m\omega}{\hbar}}x \quad \text{and} \quad K := \frac{2E}{\hbar\omega} , \quad (27.1)$$

you should find that (25.7) becomes

$$\frac{d^2\varphi}{d\xi^2} = (\xi^2 - K)\varphi . \quad (27.2)$$

Discussion 27.1

Show that using (27.1), (25.7) becomes (27.2).

Our task is now to find which values of K will lead to normalisable solutions. The first step is to factor out the asymptotic behaviour² which will simplify the power series expansions. When $|\xi| \gg K$, we get

$$\frac{d^2\varphi}{d\xi^2} \approx \xi^2\varphi , \quad (27.3)$$

which has the approximate solution

$$\varphi(\xi) \approx Ae^{-\xi^2/2} + Be^{\xi^2/2} . \quad (27.4)$$

To check that (27.4) approximately solves (27.3), notice that

$$\begin{aligned} \frac{d^2\varphi}{d\xi^2} &= A(\xi^2 - 1)e^{-\xi^2/2} + B(\xi^2 + 1)e^{\xi^2/2} \\ &\approx \xi^2\varphi . \end{aligned} \quad (27.5)$$

This reading is based on [Griffiths 2016, Section 2.3.2]

¹ When dealing with differential equations, it is often convenient to move to dimensionless variables.

² The asymptotic behaviour is how φ behaves for $|\xi|$ large.

Discussion 27.2

Fill in the missing steps in (27.5).

Since we are interested in normalisable solutions we need to set $B = 0$ and search for solutions of the form

$$\varphi(\xi) = h(\xi)e^{-\xi^2/2}, \quad (27.6)$$

where $h(\xi)$ is some unknown function of ξ . Substituting (27.6) into (27.2) gives

$$\frac{d^2h}{d\xi^2} - 2\xi \frac{dh}{d\xi} + (K-1)h = 0. \quad (27.7)$$

We now make a power series ansatz³ for $h(\xi)$:⁴

$$h(\xi) = a_0 + a_1\xi + a_2\xi^2 + \dots = \sum_{j=0}^{\infty} a_j\xi^j, \quad (27.8)$$

so that

$$\frac{dh}{d\xi} = a_1 + 2a_2\xi + 3a_3\xi^2 + \dots = \sum_{j=0}^{\infty} ja_j\xi^{j-1}, \quad (27.9)$$

and

$$\begin{aligned} \frac{d^2h}{d\xi^2} &= 2a_2 + 2 \cdot 3a_3\xi + 3 \cdot 4a_4\xi^2 + \dots \\ &= \sum_{j=0}^{\infty} j(j-1)a_j\xi^{j-2} \\ &= \sum_{j=0}^{\infty} (j+1)(j+2)a_{j+2}\xi^j. \end{aligned} \quad (27.10)$$

Discussion 27.3

Show how we go from the second line to the third line of (27.10).

Now substituting (27.8)-(27.10) into (27.7) gives

$$\sum_{j=0}^{\infty} [(j+1)(j+2)a_{j+2} - (2j+1-K)a_j] \xi^j = 0, \quad (27.11)$$

Discussion 27.4

Confirm (27.11).

Then using the uniqueness of power series expansions⁵, we conclude from (27.11) that for each j ,

$$(j+1)(j+2)a_{j+2} - (2j+1-K)a_j = 0, \quad (27.12)$$

which can be solved for, a_{j+2} , leading to the recursion formula

$$a_{j+2} = \frac{2j+1-K}{(j+1)(j+2)}a_j. \quad (27.13)$$

So given a_0 and a_1 , using (27.13), we can get all the a_j 's iteratively. For example

$$a_2 = \frac{1-K}{2}a_0$$

³ An ansatz is an educated guess that is to be confirmed.

⁴ This is also called the Frobenius method

⁵ Some mathematical details glossed over here – if two infinite power series are convergent and $\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} b_n x^n$ we can conclude that $a_n = b_n$.

$$\begin{aligned}
 a_4 &= \frac{5-K}{12}a_2 = \frac{(5-K)(1-K)}{24}a_0 \\
 a_3 &= \frac{3-K}{6}a_1 \\
 a_5 &= \frac{(7-K)(3-K)}{120}a_1 \\
 &\text{etc ...}
 \end{aligned} \tag{27.14}$$

Discussion 27.5

Show how we get the coefficients shown in (27.14) using (27.13).

Notice that we obtain all the a_j with j even from a_0 and all the a_j with j odd from a_1 , so that it is convenient to write our solution in terms of an even and odd part

$$h(\xi) = h_{\text{even}}(\xi) + h_{\text{odd}}(\xi), \tag{27.15}$$

where

$$\begin{aligned}
 h_{\text{even}} &:= a_0 + a_2\xi^2 + a_4\xi^4 + \dots \\
 h_{\text{odd}} &:= a_1\xi + a_3\xi^3 + a_5\xi^5 + \dots .
 \end{aligned} \tag{27.16}$$

Now not all solutions will be normalisable. To see which solutions are normalisable, we need to look at the behaviour of $h(\xi)$ when $|\xi|$ is large. Notice (from (27.13)) that when $|\xi|$ is big, terms in (27.8) with j large become more important, so we can approximate $h(\xi)$ using

$$a_{j+2} \approx \frac{2}{j}a_j, \tag{27.17}$$

and it follows by induction that, for j large,

$$a_j \approx \frac{C}{(j/2)!}, \tag{27.18}$$

for some constant C .

Discussion 27.6

Explain why when $|\xi| \gg 1$, we can approximate (27.8) using the large j behaviour of a_j .

Discussion 27.7

Show we get (27.17) from (27.13) when j is large.

Discussion 27.8

Use (27.17) and induction to show (27.18) for large j .

Using (27.18), we see that for $|\xi|$ large,

$$h_{\text{even}}(\xi) \approx C \sum_{j \text{ even}} \frac{\xi^j}{(j/2)!}$$

$$\approx C \sum_j \frac{\xi^{2j}}{j!} = Ce^{\xi^2}, \quad (27.19)$$

💬 Discussion 27.9

Show how the last line of (27.19) comes about.

In a similar way we can also show that for large j ,

$$h_{\text{odd}}(\xi) \approx C' \xi e^{\xi^2}. \quad (27.20)$$

💬 Discussion 27.10

Show how we get (27.20).

Now returning to (27.6), using (27.19) and (27.20), we see that for large $|\xi|$, $\varphi(\xi) \approx (C + C'\xi)e^{\xi^2/2}$, which is clearly not normalisable as it goes to ∞ as $\xi \rightarrow \pm\infty$.

It would appear as if our brute force power series method has been a waste of time since it hasn't produced any normalisable solutions but there is a loop hole. If our power series terminates at some n , then we **can** get normalisable solutions. Suppose that for some n , $a_{n+2} = 0$, from (27.13) this would require that

$$a_{n+2} = 0 = \frac{2n+1-K}{(n+1)(n+2)} a_n, \quad (27.21)$$

which in turn means that

$$K = 2n + 1. \quad (27.22)$$

For a given K , (27.22) will only kill either the even or odd series. This means that for normalisable solutions we need two conditions

- either a_0 or a_1 is zero
- $K = 2n + 1$

💬 Discussion 27.11

Explain how we can conclude that $h_n(\xi)$ is a polynomial of order n , with only even or odd powers, for n even or odd respectively.

Recalling the definition of K , we see that

$$K = \frac{2E}{\hbar\omega} = 2n + 1 \implies E = (n + \frac{1}{2})\hbar\omega, \quad (27.23)$$

which is (comfortingly) the same result we got using the algebraic method.

Notice how the quantisation of the energy comes from a rather technical consequence of normalisability – and in fact in all the

examples we considered, the quantisation seems to have arisen for unrelated technical reasons. On the other hand, it is not too hard to understand why we only get solutions for certain energies. Looking at the stationary state ansatz,

$$\psi(x, t) = \varphi(x)e^{-\frac{iEt}{\hbar}} = \varphi(x)e^{-i\bar{\omega}t}, \quad (20.9')$$

we see that the energy of a particular solution is related to the angular frequency⁶ of its phase by $E = \hbar\bar{\omega}$. The **bound** stationary states of the Schrödinger **wave** equation can be thought of as standing waves. Generically, standing wave solutions only occur at certain discrete frequencies for given boundary conditions – just think of the strings of a violin which only vibrate at certain frequencies once we fix the tension and length of the string. As we have seen when we have **unbound** travelling wave solutions, our allowed energies are no longer discrete. However, for the harmonic oscillator, $V \rightarrow \infty$ as $x \rightarrow \pm\infty$ and there are no unbound states.

Discussion 27.12

Explain why the harmonic oscillator does not have unbound states.

Returning to our expression for a_j , if we substitute (27.23) into (27.13) we get

$$a_{j+2} = \frac{-2(n-j)}{(j+1)(j+2)} a_j. \quad (27.24)$$

Using (27.24), lets look at a few simple cases

$$n = 0, a_1 = 0, \implies a_2 = \frac{-2(0-0)}{1 \cdot 2} a_0 = 0, \quad (27.25)$$

so $h_0(\xi) = a_0$ which implies $\varphi_0 = a_0 e^{-\xi^2/2}$ (which, up to a normalisation constant, is the same result we got using the algebraic method).

$$n = 1, a_0 = 0, \implies a_3 = \frac{-2(1-1)}{2 \cdot 3} a_1 = 0, \quad (27.26)$$

so $\varphi_1(\xi) = a_1 \xi e^{-\xi^2/2}$.

$$n = 2, a_1 = 0, \implies a_2 = \frac{-2(2-0)}{(0+1)(0+2)} a_0 = -2a_0, \quad \text{and } a_4 = 0, \quad (27.27)$$

so $\varphi_2(\xi) = a_0(1 - 2\xi^2)e^{-\xi^2/2}$.

In general, it is not hard to see that, $h_n(\xi)$ is a polynomial of order n , with only even or odd powers, for n even or odd respectively. Apart from a normalisation, these are called Hermite polynomials, $H_n(\xi)$. Traditionally, the coefficient in front is chosen so that the highest power has of coefficient of 2^n and the normalised harmonic oscillator solutions are

$$\varphi_n(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}} \frac{1}{\sqrt{2^n n!}} H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right) e^{-m\omega x^2/2\hbar}, \quad (27.28)$$

which are the same as the solutions found using the operator method.

⁶ The $\bar{\omega}$ here should not be confused with the ω that appears in the harmonic oscillator potential. Here $\bar{\omega}$ is the angular frequency of the phase of a particular state while ω is a parameter in the potential corresponding the angular frequency of oscillations we would have if our system was a classical spring. There is of course a relationship between the two we can read off from (26.17), namely $\bar{\omega}_n = (n + \frac{1}{2})\omega$.

$H_0(x) = 1$
$H_1(x) = 2x$
$H_2(x) = 4x^2 - 2$
$H_3(x) = 8x^3 - 12x$
$H_4(x) = 16x^4 - 48x^2 + 12$
$H_5(x) = 32x^5 - 160x^3 + 120x$

Table 27.1: First few Hermite polynomials, $H_n(x)$

❖ Discussion 27.13

Check that eq. (27.28) matches with φ_0 and φ_1 we found in chapter 26.

The first few Hermite polynomials are listed in Table 27.1.

❖ Discussion 27.14

Write a summary of the important points and formulae in this chapter.

28

Unbound states – Free Particle

We looked at states of particles trapped in a potential well. Now we will investigate unbound states. The simplest case to consider is the free particle.

A free particle does not experience any force – this means we can take $V(x) = 0$ and the Hamiltonian simplifies to

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}) = \frac{\hat{p}^2}{2m}. \quad (28.1)$$

In position space, (28.1), gives us

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \varphi_E(x) + \cancel{V(x)} \varphi_E(x) = E \varphi_E(x), \quad (28.2)$$

which we can write as

$$\frac{d^2}{dx^2} \varphi_E(x) = -\frac{2mE}{\hbar^2} \varphi_E(x). \quad (28.3)$$

Notice that this is the same equation as we had inside the square well. It is again useful to define,

$$k^2 = \frac{2mE}{\hbar^2}, \quad (28.4)$$

so that (28.3) becomes

$$\frac{d^2}{dx^2} \varphi_E(x) = -k^2 \varphi_E(x). \quad (28.5)$$

The solution to (28.5) can be written in terms of trigonometric function or complex exponentials – as we will soon see, it turns out to be convenient to use exponentials for the free particle. So, we write the solution to (28.5) as

$$\varphi_E(x) = A e^{ikx} + B e^{-ikx} \quad (28.6)$$

Now since the unbound particle isn't forced to fit in a potential well, there is no limitation on the energy which means the energy is **not** quantised. Mathematically, we have two normalisation constants A, B and the energy E . Our only constraint is normalisation which does not constrain the energy meaning, E is a continuous variable.

This reading is based in part on
[McIntyre et al. 2012, Section 6.1,2] and
[Griffiths 2016, Section 2.4]

Time Evolution

As (28.6) is already in the energy basis, we just need to multiply by the energy dependent phase factor to get time-dependence:

$$\psi_E(x, t) = \varphi_E(x)e^{-iEt/\hbar} = (Ae^{ikx} + Be^{-ikx})e^{-iEt/\hbar} \quad (28.7)$$

Defining the angular frequency,

$$\omega = \frac{E}{\hbar} \stackrel{(28.4)}{=} \frac{\hbar k^2}{2m}, \quad (28.8)$$

equation (28.7) becomes

$$\begin{aligned} \psi_E(x, t) &= (Ae^{ikx} + Be^{-ikx})e^{-i\omega t} \\ &= Ae^{ikx-i\omega t} + Be^{-ikx-i\omega t} \\ &= Ae^{ik(x-(\frac{\omega}{k})t)} + Be^{-ik(x+(\frac{\omega}{k})t)} \\ &= Ae^{ik(x-v_\phi t)} + Be^{-ik(x+v_\phi t)} \end{aligned} \quad (28.9)$$

where we have defined the **phase velocity**¹,

$$v_\phi = \frac{\omega}{k}. \quad (28.10)$$

The terms in (28.9), have the same form as a classical wave, namely a function, $f(x \pm vt)$, with the argument $x \pm vt$. This kind of wave keeps its form as it moves with speed v .

Discussion 28.1

Show how we use (28.4) in the final step of (28.8).

¹ We call v_ϕ the phase velocity because it is associated with the phase.

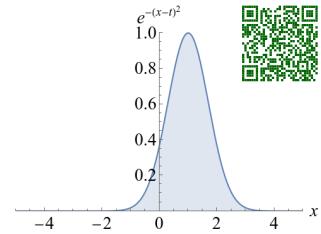


Figure 28.1: Click for animation of $f(x-t) = e^{-(x-t)^2}$.

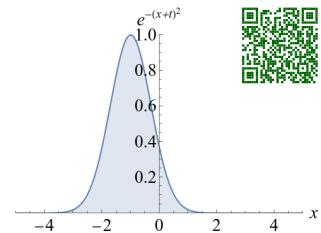


Figure 28.2: Click for animation of $f(x+t) = e^{-(x+t)^2}$.

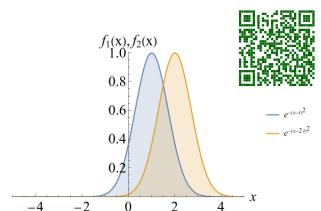


Figure 28.3: Click for animated comparison of $f_1(x-t) = e^{-(x-t)^2}$ and $f_2(x-2t) = e^{-(x-2t)^2}$.

Classical wave example

An example is the function, $f(x-t) = e^{-(x-t)^2}$, as shown in Figure 28.1. Notice that the function is moving in positive x direction. If on the other hand we take, $f(x+t) = e^{-(x+t)^2}$, we see that it moves in negative x direction as you can see if you click on Figure 28.2. Comparing $f_1(x-t) = e^{-(x-t)^2}$ and $f_2(x-2t) = e^{-(x-2t)^2}$, as Figure 28.3, we see f_2 moves twice as fast as f_1 .

Moving parts

The quantum wave function, (28.9) has two parts: $e^{+i(kx-\omega t)}$, which is moving in the positive x direction and, $e^{-i(kx+\omega t)}$, moving in negative x direction. This indicates that, when we see a wave function with spacial dependence, $e^{\pm ikx}$, the sign indicates direction of motion. It is convenient to introduce the general eigenstate notation:

$$|k\rangle \doteq \varphi_k(x) = Ae^{ikx}, \quad (28.11)$$

which uses positive and negative values for k to represent both solutions. We also use the notation,

$$|k(t)\rangle \doteq \psi_k(x, t) = A e^{ikx} e^{-iE_k t/\hbar}, \quad (28.12)$$

where

$$E_k = \frac{\hbar^2 k^2}{2m}. \quad (28.13)$$

Momentum and the de Broglie relation

The function, $\varphi_k(x)$, is also a momentum eigenstate

$$\hat{p}\varphi_k(x) = \frac{\hbar}{i} \frac{d}{dx} A e^{ikx} = (\hbar k) \varphi_k(x), \quad (28.14)$$

and we see that k is related to the momentum eigenvalue by

$$p = \hbar k. \quad (28.15)$$

Discussion 28.2

Explain why

$$k = 2\pi/\lambda \quad (28.16)$$

where λ is the wave-length.

Hint: think about why λ corresponds to the wave length in $e^{i2\pi x/\lambda}$.

Substituting (28.16) into (28.15) we get

$$p = \frac{\hbar}{2\pi} \frac{2\pi}{\lambda} = \frac{\hbar}{\lambda}, \quad (28.17)$$

which you may recognise as the de Broglie relationship between momentum and wavelength. The de Broglie relation, (28.17), is at the heart of wave-particle duality, since it relates the particle property of momentum to the wave property of wavelength

Degeneracy

An interesting new feature of the energy eigenstates, φ_E , is that, for the free particle, a momentum eigenstate has a definite energy but an energy eigenstate may not have a definite momentum.

This is because a general energy eigenstate is a superposition of two momentum eigenstates states, $e^{+i|k|x}$ and, $e^{-i|k|x}$, which have opposite momenta. There are different states corresponding to the same energy.

Terminology

When we have more than one state with the same energy we call such states **degenerate**.

For free particles (in one dimension), energy states are twice degenerate.

Something is fishy about the phase velocity

Now the phase velocity of the free particle energy eigenstate (28.9) is

$$v_\phi = \frac{\omega}{k} \stackrel{(28.8),(28.15)}{=} \frac{E/\hbar}{p/\hbar} = \frac{p^2/2m}{p} = \frac{1}{2} \frac{p}{m}, \quad (28.18)$$

but on the other hand for a classical free particle, $v_{\text{classical}} = p/m$, so

$$v_\phi = \frac{1}{2} v_{\text{classical}}. \quad (28.19)$$

It would appear as if the quantum mechanical wave travels at half the speed of the particle it is meant to represent physically! We'll see this relates to the difference between the phase velocity and what is called the **group velocity**.

While the speed of the wave seems a little strange, a more serious problem appears when we look at the probability density

$$\mathcal{P}(x) = |\varphi_k(x)|^2 = |A|^2 \quad (28.20)$$

which is plotted in Figure 28.4. You can view an animated plot of $\psi_k(x, t) = |\psi_k(x, t)|e^{-iE_k t/\hbar}$ by clicking on Figure 28.5. You can literally see the phase moving with the *phase velocity*.

Looking at Figure 28.4, notice that the probability density is constant. It extends to infinity which means that mathematically, we can not normalise this state since,

$$\int_{-\infty}^{\infty} |\psi_k|^2 dx = |A|^2 \int_{-\infty}^{\infty} dx = \infty. \quad (28.21)$$

Physically, we expect to find a particle within some limited region of space, not spread over an infinite region.

What this means is that the energy eigenstates of the free particle do not represent physical states and consequently a free particle can not have a definite energy. However, despite this, the solutions $\psi_k(x, t)$ are useful – they play an important role in the construction of physical solutions. When we had discrete energy eigenvalues, the general solution to the time-dependent Schrodinger equation can be written as a linear combination of the energy eigenstates:

$$|\psi(0)\rangle = \sum_n c_n |E_n\rangle, \quad \text{where } c_n = \langle E_n | \psi(0) \rangle. \quad (28.22)$$

The difference in this case is that, instead of a sum over a discrete index, we will have to integrate over k which is continuous. By analogy with (28.22), we define

$$\begin{aligned} c(k) &= \langle k | \psi(0) \rangle \\ &= \int_{-\infty}^{\infty} (A e^{ikx})^* \psi(x, 0) dx \end{aligned} \quad (28.23)$$

Now, the question is how exactly to express $\psi(x, 0)$, in terms of the $c(k)$, as an integral over k . The solution to this problem is given by Plancherel's theorem which states

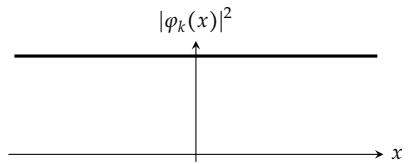


Figure 28.4: Probability density of an energy eigenstate.

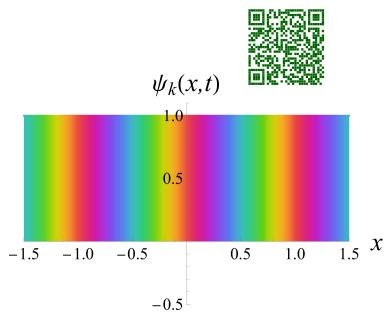


Figure 28.5: Modulus phase plot of $\psi_k(x, t) = |\psi_k(x, t)|e^{-iE_k t/\hbar}$. Click on the image to see an animation.

Plancherel's theorem

$$F(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx \Leftrightarrow f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(k) e^{ikx} dk . \quad (28.24)$$

$F(k)$ is called the Fourier transform of $f(x)$ and $f(x)$ is called the inverse Fourier transform of $F(k)$. For the theorem to hold, the integrals must actually exist². Choosing the convention, $A = 1/\sqrt{2\pi}$, (28.23) becomes

$$\begin{aligned} c(k) &= \langle k | \psi(0) \rangle \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi(x, 0) e^{-ikx} dx \end{aligned} \quad (28.25)$$

and using Plancherel's theorem we obtain

$$\psi(x, 0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} c(k) e^{ikx} dk , \quad (28.26)$$

which expresses the wave function as an integral over the continuous parameter k . We see that the initial wave function, $\psi(x, 0)$, is the inverse Fourier transform of $c(k)$. To find the time evolution of the wave function, we fold in the time-dependant phase of the energy eigenfunction in the usual way

$$\begin{aligned} \psi(x, t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \underbrace{c(k) e^{-iE_k t/\hbar}}_{\psi_k} e^{ikx} dk \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} c(k) e^{i\left(kx - \frac{\hbar k^2}{2m} t\right)} dk , \end{aligned} \quad (28.27)$$

Unlike the energy eigenstates, $|k\rangle$, the wave function (28.27) can be normalised, but it involves a range of k 's which in turn means it is associated with a range of energies and speeds. Such a wavefunction is called a *wave packet*.

Example 28.1: Gaussian wave function

Problem Consider a free particle with the initial normalised Gaussian wave function, $\psi(x, 0) = Ne^{-x^2/2a^2}$, where $N = 1/(\pi^{1/4}\sqrt{a})$. Determine $\psi(x, t)$. You may find the integral, $\int_{-\infty}^{\infty} e^{-a^2 u^2 + bu} du = \frac{\sqrt{\pi}}{a} e^{b^2/4a^2}$, useful.

Solution Using (28.25), we obtain the Fourier transform of the wave-function

$$c(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} N e^{-x^2/2a^2} e^{-ikx} dx \quad (28.28)$$

² In our case, requiring the normalisation of $\psi(x, 0)$ will guarantee that we can apply the theorem. You may be assigned a rough proof of Plancherel's theorem as an exercise.

Comparing with the given integral we identify, $a^2 = 1/2\alpha^2$, $a = 1/\sqrt{2}\alpha$, and, $b = -ik$, so that

$$\begin{aligned} c(k) &= \frac{(\pi\alpha^2)^{-1/4}}{\sqrt{2\pi}} \sqrt{\pi} \sqrt{2}\alpha e^{(-ik)^2 2\alpha^2/4} \\ &= \frac{\sqrt{\alpha}}{\pi^{1/4}} e^{-\alpha^2 k^2/2}. \end{aligned} \quad (28.29)$$

then

$$\begin{aligned} \psi(x, t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} c(k) e^{i(kx - \frac{\hbar k^2}{2m} t)} dk \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\frac{\sqrt{\alpha}}{\pi^{1/4}} e^{-\alpha^2 k^2/2} \right) e^{i(kx - \frac{\hbar k^2}{2m} t)} dk \\ &= \frac{1}{\sqrt{2\pi}} \frac{\sqrt{\alpha}}{\pi^{1/4}} \int_{-\infty}^{\infty} e^{-(\alpha^2 + \frac{i\hbar}{m} t)k^2/2 + (ix)k} dk, \end{aligned} \quad (28.30)$$

Comparing with the given integral, we identify,

$$a^2 = \left(\alpha^2 + \frac{i\hbar}{m} t \right)/2, a = \sqrt{\alpha^2 + \frac{i\hbar}{m} t}/\sqrt{2}, \text{ and, } b = ix, \text{ so that}$$

$$\begin{aligned} \psi(x, t) &= \frac{1}{\sqrt{2\pi}} \frac{\sqrt{\alpha}}{\pi^{1/4}} \frac{\sqrt{\pi}\sqrt{2}}{\sqrt{\alpha^2 + \frac{i\hbar}{m} t}} e^{(ix)^2/(4(\alpha^2 + \frac{i\hbar}{m} t)/2)} \\ &= \frac{N}{\gamma} e^{-x^2/2\gamma^2\alpha^2} \end{aligned} \quad (28.31)$$

where, $\gamma^2 = 1 + \frac{it}{\tau}$, and, $\tau = \frac{m\alpha^2}{\hbar}$. A modulus-phase plot of (28.31) is shown in Figure 28.6. If you click on the figure, you will notice that the wave function spreads out over time. The spread is related to the fact that the different components used to construct the state move at different phase velocities. A plot of $\pm|\psi|$ and the real part of (28.31) is shown in Figure 28.7.

We are now ready to return to the mismatch between the speed of the energy eigenstates, ψ_k , and the particle it represents. Firstly, since ψ_k does not represent a physical state on its own, the fact that it moves at a different speed, is not in of itself a problem. The interesting question becomes:

“How is the information about a particle’s velocity contained in the full wavefunction (28.27)?”

To answer the question, we need to come to grips with what a wave-packet really is. A wave packet is a superposition of sinusoidal waves or ripples (the ψ_k) whose amplitude is modulated by an “envelope” (the $c(k)$). For an illustration of this see Figure 28.8. The particle’s velocity is not equal to the velocity of individual ripples (which we call the phase velocity v_ϕ) but the velocity of the

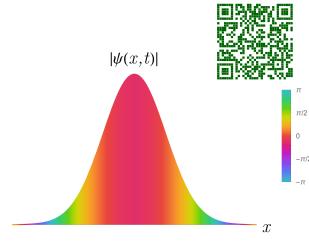


Figure 28.6: Modulus-phase plot of (28.31) at $t = \tau$. Click on the image for an animated version

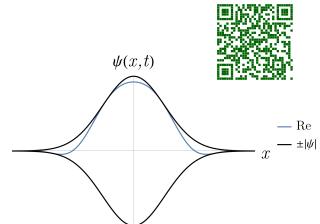


Figure 28.7: $\pm|\psi|$ and real part of (28.31) at $t = \tau$. Click on the image for an animated version

envelope (which we call the group velocity v_g).

The relationship between the group velocity and phase velocity depends on the type of wave. For example for the waves on a string, $v_g = v_\phi$, but for water waves³, $v_g = \frac{1}{2}v_\phi$. For different examples with varying relationships between v_g and v_ϕ , see some nice [animations](#) from the Institute of Sound and Vibration Research⁴.

Group and phase velocity for general waves

How can we determine the relationship between the phase and group velocity for a given wave? A general wave packet will be of the form

$$\psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} c(k) e^{i(kx - \omega t)} dk , \quad (28.32)$$

and in our case $\omega = \hbar k^2 / 2m$. In general, the relationship between ω and k is called the dispersion relation and, as we will see, it determines the relationship between the phase and group velocities. In fact, (28.32) can be applied to other kinds of waves as well. In general the relationship between ω and k is called the dispersion relation and, as we will see, it determines the relationship between the phase and group velocities.

In our investigation of how the phase and group velocities are related, we will assume that $c(k)$ is narrowly peaked about some k_0 . There is nothing wrong with considering a broad envelope but such a wave-packet will change shape quickly and the whole notion of a group velocity will not be well defined. Now, if $c(k)$ is negligible, except for some region close to k_0 , we can approximate $\omega(k)$ using a Taylor series,

$$\omega(k) \approx \omega(k_0) + \omega'_0(k - k_0) \quad (28.33)$$

where $\omega_0 = \omega(k_0)$ and $\omega'_0 = \omega'(k_0)$, without significantly effecting (28.32). It is also convenient to shift k by defining $s = k - k_0$ so that from (28.32) we obtain

$$\begin{aligned} \psi(x, t) &\approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} c(k_0 + s) e^{i[(k_0 + s)x - (\omega_0 + \omega'_0 s)t]} ds \\ &= \frac{1}{\sqrt{2\pi}} e^{i(k_0 x - \omega_0 t)} \int_{-\infty}^{\infty} c(k_0 + s) e^{is(x - \omega'_0 t)} ds \end{aligned} \quad (28.34)$$

The term in front of the integral corresponds to a sinusoidal wave travelling at speed ω_0/k_0 – these are the ripples. The ripples are multiplied⁵ by the integral (in other words the envelop) which is a function of $(x - \omega'_0 t)$. This tells us that the envelope moves at a speed ω'_0 . In summary

$$\begin{aligned} v_\phi &= \frac{\omega}{k} , \\ v_{\text{group}} &= \frac{d\omega}{dk} , \end{aligned} \quad (28.35)$$

where it is understood that (28.35) is evaluated at $k = k_0$.

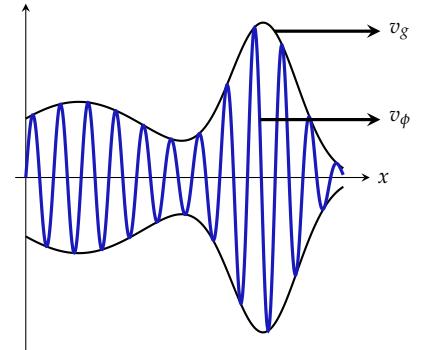


Figure 28.8: A wave packet. The overall shape of the wave packet is called the envelope which may move at a different speed to the individual ripples.

³ Next time you are looking at ripples on the surface of water, try see if you can observe this.



Applying (28.35) to our quantum mechanical free particle, we have seen that phase velocity is half the classical velocity of the particle:

$$v_\phi = \frac{\omega}{k} = \frac{1}{2} v_{\text{classical}} . \quad (28.36)$$

On the other hand, the group velocity matches with the classical velocity

$$\begin{aligned} v_{\text{group}} &= \frac{d\omega}{dk} \Big|_{k=k_0} \stackrel{(28.8),(28.15)}{=} \frac{d(\hbar\omega)}{d(\hbar k)} \Big|_{k=k_0} \\ &= \frac{dE}{dp} \Big|_{p=p_0} = \frac{d(p^2/2m)}{dp} \Big|_{p=p_0} \\ &= \frac{p_0}{m} = v_{\text{classical}} . \end{aligned} \quad (28.37)$$

We see that its the wave packet that moves with a group velocity corresponding to the speed we expect for a classical free particle.

Going back to example 28.1, as shown in Figure 28.9 the gaussian $c(k)$ is peaked at $k = 0$ so the group velocity is

$$v_{\text{group}} = \frac{\partial\omega}{\partial k} \Big|_{k=0} = \frac{\hbar k}{m} \Big|_{k=0} = 0 , \quad (28.38)$$

and the wave packet is not moving – recall that for the animation in Figure 28.6 the wave packet really just spreads out but does not move. Now suppose at $t = 0$, we have an energy eigenstate φ_{k_0} modulated by a Gaussian profile,

$$\psi(x, 0) = Ne^{-x^2/2\alpha^2} e^{ik_0 x} , \quad (28.39)$$

shown in Figure 28.11.

Discussion 28.3

Generalising the calculation in example 28.1 show that the Fourier transform of (28.39) is

$$c(k) = \frac{\sqrt{\alpha}}{\pi^{1/4}} e^{-\alpha^2(k-k_0)^2/2} . \quad (28.40)$$

As shown in Figure 28.10, we see that this wave packet, (28.40) is peaked at k_0 , so the group velocity is

$$v_{\text{group}} = \frac{\partial\omega}{\partial k} \Big|_{k=k_0} = \frac{\hbar k}{m} \Big|_{k=k_0} = \frac{\hbar k_0}{m} \stackrel{(28.15)}{=} \frac{p_0}{m} , \quad (28.41)$$

which is what we expect for a classical particle. Performing the inverse Fourier transform of (28.40) to obtain $\psi(x, t)$ is a little technical – you may be assigned this as an exercise – but you can get some insight by clicking on Figure 28.11 for an animated modulus-phase plot. Looking at the animation, you can see the Gaussian envelope, moving at the group velocity, is going twice as fast as the phase “rainbow” which is moving at the phase velocity. Clicking on Figure 28.12 you can clearly see the how the overall shape (ie. the envelope), travelling at the group velocity, travels faster than the wiggles which are travelling at the phase velocity.

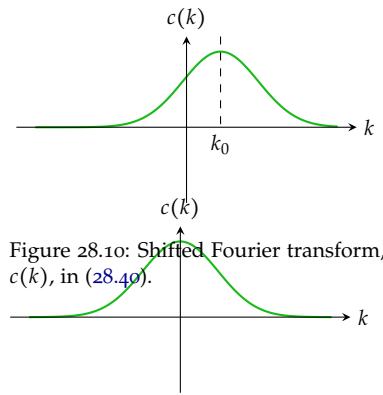


Figure 28.10: Shifted Fourier transform, $c(k)$, in (28.40).

Figure 28.9: Fourier transform, $c(k)$, of the wave function at $t = 0$ for example 28.1.

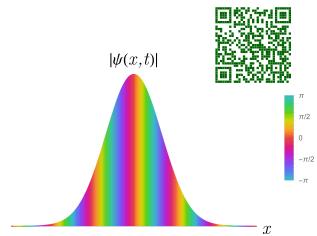


Figure 28.11: Modulus-phase plot $\psi(x, t)$ with initial state (28.39). Click for an animation. We have taken τ large so that the wave packet spreads out slowly.

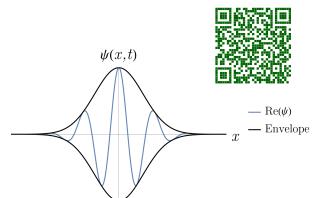


Figure 28.12: Envelope and real part of $\psi(x, t)$ with initial state (28.39). Click for an animation. Once again τ is large here.

❖ Discussion 28.4

Write a summary of the important points and formulae in this chapter.

Continuous bases

In the last chapter we encountered the continuous basis, $|k\rangle$, with our states expressed as an integral, rather than a sum, over this basis. Thus far, we have expected basis states to have two basic properties:

Basis state properties

1. orthonormal : $\langle a_i | a_j \rangle = \delta_{ij}$
2. complete: $\sum_i |a_i\rangle \langle a_i| = 1$

This reading is based on [McIntyre et al. 2012, Sections 6.1,6.2] and [Griffiths 2016, Sections 1.5, 2.4, 3.3]

To adapt orthonormality to a continuous basis, we need the **Dirac delta function**: $\delta(x - x_0)$.

Mathematical interlude - the Dirac delta function

The Dirac delta function

The dirac delta “function” (mathematicians refer to this as a distribution) is zero everywhere, except at $x = 0$, where its infinite.

$$\delta(x) = \begin{cases} \infty & x = 0, \\ 0 & \text{otherwise.} \end{cases} \quad (29.1)$$

It has the additional properties that, its integral is finite,

$$\int_{-\infty}^{\infty} \delta(x) dx = 1. \quad (29.2)$$

A schematic illustration of the delta function is shown in Figure 29.1. Roughly speaking we can think of the delta function as the limit of a set of functions, $\delta_n(x)$, whose area remains constant while its width goes to zero and its height goes to infinity as $n \rightarrow \infty$. We then have

$\lim_{n \rightarrow \infty} \int \delta_n(x) f(x) dx = \int \delta(x) f(x) dx$. A Gaussian example of such a set is shown in Figure 29.2.

Now, if $f(x)$ is a reasonable¹ function (and for example not another delta function), then from (29.1) we see that

¹ We'll leave it to mathematicians to be more precise about what “reasonable” means.

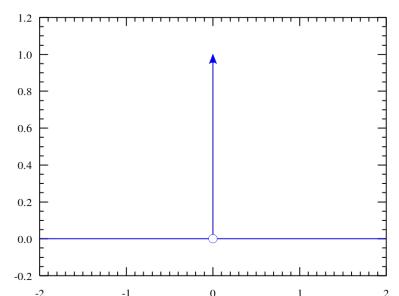


Figure 29.1: Representation of $\delta(x)$ —the height represents the area once it is integrated.

Attribution: Original SVG by Omegatron. Original PNG version by PAR. This version adapted by Qef, CC BY-SA 3.0, via Wikimedia Commons

$f(x)\delta(x)$ is zero everywhere except at $x = 0$ which tells us that

$$f(x)\delta(x) = f(0)\delta(x). \quad (29.3)$$

Discussion 29.1

Justify (29.3).

Using (29.2) and (29.3), we see that the Dirac delta function picks out the value of a function in an integral:

$$\int_{-\infty}^{\infty} f(x)\delta(x) dx \stackrel{(29.3)}{=} f(0) \int_{-\infty}^{\infty} \delta(x) dx \stackrel{(29.2)}{=} f(0). \quad (29.4)$$

We can of course shift the delta function

$$\delta(x - x_0) = \begin{cases} \infty & x = x_0, \\ 0 & \text{otherwise.} \end{cases} \quad (29.5)$$

so that (29.3) generalises to

$$f(x)\delta(x - x_0) = f(x_0)\delta(x - x_0), \quad (29.6)$$

and (29.4) becomes

$$\int_{-\infty}^{\infty} f(x)\delta(x - x_0) dx = f(x_0), \quad (29.7)$$

which picks out the value of the function at x_0 . The Dirac delta function can be thought of as a continuous analogue of the Kronecker delta that picks out the value in a sum:

$$\sum_{n=1}^{\infty} a_n \delta_{mn} = a_m. \quad (29.8)$$

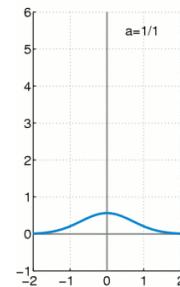


Figure 29.2: The δ -function as the limit, as $a \rightarrow 0$, of a sequence of normalised Gaussians $\delta_a(x) = \frac{e^{-x^2/a^2}}{a\sqrt{\pi}}$ with decreasing width. Click on the image for an animation of $a \rightarrow 0$.

Attribution: by Oleg Alexandrov, source, Public domain, via Wikimedia Commons

Example 29.1: Using $\delta(x)$

Example Evaluate

$$(a) I_1 = \int_0^{10} x^2 \delta(x - 5) dx$$

$$(b) I_2 = \int_0^1 \sinh^3(x^2) \delta(x - 5) dx$$

Solution

(a) $\delta(x - 5)$ picks out the value of the integrand at, $x = 5$, so $I_1 = 25$.

(b) $x = 5$ is not in the domain of integration so $I_2 = 0$.

The dirac delta function should really be thought of as something we intend to use *inside* an integral. We say that two expressions

involving delta functions, say $D_1(x)$ and $D_2(x)$, are equal if

$$\int_{-\infty}^{\infty} D_1(x)f(x) dx = \int_{-\infty}^{\infty} D_2(x)f(x) dx . \quad (29.9)$$

for all reasonable¹ functions $f(x)$.

Example 29.2: Some dirac delta-ology

Problem Show

$$\delta(kx) = \frac{1}{|k|}\delta(x). \quad (29.10)$$

Solution Consider an arbitrary test function $f(x)$, and the integral

$$\int_{-\infty}^{\infty} f(x)\delta(kx) dx . \quad (29.11)$$

Let $u = kx$, so that $x = u/k$ and $dx = \frac{1}{k}du$. Now if k is positive, the limits of the integral will remain the same but if k is negative, $x = \infty$ means that $u = -\infty$ (and vice versa) reversing the order of the limits, so

$$\int_{-\infty}^{\infty} f(x)\delta(kx) dx = \pm \int_{-\infty}^{\infty} f(u/k)\delta(u) \frac{du}{k} = \pm \frac{f(0)}{k} = \frac{f(0)}{|k|} . \quad (29.12)$$

Discussion 29.2

Looking at the first equality in (29.12), explain when we have a plus sign and when we have a minus sign. Then explain the final equality in (29.12).

From (29.12) we see that $\delta(kx)$ behaves in the same way as $\delta(x)/|k|$ under the integral sign, or in other words

$$\int_{-\infty}^{\infty} f(x)\delta(kx) dx = \frac{1}{|k|} \int_{-\infty}^{\infty} f(x)\delta(x) dx . \quad (29.13)$$

Finally, using (29.13) and the criterion (29.9) allows us to conclude that (29.10) holds.

Fourier transform of $\delta(x)$

The Fourier transform of $\delta(x)$ is

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \delta(x)e^{-ikx} dx = \frac{1}{\sqrt{2\pi}} e^0 = \frac{1}{\sqrt{2\pi}} . \quad (29.14)$$

Discussion 29.3

Using Plancherel's theorem (28.24) explain how we can conclude from (29.14) that

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dk . \quad (29.15)$$

 Discussion 29.4

Explain how we can conclude from (29.15) that

$$\delta(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dx . \quad (29.16)$$

Energy basis

We are now in a position to study how orthonormality and completeness generalise to the free particle energy basis, $|k\rangle$, from the last chapter.

Dirac Orthonormality

Looking at the inner product between two energy eigenstates

$$\begin{aligned} \langle k' | k \rangle &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-ik'x} \frac{1}{\sqrt{2\pi}} e^{ikx} dx \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i(k-k')x} dx \\ &\stackrel{(29.16)}{=} \delta(k - k') \\ &= \delta(k' - k) \end{aligned} \quad (29.17)$$

 Discussion 29.5

Explain the steps in (29.17).

We refer to states that satisfy (29.17) as **Dirac orthonormal**:

$$\langle k' | k \rangle = \delta(k' - k) \quad (29.18)$$

Completeness

To study the completeness of $|k\rangle$ basis, lets re-express

$$\psi(x, 0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} c(k) dk , \quad (28.26')$$

in Dirac notation. Firstly, $\psi(x, 0) = \langle x | \psi(0) \rangle$, since this is the position space representation of $|\psi(0)\rangle$. Then, $\frac{e^{ikx}}{\sqrt{2\pi}} = \langle x | k \rangle$, since this is the position space representation of $|k\rangle$. Finally using (28.25), $c(k) = \langle k | \psi(0) \rangle$, we can re-write (28.26) as

$$\begin{aligned} \langle x | \psi(0) \rangle &= \int_{-\infty}^{\infty} \langle x | k \rangle \langle k | \psi(0) \rangle dk \\ &= \langle x | \left(\int_{-\infty}^{\infty} |k\rangle \langle k| dk \right) |\psi(0)\rangle . \end{aligned} \quad (29.19)$$

Discussion 29.6

Explain how using (29.19) and the fact that $|\psi(0)\rangle$ is an arbitrary state, we can conclude

$$\int_{-\infty}^{\infty} |k\rangle\langle k| dk = 1 \quad (29.20)$$

In general, we say a continuous basis with states, $|z\rangle$, labelled by a continuous parameter z , is Dirac orthonormal and complete if it satisfies

Complete Dirac orthonormal basis

1. Dirac orthonormal : $\langle z'|z\rangle = \delta(z' - z)$
2. Complete: $\int |z\rangle\langle z| dz = 1$

Position Basis

Up to now, we have treated the position eigenket as an abstract ket $|x\rangle$. For $|x_0\rangle$ to be a position eigenket with eigenvalue x_0 , it must satisfy

$$\hat{x}|x_0\rangle = x_0|x_0\rangle. \quad (29.21)$$

Let, $\varphi_{x_0}(x)$, be the position space representation of $|x_0\rangle$. In other words, $\varphi_{x_0}(x) \doteq |x_0\rangle$. This means that in position space, $\varphi_{x_0}(x)$ must satisfy

$$x\varphi_{x_0}(x) = x_0\varphi_{x_0}(x). \quad (29.22)$$

The Fourier transform of, $\varphi_{x_0}(x) \doteq |x_0\rangle$, is

$$\begin{aligned} c(k) &= \langle k|x_0\rangle \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \varphi_{x_0}(x) e^{-ikx} dx. \end{aligned} \quad (29.23)$$

On the other hand we know that $\langle x_0|k\rangle$ is the positions space representation of $|k\rangle$ at x_0 , so $\langle x_0|k\rangle = \frac{1}{\sqrt{2\pi}} e^{ikx_0}$. This tells us that

$$c(k) = \langle k|x_0\rangle = (\langle x_0|k\rangle)^* = \frac{1}{\sqrt{2\pi}} e^{-ikx_0}. \quad (29.24)$$

Discussion 29.7

Explain how we can use (29.24) and Plancherel's theorem to conclude

$$\varphi_{x_0}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{2\pi}} e^{-ikx_0} \right) e^{ikx} dk \quad (29.25)$$

💬 Discussion 29.8

Explain how we can conclude from (29.25) that

$$\varphi_{x_0}(x) = \delta(x - x_0) . \quad (29.26)$$

💬 Discussion 29.9

Explain why, $\varphi_{x_0}(x) = \delta(x - x_0)$, satisfies (29.22).

💬 Discussion 29.10

Explain why the $\varphi_{x_0}(x)$ are not normalisable.

It is not hard to see that the $|x\rangle$ -basis is Dirac orthonormal

$$\begin{aligned} \langle x_1 | x_2 \rangle &= \int_{-\infty}^{\infty} (\varphi_{x_1}(x))^* \varphi_{x_2}(x) dx \\ &= \int_{-\infty}^{\infty} \delta(x - x_1) \delta(x - x_2) dx \\ &= \delta(x_1 - x_2) . \end{aligned} \quad (29.27)$$

💬 Discussion 29.11

Explain the steps in (29.27).

Completeness is also easily show. Consider an arbitrary state $|\psi\rangle$, then

$$\psi(x_0) = \langle x_0 | \psi \rangle \quad (29.28)$$

on the other hand

$$\begin{aligned} \langle x_0 | \left(\int_{-\infty}^{\infty} |x\rangle \langle x| dx \right) | \psi \rangle &= \int_{-\infty}^{\infty} \langle x_0 | x \rangle \langle x | \psi \rangle dx \\ &= \int_{-\infty}^{\infty} \delta(x - x_0) \psi(x) dx \\ &= \psi(x_0) \end{aligned} \quad (29.29)$$

💬 Discussion 29.12

Explain how comparing (29.28) and (29.29) we can conclude that

$$\int_{-\infty}^{\infty} |x\rangle \langle x| dx = 1 \quad (29.30)$$

Momentum basis

We saw before that the free particle energy eigenstates, $\varphi_k(x)$, are also a momentum eigenstate with momentum eigenvalue hk . This

is a rather special feature of the free particle but in general we can defined momentum eigenstates

$$\hat{p}\varphi_p(x) = \frac{\hbar}{i} \frac{d}{dx} A e^{ipx/\hbar} = p\varphi_p(x), \quad (29.31)$$

with the corresponding ket satisfying

$$\hat{p}|p\rangle = p|p\rangle. \quad (29.32)$$

Once again these are not normalisable states in the usual sense but we can ensure that they are Dirac normalised

$$\begin{aligned} \langle p'|p\rangle &= \int_{-\infty}^{\infty} (A e^{ip'x/\hbar})^* (A e^{ipx/\hbar}) dx \\ &= |A|^2 \int_{-\infty}^{\infty} e^{i\left(\frac{p-p'}{\hbar}\right)x} dx \\ &= |A|^2 2\pi \delta\left(\frac{p-p'}{\hbar}\right) \\ &= |A|^2 2\pi\hbar\delta(p-p') \end{aligned} \quad (29.33)$$

💬 Discussion 29.13

Explain the steps in (29.33).

Which means in if we take, $A = 1/\sqrt{2\pi\hbar}$, our Dirac normalised momentum eigenstates are

$$\varphi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} \quad (29.34)$$

💬 Discussion 29.14

Explain how we can conclude

$$|k\rangle = \sqrt{\hbar}|p\rangle \quad \text{with} \quad p = \hbar k. \quad (29.35)$$

Using (29.35), we see that the $|p\rangle$ are complete:

$$1 = \int_{-\infty}^{\infty} |k\rangle\langle k| dk = \int_{-\infty}^{\infty} \sqrt{\hbar}|p\rangle\langle p|\frac{1}{\hbar} dp = \int_{-\infty}^{\infty} |p\rangle\langle p| dp. \quad (29.36)$$

We would now like to know how we can express states in terms of the momentum basis. Remember that we can use the completeness relationship to change from one basis to another. So starting with the position representation of a general state

$$\begin{aligned} \psi(x) &= \langle x|\psi\rangle = \langle x|1|\psi\rangle \\ &= \int_{-\infty}^{\infty} \langle x|p\rangle\langle p|\psi\rangle dp \end{aligned} \quad (29.37)$$

The 1st term in (29.37), namely $\langle x|p\rangle$, is the position space representation of a momentum eigenstate onto the position basis given by

(29.34). The 2nd term, $\langle p|\psi \rangle$, is the projection of the general state, $|\psi\rangle$ onto the momentum basis. We introduce new notation, writing

$$\phi(p) = \langle p|\psi \rangle , \quad (29.38)$$

and we call, $\phi(p)$, the **momentum space wave function**. According to QM Postulate 4, if $|p\rangle$ was a normalisable eigenstate, $|\langle p|\psi \rangle|^2 = |\phi(p)|^2$, would give the probability, \mathcal{P}_p , of measuring momentum p . In the next chapter we will generalise QM Postulate 4 to a continuous basis. We will see that the square of the momentum space wave function, $|\phi(p)| = |\langle p|\psi \rangle|^2$, is not the probability but the probability **density** for momentum, just like the wave function, $\psi(x) = \langle x|\psi \rangle$, is the probability density for position.

The wave function $\psi(x)$ and the momentum space wave function $\phi(p)$ are both representations of the state $|\psi\rangle$ in different bases.

Putting (29.34), the form of $\varphi_p(x) = \langle x|p\rangle$, and the definition of $\phi(p)$, (29.38), into (29.37), leads to

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \phi(p) e^{ipx/\hbar} dp , \quad (29.39)$$

Discussion 29.15

Explain how we can conclude that, for a free particle,

$$\psi(x, t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \phi(p) e^{ip(x-(p/2m)t)/\hbar} dp . \quad (29.40)$$

² up to some factors of \hbar

Comparing (29.39) with (28.24), we see that $\psi(x)$ is essentially² the inverse Fourier transform of $\phi(p)$ and the transform is basically just a change of basis. On the other hand, writing out (29.38) in position space, we get

$$\begin{aligned} \phi(p) &= \langle p|\psi \rangle \\ &= \int_{-\infty}^{\infty} (\varphi_p(x))^* \psi(x) dx \\ &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \psi(x) e^{-ipx/\hbar} dx \end{aligned} \quad (29.41)$$

and we see that $\phi(p)$ is essentially the Fourier transform of $\psi(x)$.

Discussion 29.16

Explain how we can conclude that, for a free particle,

$$\phi(p, t) = \langle p|\psi(t) \rangle = \phi(p) e^{-ip^2 t / 2m\hbar} . \quad (29.42)$$

Discussion 29.17

Generalising the calculation in example 28.1, show that with

the state, $\psi(x, 0) = Ne^{-x^2/2\alpha^2}$,

$$\phi(p) = \frac{\sqrt{\alpha/\hbar}}{\pi^{1/4}} e^{-\alpha^2 p^2/2\hbar^2} \quad (29.43)$$

❖ Discussion 29.18

Using the expression in (29.43), show that $|\phi(p)|^2$ has dimensions of $(\text{momentum})^{-1}$. Explain why this is consistent with what we expect for a momentum probability density.

❖ Discussion 29.19

Write a summary of the important points and formulae in this chapter.

30

Momentum space and uncertainty

$|\phi(p)|^2$ is a probability density

This reading is based on [McIntyre et al. 2012, Section 6.2]

As mentioned in the last chapter, according to QM Postulate 4, if $|p\rangle$ was a normalisable eigenstate, $|\langle p|\psi \rangle|^2 = |\phi(p)|^2$, would be the probability of measuring momentum p . We will now show that $\mathcal{P}_p = |\phi(p)|^2$ is actually the momentum probability *density*. For \mathcal{P}_p to be a probability density,

- it must be normalised to 1 (when integrated¹ over p)
- we should find that² $\langle p \rangle = \int_{-\infty}^{\infty} p \mathcal{P}_p \, dp$.

¹ Just like $|\psi(x)|^2$ is normalised to 1 when integrated over x

² Just like $\langle x \rangle = \int_{-\infty}^{\infty} x |\psi(x)|^2 \, dx$

Let's see how that works. Starting with the normalisation of the original wave function:

$$\begin{aligned} 1 &= \langle \psi | \psi \rangle \\ &= \int_{-\infty}^{\infty} \langle \psi | p \rangle \langle p | \psi \rangle \, dp \\ &= \int_{-\infty}^{\infty} |\phi(p)|^2 \, dp , \end{aligned} \quad (30.1)$$

consistent with a probability density³.

Discussion 30.1

Explain the steps in (30.1).

³ Equation (30.1) is essentially due to the fact that $\phi(p)$ and $\psi(x)$ are just different representations of the same state in different bases. No matter how we represent a state, it should still be normalised.

Furthermore

$$\begin{aligned} \langle p \rangle &= \langle \psi | \hat{p} | \psi \rangle \\ &= \int_{-\infty}^{\infty} \langle \psi | \hat{p}(|p\rangle\langle p|) | \psi \rangle \, dp \\ &= \int_{-\infty}^{\infty} \langle \psi | p | p \rangle \langle p | \psi \rangle \, dp \\ &= \int_{-\infty}^{\infty} p \langle \psi | p \rangle \langle p | \psi \rangle \, dp \\ &= \int_{-\infty}^{\infty} p |\phi(p)|^2 \, dp . \end{aligned} \quad (30.2)$$

Discussion 30.2

Explain the steps in (30.2) .

❖ Discussion 30.3

Explain how we can infer from (30.1) and (30.2) that $|\phi(p)|^2$ is the probability density for momentum. This means that the chance of finding a state to have momentum $p_1 < p < p_2$ is

$$\mathcal{P}_{p_1 < p < p_2} = \int_{p_1}^{p_2} |\phi(p)|^2 dp . \quad (30.3)$$

❖ Discussion 30.4

Explain how we can conclude from (30.2) that with respect to the momentum basis, $\hat{p} = p$.

📘 Example 30.1:

Problem Given the momentum space wave function,

$\phi(p) = \frac{\sqrt{\alpha/\hbar}}{\pi^{1/4}} e^{-\alpha^2 p^2 / 2\hbar^2}$, what is the probability that the momentum lies between $\frac{\hbar}{\alpha}$ and $\frac{2\hbar}{\alpha}$?

Solution Using (30.3),

$$\begin{aligned} \mathcal{P}_{\frac{\hbar}{\alpha} < p < \frac{2\hbar}{\alpha}} &= \int_{\frac{\hbar}{\alpha}}^{\frac{2\hbar}{\alpha}} |\phi(p)|^2 dp \\ &= \frac{\alpha}{\hbar\sqrt{\pi}} \int_{\frac{\hbar}{\alpha}}^{\frac{2\hbar}{\alpha}} e^{-\alpha^2 p^2 / \hbar^2} dp . \end{aligned} \quad (30.4)$$

Letting $u = \alpha p / \hbar$, we find

$$\mathcal{P}_{\frac{\hbar}{\alpha} < p < \frac{2\hbar}{\alpha}} = \frac{1}{\sqrt{\pi}} \int_1^2 e^{-u^2} du . \quad (30.5)$$

Now the integral in (30.5) can be evaluated numerically. For example the mathematica code,

N[Integrate[Exp[-u^2], {u, 1, 2}]/Sqrt[Pi]], gives us the value $\mathcal{P}_{\frac{\hbar}{\alpha} < p < \frac{2\hbar}{\alpha}} \approx 0.076$. The function **N[]** is used to give a numerical value.

❖ Discussion 30.5

Consider

$$\begin{aligned} \langle x \rangle &= \langle \psi | \hat{x} | \psi \rangle \\ &= \int_{-\infty}^{\infty} \langle \psi | (|p\rangle \langle p|) \hat{x} | \psi \rangle dp \\ &= \int_{-\infty}^{\infty} \phi^*(p) \langle p | \hat{x} | \psi \rangle dp \end{aligned} \quad (30.6)$$

Explain the steps in (30.6). Then

$$\begin{aligned}\langle p|\hat{x}|\psi\rangle &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\hbar}} e^{-ipx/\hbar} x \psi(x) dx \\ &= i\hbar \frac{d}{dp} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\hbar}} e^{-ipx/\hbar} \psi(x) dx \\ &= i\hbar \frac{d}{dp} \phi(p).\end{aligned}\quad (30.7)$$

Putting (30.7) into (30.6) we get

$$\langle \hat{x} \rangle = \int_{-\infty}^{\infty} \phi^*(p) \left(i\hbar \frac{d}{dp} \right) \phi(p) dp. \quad (30.8)$$

Explain the steps in (30.7). Explain how we can conclude from (30.8) that in momentum space $\hat{x} \doteq i\hbar \frac{d}{dp}$.

In summary we have the operator representations

Position space	Momentum space
$\hat{x} \doteq x$	$\hat{x} \doteq i\hbar \frac{d}{dp}$
$\hat{p} \doteq -i\hbar \frac{d}{dx}$	$\hat{p} \doteq p$

Generalised 4th QM postulate

We can now generalise the 4th to Dirac normalised eigenstates

4th QM postulate

- (a) For **normalised eigenstates**, the probability of obtaining the eigenvalue a_n in a measurement of the observable A on the system in the state $|\psi\rangle$ is $\mathcal{P}_{a_n} = |\langle a_n | \psi \rangle|^2$ where $|a_n\rangle$ is the normalized eigenvector of \hat{A} corresponding to the eigenvalue a_n .
- (b) For **Dirac normalised eigenkets**, the probability of an eigenvalue between a and $a + da$, in a measurement of the observable A , on the system in the state $|\psi\rangle$, is $\mathcal{P}_a = |\langle a | \psi \rangle|^2 da$ where $|a\rangle$ is the Dirac normalized eigenket of \hat{A} corresponding to the eigenvalue a .

Uncertainty principle

The Fourier relation between momentum space and position space is related to Heisenberg's uncertainty principle. Recall from section 10.3 that two observables can not be measured at the same time if

they do not commute:

$$(\Delta A)^2 (\Delta B)^2 \geq \left(\frac{1}{2i} \langle [\hat{A}, \hat{B}] \rangle \right)^2 \quad (30.9)$$

where the uncertainty corresponds to the standard deviation

$$\Delta A = \sqrt{\langle (\hat{A} - \langle \hat{A} \rangle)^2 \rangle} = \sqrt{\langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2} \quad (30.10)$$

In particular lets look at position and momentum. Recall that

$$[\hat{x}, \hat{p}] = i\hbar. \quad (25.16')$$

Discussion 30.6

Show that we also get (25.16) if we use the momentum space representation of the operators.

Since they do not commute, position and momentum are not compatible observables, (25.16) and the uncertainty principle tells us that,

$$\Delta x \Delta p \geq \frac{\hbar}{2} \quad (30.11)$$

Discussion 30.7

Show how we get (30.11) from (30.9).

So, momentum and position are not independent: the more we know about one, the less we know about the other. This tells us that we can not precisely predict the trajectory of a particle in quantum mechanics but we **can** make predictions about finding the particle in some region of space or the probability that its momentum is within some range of values. It turns out that the Fourier relations between momentum and position, (29.39) and (29.41), helps us understand this.

We will look at a Gaussian wave functions as it is relatively easy to integrate and is a good model of many experimental situations. Recall that in chapter 19 we found for the wave function,

$$\psi(x) = Ne^{-x^2/2\alpha^2} \quad \text{where } N = (1/\alpha^2\pi)^{1/4}, \quad (30.12)$$

that,

$$\begin{aligned} \langle \hat{x} \rangle &= N^2 \int_{-\infty}^{\infty} x e^{-x^2/\alpha^2} dx = 0, \\ \langle \hat{x}^2 \rangle &= N^2 \int_{-\infty}^{\infty} x^2 e^{-x^2/\alpha^2} dx = \alpha^2/2, \end{aligned} \quad (30.13)$$

so that $\Delta x = \sqrt{\langle \hat{x}^2 \rangle - \langle \hat{x} \rangle^2} = \alpha/\sqrt{2}$.

We also saw in chapter 29 that the momentum space wave function, given the wavefunction (30.12), is

$$\phi(p) = \frac{\sqrt{\alpha/\hbar}}{\pi^{1/4}} e^{-\alpha^2 p^2/2\hbar^2} \quad (30.14)$$

Discussion 30.8

Explain how we can use (30.14) and integrals analogous to (30.13), to conclude that, for this state,

$$\langle \hat{p} \rangle = 0$$

$$\langle \hat{p}^2 \rangle = \hbar^2/2\alpha^2 \quad (30.15)$$

From (30.15), we find $\Delta p = \sqrt{\langle \hat{p}^2 \rangle - \langle \hat{p} \rangle^2} = \hbar/\sqrt{2}\alpha$. We see that the free particle Gaussian wave function has the *minimum* position-momentum uncertainty allowed:

$$\Delta x \Delta p = \frac{\alpha}{\sqrt{2}} \frac{\hbar}{\sqrt{2}\alpha} = \frac{\hbar}{2}. \quad (30.16)$$

We can get an feeling for (30.16) from Figure 30.1. In Figure 30.1 (a) we see a wave function that is highly peaked Gaussian in position space meaning Δx is small. On the other hand when we Fourier transform it to momentum space, we get a broad Gaussian meaning Δp is large. The smaller we try make Δx , the larger Δp will be. Figure 30.1 (b) shows the opposite situation – a narrow Gaussian in momentum space corresponds to a broad Gaussian in position space.

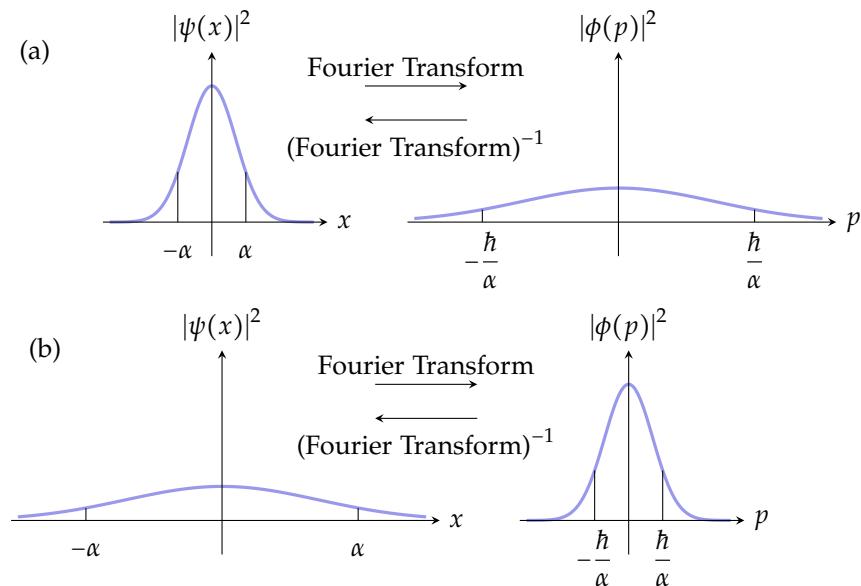


Figure 30.1: Comparison of Gaussian wave functions in position and momentum space. A narrow function in position space will be broad in momentum space and visa versa.

While we have specifically looked at Gaussian wave functions, what the uncertainty relation (30.11) tells us is that, it is generally true that a wave function that is concentrated in a small region of space, meaning we can be more certain about its position, will be spread out in momentum space, meaning we will be less certain about its momentum (and visa versa).

This is at the heart of wave-particle duality. The more concentrated the position space wave function is, the more particle-like the state is with a relatively definite position. On the other hand the

more concentrated the momentum space wave function is, the more like a wave-like the state is with a relatively definite wave-length.

Extreme wave-particle duality

If we want to be *completely* certain about a particle's position, we could look at what happens when the width of our wave function goes to zero. As mentioned in Chapter 29, if take the limit as $\alpha \rightarrow 0$,

$$\lim_{\alpha \rightarrow 0} \left(\frac{1}{\alpha^2 \pi} \right)^{1/4} e^{-(x-x_0)^2/2\alpha^2} = \delta(x - x_0), \quad (30.17)$$

which gives us a delta function.

Discussion 30.9

Explain why (30.17) does not correspond to a physical state. Show or review why in momentum space, (30.17), corresponds to $\phi_{x_0}(p) = \frac{1}{\sqrt{2\pi\hbar}} e^{-ipx_0/\hbar}$. With reference to the form of ϕ_{x_0} , explain why if we were completely certain about a particle's position we would be completely uncertain about its momentum.

This also gives us another way of seeing why the delta function, (30.17), is an eigenfunction of position – if we are certain about an observable we must have an eigenfunction of the corresponding operator. Although (30.17) does not correspond to a physical state, for some purposes, we can use it to model situations where our uncertainty in a particles positions is negligible.

Discussion 30.10

Explain how if we want to be *completely* certain about a particle's momentum, we are lead to the delta function $\delta(p - p_0)$ which has th position space representation $\varphi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar}$. Discuss why if we were completely certain about a particle's momentum we would be completely uncertain about its position. Once again, although $\varphi_p(x)$ is not a physical state, it is useful for modelling situations where our uncertainty in a particle's momentum is negligible.

Some of the results of this discussion are summaried in Table 30.3 and illustrated in Figure 30.2. Figure 30.3 illustrates the position and momentum eigenvalue equations in different representations.

Discussion 30.11

Write a summary of the important points and formulae in this chapter.

	Position Space	Momentum Space
Position eigenstate	$ x_0\rangle \doteq \delta(x - x_0)$	$ x_0\rangle \doteq \frac{1}{\sqrt{2\pi\hbar}} e^{-ipx_0/\hbar}$
Momentum eigenstate	$ p_0\rangle \doteq \frac{1}{\sqrt{2\pi\hbar}} e^{ip_0 x/\hbar}$	$ p_0\rangle \doteq \delta(p - p_0)$

Table 30.3: Summary of momentum and position eigenfunctions

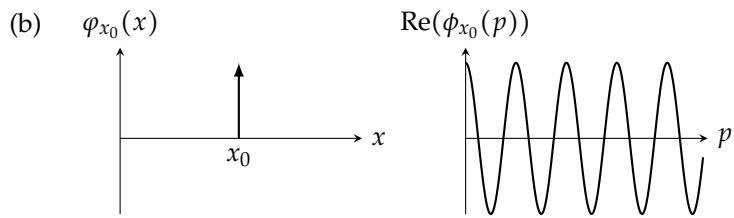
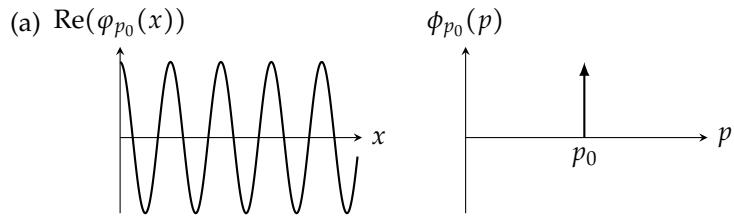


Figure 30.2: (a) Momentum eigenstate wave function and its corresponding delta-function momentum distribution, and (b) position eigenstate wave function and its corresponding infinite extent momentum distribution.

$\hat{x}|x_0\rangle = x_0|x_0\rangle$

Position space Momentum space

$$x\delta(x - x_0) = x_0\delta(x - x_0)$$

$$i\hbar \frac{d}{dp} \left(\frac{1}{\sqrt{2\pi\hbar}} e^{-ipx_0/\hbar} \right) = x_0 \left(\frac{1}{\sqrt{2\pi\hbar}} e^{-ipx_0/\hbar} \right)$$

Figure 30.3: Different representations of the position and momentum eigenvalue equations

$\hat{p}|p_0\rangle = p_0|p_0\rangle$

Position space Momentum space

$$-i\hbar \frac{d}{dx} \left(\frac{1}{\sqrt{2\pi\hbar}} e^{ip_0 x/\hbar} \right) = p_0 \left(\frac{1}{\sqrt{2\pi\hbar}} e^{ip_0 x/\hbar} \right)$$

$$p\delta(p - p_0) = p_0\delta(p - p_0)$$

31

Scattering and tunnelling

Unbound states and scattering

This reading is based on [McIntyre et al. 2012, Section 6.3-6]

To complete our introduction to quantum mechanics, we look at unbound states of potential wells. Unbound states have energy greater than the potential at infinity. This is in contrast to bound states whose energy is less than the potential at infinity. A potential well with bound and unbound states is illustrated in Figure 31.1.

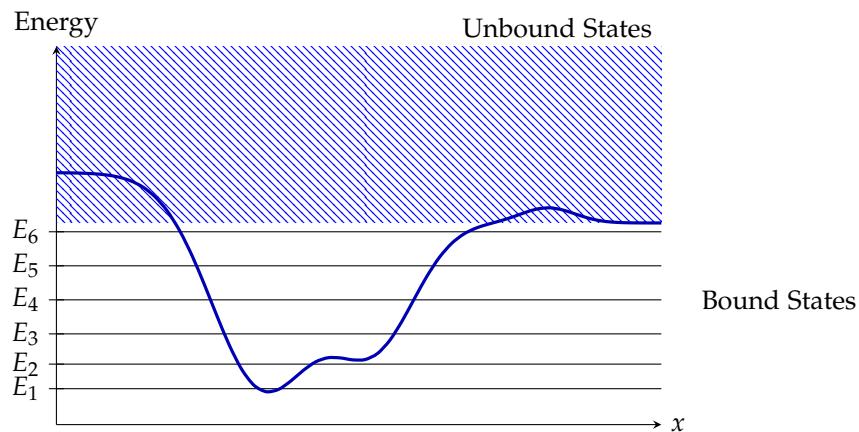


Figure 31.1: Bound ($E < E(\infty)$) and unbound ($E > E(\infty)$) states of a potential well.

Bound states are constrained to “fit” in a potential which leads to their energy being quantised. Unbound states are not constrained to “fit” inside a potential and their energy is continuous. The unbound states are however affected by the potential energy well which causes the wave to “scatter”. For this reason we also call the unbound states of a potential well **scattering states**.

Finite square well scattering states

To study the finite square well scattering states it turns out to be convenient to take energy to be zero at $x = \pm\infty$:

$$V(x) = \begin{cases} 0, & x < -a \\ -V_0, & -a < x < a \\ 0, & x > a. \end{cases} \quad (31.1)$$

With this choice, bound states have $E < 0$ and scattering states have $E > 0$. Taking $V = 0$ is a common choice when studying scattering problems. The potential (31.1) is shown in Figure 31.2.

Now that we have a potential, the next step is to solve the energy eigenvalue equation:

$$\begin{aligned} \left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} - V_0 \right) \varphi_E(x) &= E \varphi_E(x), & |x| < a \\ \left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} - 0 \right) \varphi_E(x) &= E \varphi_E(x), & |x| > a \end{aligned} \quad (31.2)$$

Defining

$$\begin{aligned} k_1 &= \sqrt{\frac{2mE}{\hbar^2}}, \\ k_2 &= \sqrt{\frac{2m(E + V_0)}{\hbar^2}}, \end{aligned} \quad (31.3)$$

(31.2) becomes

$$\begin{aligned} \frac{d^2 \varphi_E(x)}{dx^2} &= -k_2^2 \varphi_E(x), & |x| < a \\ \frac{d^2 \varphi_E(x)}{dx^2} &= -k_1^2 \varphi_E(x), & |x| > a \end{aligned} \quad (31.4)$$

Since we will be interested in the reflection and transmission of an incoming wave, we write solutions to (31.4) in terms of exponentials which are momentum eigenstates:

$$\varphi_E(x) = \begin{cases} Ae^{ik_1 x} + Be^{-ik_1 x}, & x < -a \\ Ce^{ik_2 x} + De^{-ik_2 x}, & -a < x < a \\ Fe^{ik_1 x} + Ge^{-ik_1 x}, & x > a. \end{cases} \quad (31.5)$$

Equation (31.5) has seven unknowns : A, B, C, D, F, G and E but only 4 boundary conditions and the energy is **not** quantised.

In a typical scattering experiment, particles are shot at each other to study their interactions. You will probably have heard of the Rutherford scattering experiment which demonstrated the atomic nature of matter. In that experiment, α -particles (i.e He²⁺ ions) are shot at thin sheets of gold.

In studying scattering experiments, we will consider one particle, the target, as fixed and the other particle, the projectile, as moving. The potential energy then corresponds to the interaction between particles. Our projectile comes from some source which we take to be $x = -\infty$. We assume that the incoming particle has some energy E . Figure 31.3 illustrates a scattering experiment for the finite square well potential.

In Figure 31.3, there is an incoming wave, with amplitude, A , a reflected wave with amplitude, B , and a transmitted wave with amplitude, F . Given the set up of the experiment, there is no wave coming in from $x = \infty$, so we set $G = 0$. With this setup, our wave function is¹,

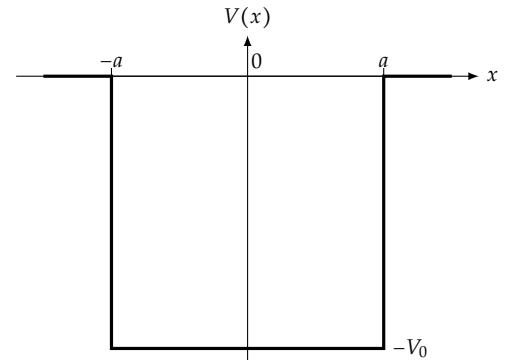


Figure 31.2: Finite square well.

¹ Note that we have only considered one incoming momentum eigenstate, $Ae^{ik_1 x}$, but a more realistic normalisable incoming state would be a wave packet superposition of momentum eigenstates. After having solved the problem at some momentum k_1 , as we are doing here, one can reconstruct the full wave packet behaviour by integrating over the incoming momentum distribution.

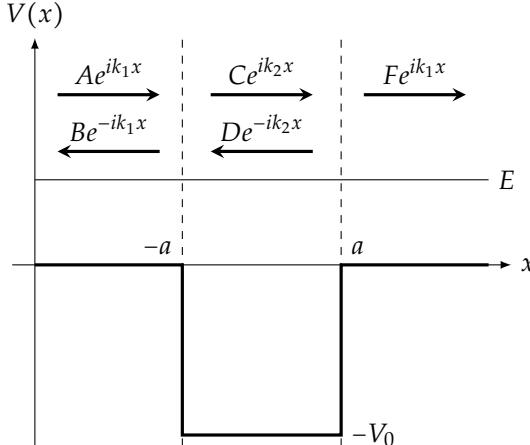


Figure 31.3: Finite square well scattering experiment.

$$\varphi_E(x) = \begin{cases} Ae^{ik_1 x} + Be^{-ik_1 x}, & x < -a \\ Ce^{ik_2 x} + De^{-ik_2 x}, & -a < x < a \\ Fe^{ik_1 x}, & x > a. \end{cases} \quad (31.6)$$

This means that we now have 5 unknowns and 4 boundary conditions. Now, if we increase the amplitude of the incoming wave, there will be an increase in the amplitude of the transmitted and reflected waves. This means that we are really interested in the ratios B/A and F/A .

Now applying boundary conditions at $x = \pm a$ to (31.6), we get

$$\begin{aligned} \varphi_E(-a) : \quad & Ae^{-ik_1 a} + Be^{ik_1 a} = Ce^{-ik_2 a} + De^{ik_2 a} \\ \frac{d\varphi_E(x)}{dx} \Big|_{x=-a} : \quad & ik_1 Ae^{-ik_1 a} - ik_1 Be^{ik_1 a} = ik_2 Ce^{-ik_2 a} - ik_2 De^{ik_2 a} \\ \varphi_E(a) : \quad & Ce^{ik_2 a} + De^{-ik_2 a} = Fe^{ik_1 a} \\ \frac{d\varphi_E(x)}{dx} \Big|_{x=a} : \quad & ik_2 Ce^{ik_2 a} - ik_2 De^{-ik_2 a} = ik_1 Fe^{ik_1 a} \end{aligned} \quad (31.7)$$

Since C and D relate to waves inside the potential, which we can not usually measure directly, our first step is to eliminate them. One can solve the last two equations in (31.7) for C and D in terms of F . Then eliminating C and D from the first two equations, and solving for the ratios F/A and B/A one finds

$$\begin{aligned} \frac{F}{A} &= \frac{e^{-2ik_1 a}}{\cos(2k_2 a) - i \frac{k_2^2 + k_1^2}{2k_1 k_2} \sin(2k_2 a)} \\ \frac{B}{A} &= i \frac{F}{A} \frac{k_2^2 - k_1^2}{2k_1 k_2} \sin(2k_2 a) \end{aligned} \quad (31.8)$$

Transmission probability

Now F/A is the ratio of the amplitude of transmitted wave to amplitude of incoming wave. This means that, $|F|^2/|A|^2$, gives the relative probability, T , that an incident particle is transmitted.

T is called the transmission coefficient and from (31.8) one can find

$$T = \frac{|F|^2}{|A|^2} = \frac{1}{1 + \frac{(k_1^2 - k_2^2)^2}{4k_1^2 k_2^2} \sin^2(2k_2 a)} . \quad (31.9)$$

Using (31.3) we can express (31.9) in terms of energy:

$$T = \frac{1}{1 + \frac{V_0^2}{4E(E+V_0)} \sin^2\left(\frac{2a}{\hbar} \sqrt{2m(E+V_0)}\right)} \quad (31.10)$$

Reflection coefficient

We also define, $R = |B|^2/|A|^2$, the relative probability that an incident particle is reflected.

From (31.8) one can show

$$R = \frac{|B|^2}{|A|^2} = \frac{1}{1 + \frac{4k_1^2 k_2^2}{(k_1^2 - k_2^2)^2 \sin^2(2k_2 a)}} \quad (31.11)$$

Since there is no absorption of particles by the well, we must have

$$R + T = 1 \quad (31.12)$$

In classical mechanics, an unbound particle would be slowed by such a well but not reflected. The fact that, $T \neq 0$, demonstrates another wave-like property of quantum particles.

A plot of R and T as a function of energy is shown in Figure 31.4.

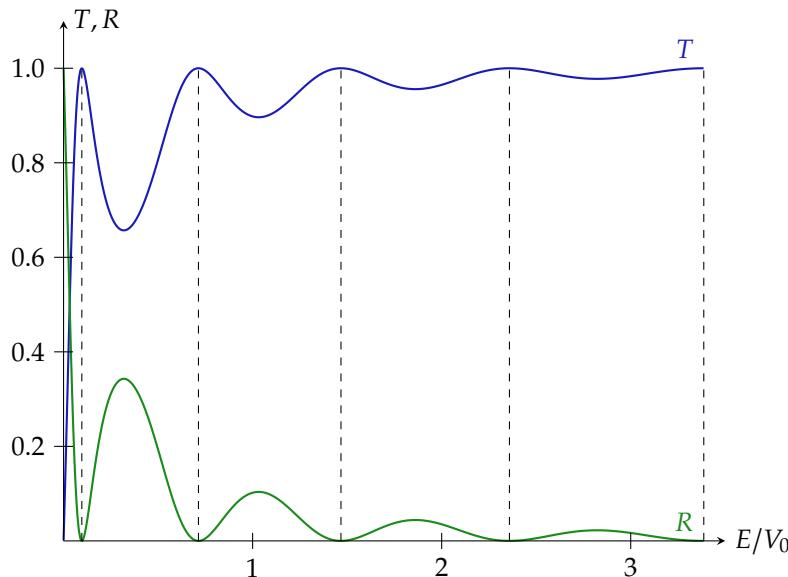


Figure 31.4: R and T as a function of energy for the finite square well. For this plot $\frac{2a}{\hbar} \sqrt{2mV_0} = 12$.

We can see from Figure 31.4 that when the energy is large, $T \rightarrow 1$ and $R \rightarrow 0$. The transmission coefficient is also equal to one for certain special energies, called resonances. Looking at (31.9), we can see that these resonances occur when $\sin(2k_2 a) = 0$ which happens when

$$2k_2a = n\pi \quad (31.13)$$

Writing (31.13) in terms of the wavelength inside the well, $\lambda_2 = 2\pi/k_2$ we get

$$2\frac{2\pi}{\lambda_2}a = n\pi \implies 2a = \frac{n}{2}\lambda_2. \quad (31.14)$$

This means that we get resonance whenever the well contains an integer number of half wavelengths. This is analogous to an effect in optics, where light undergoes multiple reflections going through a glass slab as shown in Figure 31.5. When the wavelength is a half integer multiple of the thickness of the glass, the reflected waves interfere destructively, so that there is no reflected light wave while the transmitted waves interfere constructively.

Discussion 31.1

Rewrite the resonance condition, (31.13) in terms of energy and show that

$$E = -V_0 + \frac{n^2\pi^2\hbar^2}{2m(2a)^2} \quad (31.15)$$

Does the second term on the RHS of (31.15) look familiar? Can you explain why?

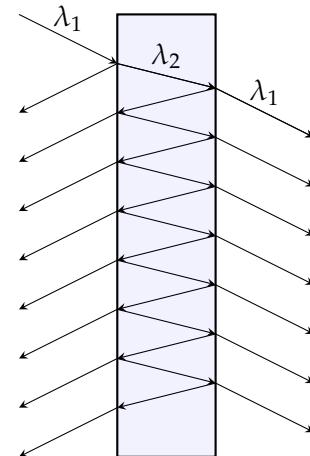


Figure 31.5: Multiple internal reflections in a glass slab.

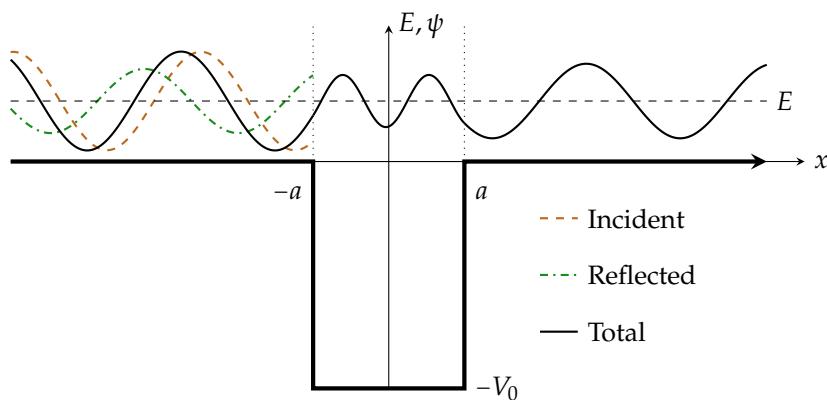


Figure 31.6: Real part of a wave function partially reflected and transmitted through a finite square well.

Tunnelling and potential barriers

Rather than considering a potential well we now consider a potential barrier, illustrated in Figure 31.7,

$$V(x) = \begin{cases} 0, & x < -a \\ V_0, & -a < x < a \\ 0, & x > a. \end{cases} \quad (31.16)$$

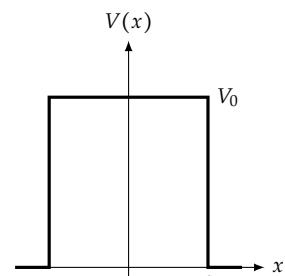


Figure 31.7: Square potential barrier.

The square potential barrier, (31.16), leads to the energy eigenvalue equations

$$\begin{aligned} \left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V_0 \right) \varphi_E(x) &= E \varphi_E(x) , \quad |x| < a \\ \left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} - 0 \right) \varphi_E(x) &= E \varphi_E(x) , \quad |x| > a \end{aligned} \quad (31.17)$$

Rather than solving eq. (31.17), using the principle that laziness can be a virtue, we will reuse our results from the potential well by noting that just changing $V_0 \rightarrow -V_0$ transforms a well into a barrier. Doing this in the expressions for T and R from the potential well leads to

$$T = \frac{1}{1 + \frac{V_0^2}{4E(E-V_0)} \sin^2\left(\frac{2a}{\hbar}\sqrt{2m(E-V_0)}\right)} \quad (31.18)$$

and

$$R = \frac{1}{1 + \frac{4E(E-V_0)}{V_0^2 \sin^2\left(\frac{2a}{\hbar}\sqrt{2m(E-V_0)}\right)}} \quad (31.19)$$

A big difference between the well and the barrier is that we can consider the case in which $E < V_0$. Classically if a particle's energy is below a barrier's height, it can not go beyond the barrier. In quantum mechanics, it is possible for a particle to go through the barrier and come out the other side. This phenomena is called **quantum mechanical tunnelling**. A square potential energy barrier gives us a simple model of tunnelling.

Discussion 31.2

Using the fact that $\sin x = \frac{e^{ix} - e^{-ix}}{2i}$ and $\sinh x = \frac{e^x - e^{-x}}{2}$, show that

$$\sin(ix) = i \sinh(x) \quad (31.20)$$

Notice that when $E < V_0$, the argument of the "sin" in eqs. (31.18) and (31.19) becomes imaginary. Using Discussion 31.2, eqs. (31.18) and (31.19) become

$$T = \frac{1}{1 + \frac{V_0^2}{4E(V_0-E)} \sinh^2\left(\frac{2a}{\hbar}\sqrt{2m(V_0-E)}\right)} \quad (31.21)$$

$$R = \frac{1}{1 + \frac{4E(V_0-E)}{V_0^2 \sinh^2\left(\frac{2a}{\hbar}\sqrt{2m(V_0-E)}\right)}} \quad (31.22)$$

Discussion 31.3

Using eq. (31.20) explain how we go from eqs. (31.18) and (31.19) to eqs. (31.21) and (31.22).

Note that classically, the transmission coefficient T would be zero. Plotting (31.18)-(31.22) together leads to the Figure 31.8

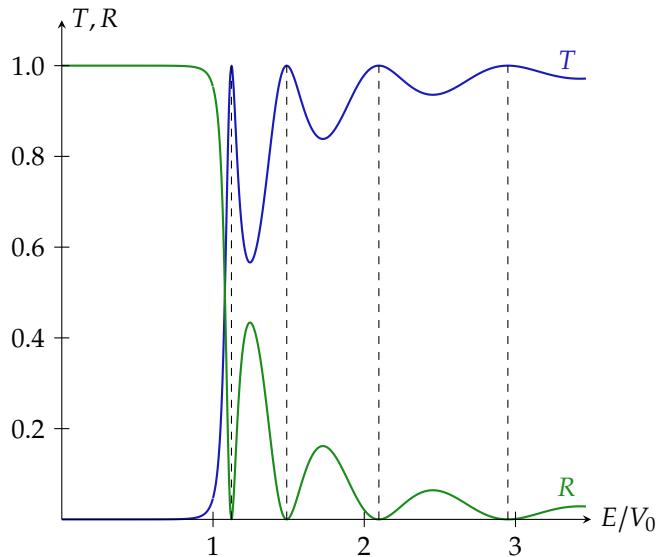


Figure 31.8: Tunnelling reflection and transmission coefficients. For this plot $\frac{2a}{\hbar} \sqrt{2mV_0} = 9$.

From the figure we see that when, $E/V_0 < 1$, T is small and the probability of tunnelling is nearly zero except near the top of the barrier. When the energy is larger than the energy of the barrier, i.e. $E/V_0 > 1$ (and we have scattering rather than tunnelling), transmission becomes large as expected and goes to 1 for large E . Notice that we still get resonances as for the finite square well scattering. A plot of a tunneling wave function is shown in Figure 31.9. In the forbidden region we have real exponential solutions not complex exponentials.

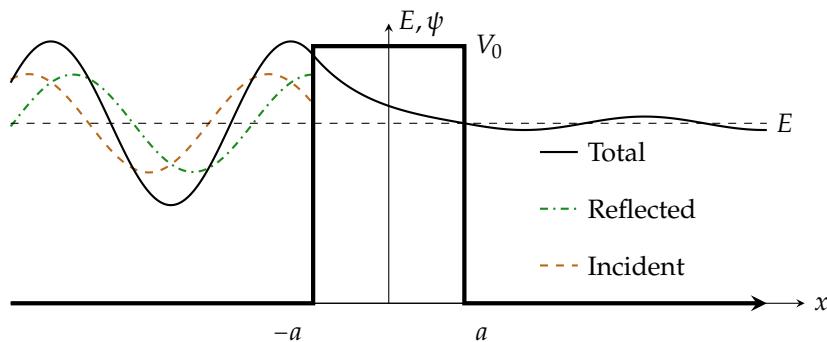


Figure 31.9: Real part of a tunnelling wave function

If you are curious, you can play with simulations of more realistic wave packets interacting with various potentials on the PhET website here².



An impressive application of quantum tunnelling is the scanning tunneling microscope which won Gerd Binnig and Heinrich Rohrer the Nobel Prize in physics in 1986. As shown in fig. 31.10, a tiny, very sharp conduction tip is brought close to a sample. The space between the tip and the sample acts like a potential barrier. Modelling this situation as a square potential

barrier, the tunneling probability is approximately

$$T \propto e^{-2qd}, \quad (31.23)$$

where d is the distance between the tip and the sample and $q = \sqrt{2m(V_0 - E)/\hbar^2}$ is the decay constant in the barrier. In the microscope, a small voltage is applied which induces a current from the tip to the sample. Since this current is only due to tunnelling it will be proportional to the tunnelling probability so that

$$I = I_0 e^{-2qd}. \quad (31.24)$$

The exponential factor makes the current very sensitive to the distance between the tip and the sample allowing measurements to be made in the nanometre range. Moving (or scanning) the tip over the sample while measuring the current allows one to determine the shape of the sample surface on an atomic scale. For example, fig. 31.11 shows a 7nm section of a carbon nanotube.

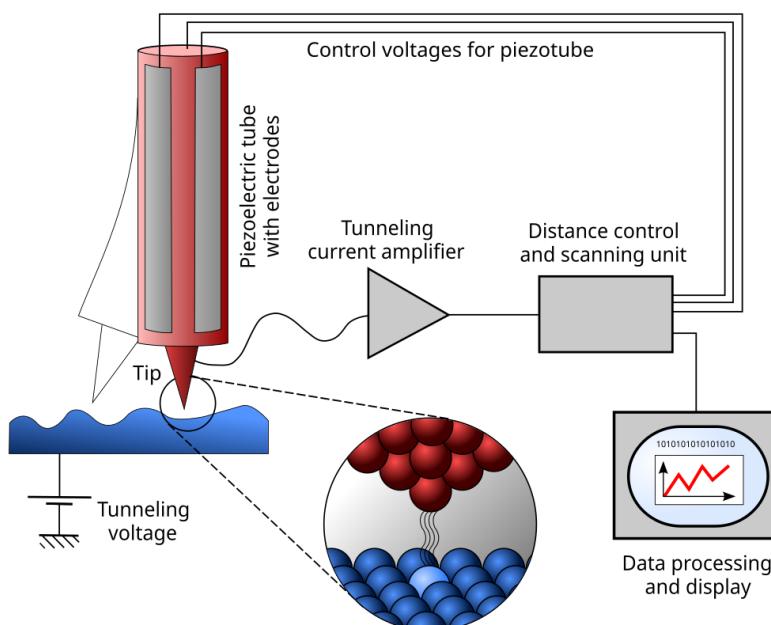


Figure 31.10: Schematic of a scanning tunnelling microscope

Attribution: Image by Michael Schmid and Grzegorz Pietrzak, License: CC BY-SA 2.0 AT.

Discussion 31.4

Write a summary of the important points and formulae in this chapter.

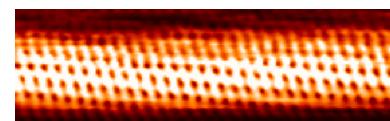


Figure 31.11: 7 nm long part of a single-walled carbon nanotube

Attribution: Image by Taner Yildirim (The National Institute of Standards and Technology - NIST), Public domain, via Wikimedia Commons

32

Quantum Mechanics in 3D

We saw that the Schrödinger equation in position space can be written

$$i\hbar \frac{\partial}{\partial t} \psi = \hat{H}\psi . \quad (32.1)$$

In three dimensions we have the Hamiltonian

$$H = \frac{1}{2m}mv^2 + V = \frac{1}{2m} \left(p_x^2 + p_y^2 + p_z^2 \right) + V(x, y, z, t) . \quad (32.2)$$

Generalising the position space prescription for \hat{p}_x we get:

$$\hat{p}_x \doteq \frac{\hbar}{i} \frac{\partial}{\partial x} , \quad \hat{p}_y \doteq \frac{\hbar}{i} \frac{\partial}{\partial y} , \quad \hat{p}_z \doteq \frac{\hbar}{i} \frac{\partial}{\partial z} , \quad (32.3)$$

which can compactly be written $\hat{p} \doteq (\hbar/i)\vec{\nabla}$. We also generalise the the prescription for \hat{x} to

$$\hat{x} \doteq x , \quad \hat{y} \doteq y , \quad \hat{z} \doteq z . \quad (32.4)$$

Discussion 32.1

Show that

$$\begin{aligned} [\hat{x}_i, \hat{x}_j] &= 0 , \\ [\hat{p}_i, \hat{p}_j] &= 0 , \\ [\hat{x}_i, \hat{p}_j] &= i\hbar\delta_{ij} , \end{aligned} \quad (32.5)$$

where $x_1 = x$, $x_2 = y$, $x_3 = z$, $p_1 = p_x$, $p_2 = p_y$ and $p_3 = p_z$.

Discussion 32.2

Show that \hat{x} and \hat{p} are Hermitian, or in other words that

$$\begin{aligned} \hat{x}_i &= \hat{x}_i^\dagger \\ \hat{p}_i &= \hat{p}_i^\dagger . \end{aligned} \quad (32.6)$$

What is the significance of this?

Discussion 32.3

How does the momentum space representation of \hat{p} and \hat{x} generalise to three dimensions?

Using (32.3)-(32.4), the Hamiltonian operator in position space is

$$\begin{aligned}\hat{H} &\doteq \frac{1}{2m} \left(\left(\frac{\hbar}{i} \frac{\partial}{\partial x} \right)^2 + \left(\frac{\hbar}{i} \frac{\partial}{\partial y} \right)^2 + \left(\frac{\hbar}{i} \frac{\partial}{\partial z} \right)^2 \right) + V(x, y, z, t) \\ &\doteq \frac{-\hbar^2}{2m} \nabla^2 + V(x, y, z, t)\end{aligned}\quad (32.7)$$

where

$$\nabla^2 = \vec{\nabla} \cdot \vec{\nabla} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} . \quad (32.8)$$

is called the **Laplacian**. Putting (32.7) into (32.1) results in

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + V\psi . \quad (32.9)$$

If the potential, V , is independent of time, we can find separable solutions to (32.9) of the form

$$\psi(x, y, z, t) = \varphi(x, y, z) e^{iEt/\hbar} \quad (32.10)$$

where, φ , which satisfies the energy eigenvalue equation¹, $\hat{H}\varphi = E\varphi$, or in other words

$$\frac{-\hbar^2}{2m} \nabla^2 \varphi + V\varphi = E\varphi . \quad (32.11)$$

¹ ie. time-independent Schrödinger equation

Discussion 32.4

Suppose that we have a time-independent potential. Generalising the derivation of (20.9) in chapter 20, show that if we make the separation ansatz

$$\psi(x, y, z, t) = \varphi(x, y, z)f(t) .$$

for (32.9), we find

$$\psi(x, y, z, t) = \varphi(x, y, z)e^{-iEt/\hbar} .$$

where $\varphi(x, y, z)$ satisfies the energy eigenvalue equation (32.11).

Assuming completeness, the general solution² is then a sum of stationary states,

$$\psi(\vec{r}, t) = \sum_n c_n \varphi_n(\vec{r}) e^{-iE_n t/\hbar} . \quad (32.12)$$

where the φ_n satisfy (32.11) with energy eigenvalue E_n .

Extending the idea of a probability density to three dimensions, the probability of finding a particle in an infinitesimal volume is,

² Equation (32.12) applies to bound states. Scattering states need to be constructed out of a wave packets whose components satisfy (32.11) with $E > V(r = \infty)$.

$|\psi|^2 dV$, where, $d^3\vec{r} = dx dy dz$, so that the normalisation condition becomes

$$\int |\psi|^2 dV = 1, \quad (32.13)$$

where the integral is over all space.

💬 Discussion 32.5

What is dV in Cartesian coordinates and what are the limits of integration of (32.13)?

Schrödinger equation in spherical coordinates

Now we'll specialise to the case where we have a spherically symmetric potential, $V = V(r)$. Since the potential only depends on r it is good to use spherical coordinates:

$$\begin{aligned} z &= r \cos \theta \\ x &= r \sin \theta \cos \phi \\ y &= r \sin \theta \sin \phi \end{aligned} \quad (32.14)$$

💬 Discussion 32.6

The definition of θ and ϕ is shown in Figure 32.1. Can you show how we get (32.14) from Figure 32.1?

In spherical coordinates, the Laplacian becomes³

$$\nabla^2 \varphi = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \varphi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \varphi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \varphi}{\partial \phi^2}. \quad (32.15)$$

Then making the ansatz $\varphi(r, \theta, \phi) = R(r)Y(\theta, \phi)$ and using (32.15), the Schrödinger equation (32.11), after a bit of rearrangement, can be written

$$\begin{aligned} \frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) - \frac{2mr^2}{\hbar^2} (V(r) - E) \\ = -\frac{1}{Y} \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Y}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 Y}{\partial \phi^2} \right). \end{aligned} \quad (32.16)$$

💬 Discussion 32.7

Show how we get (32.16).

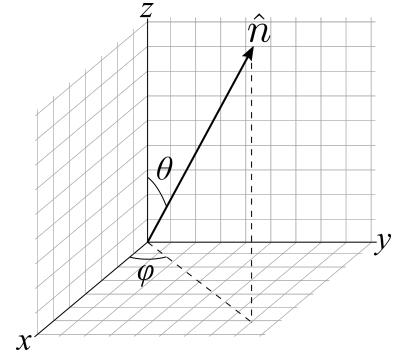


Figure 32.1: General direction along which to measure the spin component.

Modification of [this image](#)
Attribution: Andeggs, Public domain, via Wikimedia Commons

³ For a derivation of (32.15) much too long to fit into the margin, see [Pérez-Martínez et al. 2021].

The LHS of (32.16) is only a function of r and the RHS is only a function of (θ, ϕ) – as is usual for separation of variables we conclude that each side of the equation must be constant. For later convenience we call the separation constant, $l(l+1)$, so that we get

the following two equations

$$\frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) - \frac{2mr^2}{\hbar^2} (V(r) - E) = l(l+1), \quad (32.17)$$

$$\frac{1}{Y} \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Y}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 Y}{\partial \phi^2} \right) = -l(l+1). \quad (32.18)$$

Angular equations

Defining $Y(\theta, \phi) = \Theta(\theta)\Phi(\phi)$, and substituting into (32.18), we get (after some rearrangement)

$$\begin{aligned} & \frac{1}{\Theta} \left(\sin \theta \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) \right) + l(l+1) \sin^2 \theta \\ &= -\frac{1}{\Phi} \frac{d^2 \Phi}{d\phi^2}. \end{aligned} \quad (32.19)$$

Discussion 32.8

Show how we get (32.19).

Letting our separation constant be m^2 , we get

$$0 = \frac{d^2 \Phi}{d\phi^2} + m^2 \Phi, \quad (32.20)$$

$$0 = \sin \theta \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) + (l(l+1) \sin^2 \theta - m^2) \Theta. \quad (32.21)$$

Discussion 32.9

Explain how we get from (32.19) to (32.20) and (32.21).

The equation for Φ is easily solved⁴ leading to

$$\Phi = e^{im\phi}. \quad (32.22)$$

Now if we shift $\phi \rightarrow \phi + 2\pi$, we return to the same point in space so that the value of Φ should be the same – this means that we require

$$\Phi(\phi + 2\pi) = \Phi(\phi), \quad (32.23)$$

or in other words

$$e^{im(\phi+2\pi)} = e^{im\phi} \implies e^{im2\pi} = 1 \implies m \in \mathbb{Z}. \quad (32.24)$$

The equation for Θ , (32.21), is not so simple but fortunately the solution is known⁵:

$$\Theta(\theta) = AP_l^m(\cos \theta), \quad (32.25)$$

where A is an integration constant and $P_l^m(x)$ are called the **associated Legendre functions** given by:

$$P_l^m(x) = (1-x^2)^{|m|/2} \left(\frac{d}{dx} \right)^{|m|} P_l(x), \quad (32.26)$$

⁴ Strictly speaking there are two solutions $e^{im\phi}$ and $e^{-im\phi}$ but we can cover both cases by letting m be positive or negative.

⁵ The solution can be found using the same series solution method we used for the Harmonic oscillator. If you are interested, any good mathematical physics book should cover solving (32.21). We will however look at using an operator method for finding the $Y(\theta, \phi)$ in Readings 33 and 34. Equation (32.25) may concern you since (32.20) is a second order differential equation so there should be two linearly independent solutions for any l and m . In fact such solutions do exist but they blow up at $\theta = 0$ or π which makes them physically unacceptable.

where the l 'th Legendre polynomial is defined by

$$P_l(x) = \frac{1}{2^l l!} \left(\frac{d}{dx} \right)^l (x^2 - 1)^l . \quad (32.27)$$

So for example

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= \frac{1}{2}(2x) = x \\ P_2(x) &= \frac{1}{2^2 2!} \left(\frac{d}{dx} \right)^2 (x^2 - 1)^2 = \frac{1}{8} \frac{d}{dx} (2(x^2 - 1)2x) = \frac{1}{2} (3x^2 - 1) \\ P_3(x) &= \frac{1}{2} (5x^3 - 3x) \\ P_4(x) &= \frac{1}{8} (35x^4 - 30x^2 + 3) \\ \text{etc...} & \end{aligned} \quad (32.28)$$

From (32.26) we see that $P_l(x)$ is a polynomial of degree l but $P_l^m(x)$ is in general not a polynomial – if m is odd, there is a factor of $\sqrt{1-x^2}$, so for example

$$P_2^1 = (1-x^2)^{1/2} \frac{d}{dx} \left(\frac{1}{2} (3x^2 - 1) \right) = 3x\sqrt{1-x^2} . \quad (32.29)$$

What we need is $P_l^m(\cos \theta)$ and⁶ $\sqrt{1-\cos^2 \theta} = |\sin \theta|$, so $\Theta(\theta)$ is always a polynomial in $\cos \theta$, although possibly multiplied by $\sin \theta$.

Notice that l must be an integer for (32.27) to make sense and for $|m| > l$, $P_l^m = 0$. This means that for any given l , there are $2l+1$ possible values of m :

$$l = 0, 1, 2, \dots ; \quad m = -l, -l+1, \dots, -1, 0, 1, \dots, l . \quad (32.30)$$

Discussion 32.10

Does (32.30) look familiar? Where have you seen something similar before? What is different? All will be revealed in Reading 34.

As shown Figure 32.3, in spherical coordinates, $dV = r^2 \sin \theta dr d\theta d\phi$, so

$$\begin{aligned} 1 &= \int |\psi|^2 dV = \int_0^\infty \int_0^\pi \int_0^{2\pi} |\psi|^2 r^2 \sin \theta dr d\theta d\phi \\ &= \left(\int_0^\infty |R|^2 r^2 dr \right) \left(\int_0^\pi \int_0^{2\pi} |Y|^2 \sin \theta d\theta d\phi \right) , \end{aligned} \quad (32.31)$$

and if we independently normalise

$$1 = \int_0^\infty |R|^2 r^2 dr \quad (32.32)$$

$$1 = \int_0^\pi \int_0^{2\pi} |Y|^2 \sin \theta d\theta d\phi , \quad (32.33)$$

then (32.31) will be satisfied. The normalised functions satisfying (32.33) are called **spherical harmonics** which are given by⁷

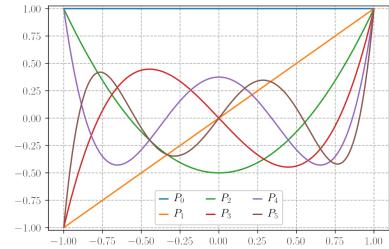


Figure 32.2: Graphs of Legendre polynomials (up to $n = 5$). [source](#).

⁶ In general $\sqrt{1-\cos^2 \theta} = |\sin \theta|$ but since, in conventional spherical coordinates $0 \leq \theta \leq \pi$, $\sin \theta$ is always positive.

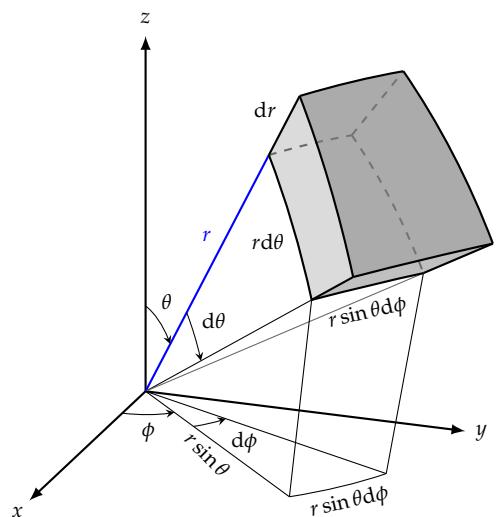


Figure 32.3: Spherical volume element
Image attribution: Alexandros Tsagkaropoulos, original image source

⁷ Although we will not look at it explicitly in this course, we will learn techniques in Reading 34 that will allow you obtain expressions for the spherical harmonics.

$$Y_l^m(\theta, \phi) = (-1)^{(m+|m|)/2} \sqrt{\frac{(2l+1)}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos \theta) e^{im\phi}, \quad (32.34)$$

The strange sign convention in (32.34) ensures the following identity

$$Y_l^{-m}(\theta, \phi) = (-1)^m (Y_l^m(\theta, \phi))^* \quad (32.35)$$

The spherical harmonics are orthogonal

$$\int_0^\pi \int_0^{2\pi} [Y_l^m(\theta, \phi)]^* Y_{l'}^{m'}(\theta, \phi) \sin \theta d\theta d\phi = \delta_{ll'} \delta_{mm'} \quad (32.36)$$

Radial equation

If we let

$$u(r) = rR(r), \quad (32.37)$$

the r dependent part of the Schrödinger equation, (32.17), becomes

$$-\frac{\hbar^2}{2m} \frac{d^2u}{dr^2} + \left(V + \frac{\hbar^2}{2m} \frac{l(l+1)}{r^2} \right) u = Eu, \quad (32.38)$$

which is called the **radial equation**.

Discussion 32.11

Show how we go from (32.17) to (32.38).

Notice that (32.38) looks just like the 1D time-independent Schrödinger equation, except that we effectively have a new potential, V_{eff} , with a new term

$$V_{\text{eff}} = V + \frac{\hbar^2}{2m} \frac{l(l+1)}{r^2}. \quad (32.39)$$

The extra piece is called the **centrifugal term**, and it tends to throw a particle outward – just like the centrifugal inertial force. The effective potential, V_{eff} , for the Hydrogen atom, for a few l values, is shown in Figure 32.4.

Discussion 32.12

Looking at Figure 32.4 can you see why we say that the centrifugal term tends to throw the particle outward?

Finally from the definition of u and (32.32), we see that the normalisation condition for u is

$$\int_0^\infty |u|^2 dr = 1. \quad (32.40)$$

Discussion 32.13

Write a summary of the important points and formulae in this chapter.

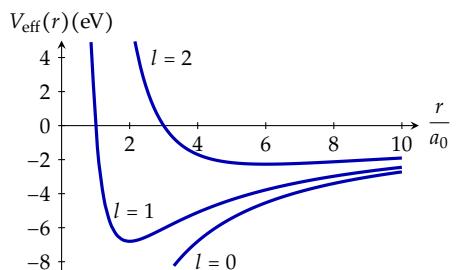


Figure 32.4: Effective potential for the Hydrogen atom for $l = 0, 1, 2$.

33

Angular momentum

Before studying the hydrogen atom, we will look at angular momentum, which as we will see, plays an important role when we have a spherically symmetric potential, $V = V(r)$.

Classically the angular momentum of a particle is

$$\vec{L} = \vec{r} \times \vec{p} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ x & y & z \\ p_x & p_y & p_z \end{vmatrix} \quad (33.1)$$

so

$$\begin{aligned} L_x &= yp_z - zp_y, \\ L_y &= zp_x - xp_z, \\ L_z &= xp_y - yp_x, \end{aligned} \quad (33.2)$$

and to get the position space representation of the quantum mechanical operator, \hat{L} , we replace \vec{p} with $(\frac{\hbar}{i})\vec{\nabla}$.

Discussion 33.1

Show that \hat{L}_x , \hat{L}_y and \hat{L}_z are hermitian.

33.1 Eigenvalues

We can get the eigenvalues of \hat{L} using just algebraic methods (similar to what we did with the harmonic oscillator)¹. Note that some of the discussion points that follow overlap with what we did in chapter 26, but they are rather fundamental to manipulating operators and it is a good idea for you to review and rederive them again.

Unlike, for instance, the components of \hat{x} , the components of \hat{L} do not commute. To evaluate the commutators between the components of angular momentum, it is helpful to use some basic properties of commutators.

Discussion 33.2

Show that $[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}]$ and $[\hat{A}, \hat{A}] = 0$.

This reading is based on [Griffiths 2016, Section 4.3]

¹ Unlike the harmonic oscillator, we can apparently **only** get certain solutions – corresponding to half-integer spin – if we use the algebraic technique.

Discussion 33.3

Show that $[\hat{A} + \hat{B}, \hat{C}] = [\hat{A}, \hat{C}] + [\hat{B}, \hat{C}]$.

Discussion 33.4

Show that if \hat{B} commutes with both \hat{A} and \hat{C} , then $[\hat{A}\hat{B}, \hat{C}] = \hat{B}[\hat{A}, \hat{C}] = [\hat{A}, \hat{C}]\hat{B}$.

Now

$$\begin{aligned} [\hat{L}_x, \hat{L}_y] &= [\hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \hat{z}\hat{p}_x - \hat{x}\hat{p}_z] \\ &= [\hat{y}\hat{p}_z, \hat{z}\hat{p}_x] - [\hat{y}\hat{p}_z, \hat{x}\hat{p}_z] - [\hat{z}\hat{p}_y, \hat{z}\hat{p}_x] + [\hat{z}\hat{p}_y, \hat{x}\hat{p}_z] \\ &= \hat{y}\hat{p}_x[\hat{p}_z, \hat{z}] + \hat{x}\hat{p}_y[\hat{z}, \hat{p}_z] \\ &= i\hbar(\hat{x}\hat{p}_y - \hat{y}\hat{p}_x) \\ &= i\hbar\hat{L}_z \end{aligned} \quad (33.3)$$

Discussion 33.5

Explain the second, third, fourth and fifth equalities in (33.3).

Notice that from, (33.2), we can cycle from $L_x \rightarrow L_y \rightarrow L_z \rightarrow L_x$ by cyclically permuting $x \rightarrow y \rightarrow z \rightarrow x$. This means that without doing any further calculation, we can read off the other commutators

$$[\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x, \quad [\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y. \quad (33.4)$$

Discussion 33.6

Show how we can use cyclic permutation to go from (33.3) to (33.4)

In summary,

Angular momentum commutators

$$\begin{aligned} [\hat{L}_x, \hat{L}_y] &= i\hbar\hat{L}_z \\ [\hat{L}_y, \hat{L}_z] &= i\hbar\hat{L}_x \\ [\hat{L}_z, \hat{L}_x] &= i\hbar\hat{L}_y \end{aligned} \quad (33.5)$$

Equations (33.5) are the same as the commutation relations that we had for spin. We see that spin angular momentum and orbital angular momentum have something in common. This justifies referring to both as angular momenta. Now, we saw that the total spin operator, $\hat{S}^2 = \hat{\vec{S}} \cdot \hat{\vec{S}}$, commuted with all the spin components – let's check if the same is true for the total orbital angular momentum. Since

$$\vec{L}^2 = \vec{L} \cdot \vec{L} = L_x^2 + L_y^2 + L_z^2. \quad (33.6)$$

we have that

$$[\hat{L}^2, \hat{L}_x] = [\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2, \hat{L}_x] = [\hat{L}_x^2, \hat{L}_x] + [\hat{L}_y^2, \hat{L}_x] + [\hat{L}_z^2, \hat{L}_x] \quad (33.7)$$

Before proceeding with (33.7) we need some more useful commutator identities

❖ Discussion 33.7

Show that

- 1) $[\hat{A}^2, \hat{A}] = 0$
- 2) $[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B}$
- 3) $[\hat{A}^2, \hat{B}] = \hat{A}[\hat{A}, \hat{B}] + [\hat{A}, \hat{B}]\hat{A}$

Using these results (33.7) becomes

$$\begin{aligned} [\hat{L}^2, \hat{L}_x] &= \hat{L}_y[\hat{L}_y, \hat{L}_x] + [\hat{L}_y, \hat{L}_x]\hat{L}_y + \hat{L}_z[\hat{L}_z, \hat{L}_x] + [\hat{L}_z, \hat{L}_x]\hat{L}_z \\ &= \hat{L}_y(-i\hbar\hat{L}_z) + (-i\hbar\hat{L}_z)\hat{L}_y + \hat{L}_z(i\hbar\hat{L}_y) + (i\hbar\hat{L}_y)\hat{L}_z \\ &= 0 \end{aligned} \quad (33.8)$$

❖ Discussion 33.8

Explain the steps in (33.8).

❖ Discussion 33.9

Show that in addition to (33.8) we also have

$$[\hat{L}^2, \hat{L}_y] = [\hat{L}^2, \hat{L}_z] = 0.$$

In summary, \hat{L}^2 commutes with all the components of \vec{L} :

Angular momentum commutators

$$[\hat{L}^2, \hat{L}_x] = 0 \quad (33.9)$$

$$[\hat{L}^2, \hat{L}_y] = 0 \quad (33.10)$$

$$[\hat{L}^2, \hat{L}_z] = 0 \quad (33.11)$$

So, spin and angular momentum have the same commutation relations. We already know what the possible eigenvalues of the spin operator is so (see (11.15)) we expect angular momentum to have the same spectrum². We will not see how we can find out the spectrum of $\hat{\vec{L}}$ using its commutation relations.

We can use “ladder” operators, (similar to \hat{a}_{\pm} for the harmonic oscillator) to determine the spectrum³ of \hat{L}^2 and \hat{L}_z . Letting

$$\hat{L}_{\pm} := \hat{L}_x \pm i\hat{L}_y, \quad (33.12)$$

² Here, the word “spectrum” refers to the set of possible eigenvalues $S^2 = \hbar^2 s(s+1)$, where $s = 0, \frac{1}{2}, 1, \dots$, and $S_z = \hbar m$ where $m = -s, -s+1, \dots, s-1, s$.

³ Since spin and angular momentum have the same commutation relations, the same method applies to spin.

we obtain the following commutators

$$[\hat{L}^2, \hat{L}_\pm] = [\hat{L}^2, \hat{L}_x] \pm i[\hat{L}^2, \hat{L}_y] = 0 , \quad (33.13)$$

and

$$\begin{aligned} [\hat{L}_z, \hat{L}_\pm] &= [\hat{L}_z, \hat{L}_x] \pm i[\hat{L}_z, \hat{L}_y] \\ &= i\hbar\hat{L}_y \pm i(-i\hbar\hat{L}_x) \\ &= \pm\hbar(\hat{L}_x \pm i\hat{L}_y) \\ &= \pm\hbar\hat{L}_\pm . \end{aligned} \quad (33.14)$$

Theorem

If $|\psi\rangle$ is an eigenstate of \hat{L}^2 and \hat{L}_z , with eigenvalues λ and μ respectively, then $\hat{L}_\pm|\psi\rangle$ is also an eigenstate of \hat{L}^2 and \hat{L}_z , with eigenvalues λ and $\mu \pm \hbar$ respectively.

Proof.

$$\hat{L}^2(\hat{L}_\pm|\psi\rangle) \stackrel{(33.13)}{=} \hat{L}_\pm\hat{L}^2|\psi\rangle = \lambda(\hat{L}_\pm|\psi\rangle) , \quad (33.15)$$

and

$$\begin{aligned} \hat{L}_z(\hat{L}_\pm|\psi\rangle) &= [\hat{L}_z, \hat{L}_\pm]|\psi\rangle + \hat{L}_\pm\hat{L}_z|\psi\rangle \\ &= (\mu \pm \hbar)(\hat{L}_\pm|\psi\rangle) \end{aligned} \quad (33.16)$$

□

Discussion 33.10

Explain the equality in the 1st and 2nd lines of (33.16).

So from the theorem above, we see that (as shown in figure 33.1) for a given λ , there is a ladder of states whose \hat{L}_z eigenvalues are separated by \hbar .

If we keep on increasing L_z , eventually it will exceed the total angular momentum which is not possible since

$$\langle L^2 \rangle = \underbrace{\langle L_x^2 \rangle}_{\geq 0} + \underbrace{\langle L_y^2 \rangle}_{\geq 0} + \langle L_z^2 \rangle \implies \langle L_z^2 \rangle \leq \langle L^2 \rangle , \quad (33.17)$$

but

$$\langle \psi | \hat{L}_z^2 | \psi \rangle = \langle \hat{L}_z \psi | \hat{L}_z \psi \rangle = \mu^2 , \quad (33.18)$$

and

$$\langle \psi | \hat{L}^2 | \psi \rangle = \langle \psi | \lambda | \psi \rangle = \lambda . \quad (33.19)$$

Discussion 33.11

Explain the equalities in (33.18).

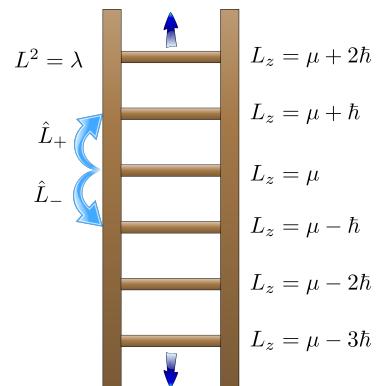


Figure 33.1: Ladder of \hat{L}_z eigenstates for a given \hat{L}^2 eigenvalue λ .

From (33.17), (33.18) and (33.19) we conclude that

$$\mu^2 \leq \lambda , \quad (33.20)$$

which in turn implies that there must be some top rung to our ladder, $|\psi_t\rangle$. This means that at some point, rather than generating a new eigenfunction, we must have that

$$\hat{L}_+ |\psi_t\rangle = 0 . \quad (33.21)$$

Let⁴, $\hbar l$, be the \hat{L}_z eigenvalue of $|\psi_t\rangle$, so

$$\hat{L}_z |\psi_t\rangle = \hbar l |\psi_t\rangle , \quad \hat{L}^2 |\psi_t\rangle = \lambda |\psi_t\rangle . \quad (33.22)$$

To proceed we need the following useful relations

$$\begin{aligned} \hat{L}_\pm \hat{L}_\mp &= (\hat{L}_x \pm i\hat{L}_y)(\hat{L}_x \mp i\hat{L}_y) \\ &= \hat{L}_x^2 + \hat{L}_y^2 + i(\underbrace{\hat{L}_x \hat{L}_y - \hat{L}_y \hat{L}_x}_{i\hbar \hat{L}_z}) \\ &= \hat{L}_x^2 + \hat{L}_y^2 \pm \hbar \hat{L}_z \\ &= \hat{L}^2 - \hat{L}_z^2 \pm \hbar \hat{L}_z , \end{aligned} \quad (33.23)$$

which implies

$$\hat{L}^2 = \hat{L}_\pm \hat{L}_\mp + \hat{L}_z^2 \mp \hbar \hat{L}_z . \quad (33.24)$$

so

$$\begin{aligned} \lambda |\psi_t\rangle &\stackrel{(33.22)}{=} \hat{L}^2 |\psi_t\rangle \\ &\stackrel{(33.24)}{=} (\hat{L}_- \hat{L}_+ + \hat{L}_z^2 + \hbar \hat{L}_z) |\psi_t\rangle \\ &\stackrel{(33.21), (33.22)}{=} (\hbar^2 l^2 + \hbar^2 l) |\psi_t\rangle \\ &= \hbar^2 l(l+1) |\psi_t\rangle , \end{aligned} \quad (33.25)$$

which gives us λ in terms of the highest eigenvalue of \hat{L}_z – ie. $\lambda = \hbar^2 l(l+1)$.

Now returning to $\mu^2 \leq \lambda$, we see there must be a lowest rung as well,

$$\hat{L}_- |\psi_b\rangle = 0 . \quad (33.26)$$

Let, $\hbar \bar{l}$, be the \hat{L}_z eigenvalue of $|\psi_b\rangle$, so

$$\hat{L}_z |\psi_b\rangle = \hbar \bar{l} |\psi_b\rangle , \quad \hat{L}^2 |\psi_b\rangle = \lambda |\psi_b\rangle , \quad (33.27)$$

and

$$\begin{aligned} \hat{L}^2 |\psi_b\rangle &= (\hat{L}_+ \hat{L}_- + \hat{L}_z^2 - \hbar \hat{L}_z) |\psi_b\rangle \\ &= (\hbar^2 \bar{l}^2 - \hbar^2 \bar{l}) |\psi_b\rangle \\ &= \hbar^2 \bar{l}(\bar{l}-1) |\psi_b\rangle , \end{aligned} \quad (33.28)$$

but since the eigenvalue of \hat{L}^2 is constant on the ladder, we must have

$$\bar{l}(\bar{l}-1) = l(l+1) , \quad (33.29)$$

which implies

$$\bar{l} = -l \text{ or } \bar{l} = l+1 , \quad (33.30)$$

⁴ At this point this is an arbitrary choice albeit with the correct units.

 **Discussion 33.12**

Using the quadratic formula, show how (33.30) follows from (33.29)

Hint: $4l^2 + 4l + 1 = (2l + 1)^2$.

We can not have the lowest rung of the ladder, \bar{l} , larger than the highest rung, l , which means we should discard the second option in (33.30).

If we write the eigenvalues of \hat{L}_z as $\hbar m$, we see that m goes in integer steps from $-l$ to l (since the raising operator raises the eigenvalues by \hbar). Suppose our ladder has N rungs, this implies that

$$\begin{aligned} l &= -l + N \quad \text{where } N \text{ is a non-negative integer} \\ \implies 2l &= N \\ \implies l &= \frac{N}{2}, \end{aligned} \tag{33.31}$$

which tells us that l is a non-negative integer or half integer.

So finally we conclude that the eigenstates are characterised by two numbers l and m . We will write these eigenstates as $|lm\rangle$ with

$$\hat{L}^2 |lm\rangle = \hbar^2 l(l+1) |lm\rangle, \quad \hat{L}_z |lm\rangle = \hbar m |lm\rangle, \tag{33.32}$$

where $l = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$ with $m = -l, -l+1, \dots, -1, 0, 1, \dots, l$. As expected, this is the same as the spectrum for \hat{S}_z (11.15). We will however see in Reading 34, that, unlike spin, for angular momentum, m and l must be integers.

 **Discussion 33.13**

Do \hat{L}_{\pm} correspond to any observable quantities?

 **Discussion 33.14**

Write a summary of the important points and formulae in this chapter.

34

Angular momentum and Spherical Harmonics

Eigenfunctions

This reading is based on [Griffiths 2016, Section 4.3]

It turns out that we have already seen the eigenfunctions of \hat{L}^2 and \hat{L}_z – they are the spherical harmonics. To see this we first have to write the position space representation of angular momentum operator in spherical coordinates. Now

$$\hat{\vec{L}} \doteq \left(\frac{\hbar}{i} \right) \vec{r} \times \vec{\nabla}, \quad (34.1)$$

and using the expression for the gradient in spherical coordinates¹

$$\vec{\nabla} = \hat{r} \frac{\partial}{\partial r} + \hat{\theta} \frac{1}{r} \frac{\partial}{\partial \theta} + \hat{\phi} \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi}, \quad (34.2)$$

together with the fact that

$$\vec{r} = r \hat{r}, \quad (34.3)$$

and (using Figure 34.1 and the right-hand rule)

$$\begin{aligned} \hat{r} \times \hat{r} &= 0 \\ \hat{r} \times \hat{\theta} &= \hat{\phi} \\ \hat{r} \times \hat{\phi} &= -\hat{\theta} \end{aligned} \quad (34.4)$$

we get

$$\begin{aligned} \hat{\vec{L}} &\doteq \frac{\hbar}{i} (r \hat{r}) \times \left(\hat{r} \frac{\partial}{\partial r} + \hat{\theta} \frac{1}{r} \frac{\partial}{\partial \theta} + \hat{\phi} \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} \right) \\ &\doteq \frac{\hbar}{i} \left(\hat{\phi} \frac{\partial}{\partial \theta} - \hat{\theta} \frac{1}{\sin \theta} \frac{\partial}{\partial \phi} \right). \end{aligned} \quad (34.5)$$

Discussion 34.1

Fill in the missing details in (34.5).

¹ Note that in (34.2), the hats, $\hat{\cdot}$, refer to unit vectors not operators.

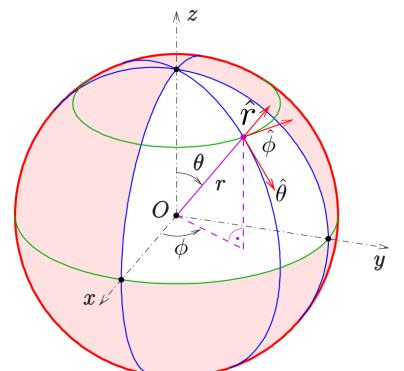


Figure 34.1: Unit vectors in spherical coordinates.

image source <https://commons.wikimedia.org/wiki/File:Kugelkoord-lokale-Basis-s.svg>

Now in Cartesian coordinates,

$$\begin{aligned} \hat{\theta} &= (\cos \theta \cos \phi) \hat{i} + (\cos \theta \sin \phi) \hat{j} - \sin \theta \hat{k} \\ \hat{\phi} &= -\sin \phi \hat{i} + \cos \phi \hat{j} \end{aligned} \quad (34.6)$$

so that

$$\hat{L} \doteq \frac{\hbar}{i} \left[\hat{i} \left(-\sin \phi \frac{\partial}{\partial \theta} - \cos \phi \cot \theta \frac{\partial}{\partial \phi} \right) + \hat{j} \left(\cos \phi \frac{\partial}{\partial \theta} - \sin \phi \cot \theta \frac{\partial}{\partial \phi} \right) + \hat{k} \frac{\partial}{\partial \phi} \right], \quad (34.7)$$

from which we can read off the Cartesian components of \hat{L} and find

$$\hat{L}_{\pm} = \hat{L}_x \pm i \hat{L}_y \doteq \pm \hbar e^{\pm i \phi} \left(\frac{\partial}{\partial \theta} \pm i \cot \theta \frac{\partial}{\partial \phi} \right), \quad (34.8)$$

from which we find

$$\hat{L}_+ \hat{L}_- \doteq -\hbar^2 \left(\frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \cot^2 \theta \frac{\partial^2}{\partial \phi^2} + i \frac{\partial}{\partial \phi} \right), \quad (34.9)$$

and

$$\hat{L}^2 \doteq -\hbar^2 \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right). \quad (34.10)$$

Discussion 34.2

Show how we get (34.8) from (34.7).

Discussion 34.3

Show how we get (34.10).

Hint: Use (33.24)

Let $F_l^m = F_l^m(\theta, \phi)$ be the position space representation of $|lm\rangle$, then

$$\begin{aligned} \hat{L}^2 |lm\rangle &\doteq \hbar^2 l(l+1) F_l^m \\ &\stackrel{(34.10)}{=} -\hbar^2 \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right) F_l^m \end{aligned} \quad (34.11)$$

Furthermore,

$$\begin{aligned} \hat{L}_z |lm\rangle &\doteq \hbar m F_l^m \\ &\stackrel{(34.7)}{=} \frac{\hbar}{i} \frac{\partial}{\partial \phi} F_l^m \end{aligned} \quad (34.12)$$

Cancelling the factors of \hbar , (34.11) and (34.12) give

$$\begin{aligned} - \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right) F_l^m &= l(l+1) F_l^m \\ \frac{\partial}{\partial \phi} F_l^m &= im F_l^m \end{aligned} \quad (34.13)$$

Comparing with (34.13) with (32.18) and (32.21), we see that F_l^m satisfies the same differential equations as the spherical harmonics. In other words, the spherical harmonics are the position space representation of the simultaneous eigenfunctions of \hat{L}^2 and L_z with eigenvalues $l(l+1)$ and m respectively.

Now, we saw in Reading 32 that single-valuedness of the wave function (see (32.24)) implies that m , and hence l must be integers. On the other hand, spin is apparently does not have a position space representation which is why we think of the total spin as simply a intrinsic property of a particle².

² Personally, I have always found this rather mysterious.

Discussion 34.4

- a) What is $\hat{L}_+ Y_l^l$? (No calculation allowed!)
- b) Use the result of (a), together with (34.7), (34.8) and the fact that $\hat{L}_z Y_l^l = \hbar l Y_l^l$, to determine $Y_l^l(\theta, \phi)$, up to a normalization constant.

Hint
$$\frac{1}{f(x)} \frac{df}{dx} = \frac{d \ln(f(x))}{dx}$$

Hint
$$\int \cot \theta d\theta = \ln(\sin \theta) + C$$
 or in other words

$$\frac{d \ln \sin x}{dx} = \cot x.$$

- c) Normalise Y_0^0 .

Discussion 34.5

If

$$Y_2^1(\theta, \phi) = \sqrt{15/8\pi} \sin \theta \cos \theta e^{i\phi}, \quad (34.14)$$

apply the raising operator to find Y_2^2 , up to a normalisation constant. Check you answer using Table 34.1.

Discussion 34.6

If the electron were a classical solid sphere, with radius

$$r_c = \frac{e^2}{4\pi\epsilon_0 mc^2} \quad (34.15)$$

(the so-called classical electron radius, obtained by assuming the electron's mass is attributable to energy stored in its electric field, via the Einstein formula $E = mc^2$), and its angular momentum is $(1/2)\hbar$, then how fast (in m/s) would a point on the "equator" be moving? Does this model make sense? (Actually, the radius of the electron is known experimentally to be much less than r_c , but this only makes matters worse).

Normalisation

Just like we used raising and lowering operators to find normalisation factors for the harmonic oscillator wave functions, we can do the same for spherical harmonics.

Discussion 34.7

Show that $(\hat{L}_+)^{\dagger} = \hat{L}_-$ and similarly $(\hat{L}_-)^{\dagger} = \hat{L}_+$.

Discussion 34.8

Using Discussion 34.7, explain why, $\langle \psi | \hat{L}_- \hat{L}_+ \psi \rangle = \langle \hat{L}_+ \psi | \hat{L}_+ \psi \rangle$, and similarly $\langle \psi | \hat{L}_+ \hat{L}_- \psi \rangle = \langle \hat{L}_- \psi | \hat{L}_- \psi \rangle$.

Discussion 34.9

Recall that

$$\hat{L}_{\pm} \hat{L}_{\mp} = \hat{L}^2 - \hat{L}_z^2 \pm \hbar \hat{L}_z . \quad (33.23')$$

Using Discussion 34.8, show that, up to a phase,

$$\hat{L}_{\pm} |lm\rangle = \hbar \sqrt{l(l+1) - m(m \pm 1)} |l(m \pm 1)\rangle . \quad (34.16)$$

Discussion 34.10

Check that your result for Discussion box 34.5 is consistent with (34.16).

Now, using Discussion (34.4), once we have normalised Y_l^l , we can in principle use (34.16) to obtain the other normalised spherical harmonics with the same L^2 eigenvalue. We will not do this in detail but Table 34.1 list some of the spherical harmonics that you could obtain this way.

l	m	$Y_l^m(\theta, \phi)$
0	0	$Y_0^0 = \sqrt{\frac{1}{4\pi}}$
1	0	$Y_1^0 = \sqrt{\frac{3}{4\pi}} \cos \theta$
	± 1	$Y_1^{\pm 1} = \mp \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\phi}$
2	0	$Y_2^0 = \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1)$
	± 1	$Y_2^{\pm 1} = \mp \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{\pm i\phi}$
	± 2	$Y_2^{\pm 2} = \mp \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{\pm 2i\phi}$
3	0	$Y_3^0 = \sqrt{\frac{7}{16\pi}} (5 \cos^3 \theta - 3 \cos \theta)$
	± 1	$Y_3^{\pm 1} = \mp \sqrt{\frac{21}{64\pi}} \sin \theta (5 \cos^2 \theta - 1) e^{\pm i\phi}$
	± 2	$Y_3^{\pm 2} = \sqrt{\frac{105}{32\pi}} \sin^2 \theta \cos \theta e^{\pm 2i\phi}$
	± 3	$Y_3^{\pm 3} = \mp \sqrt{\frac{35}{64\pi}} \sin^3 \theta e^{\pm 3i\phi}$

Table 34.1: Selection of spherical harmonics

34.1 Plotting spherical harmonics

One way to represent spherical harmonics is using three-dimensional polar plots. In these plots, the spherical harmonics are represented as three dimensional surfaces. The distance, r , of the surface from the origin at a particular value of (θ, ϕ) is given by $|Y_l^m|^2$. For example, $|Y_0^0|^2 = 1/4\pi$ which is constant of a function of θ and ϕ meaning the polar plot surface is a sphere. The polar plots of the first few spherical harmonics are shown in Table 34.2.

Discussion 34.11

In a similar manner to the discussion in the preceding paragraph about $|Y_0^0|$, using the expression in Table 34.1, qualitatively explain the form of the polar plot for $|Y_1^0|^2$ in Table 34.2.

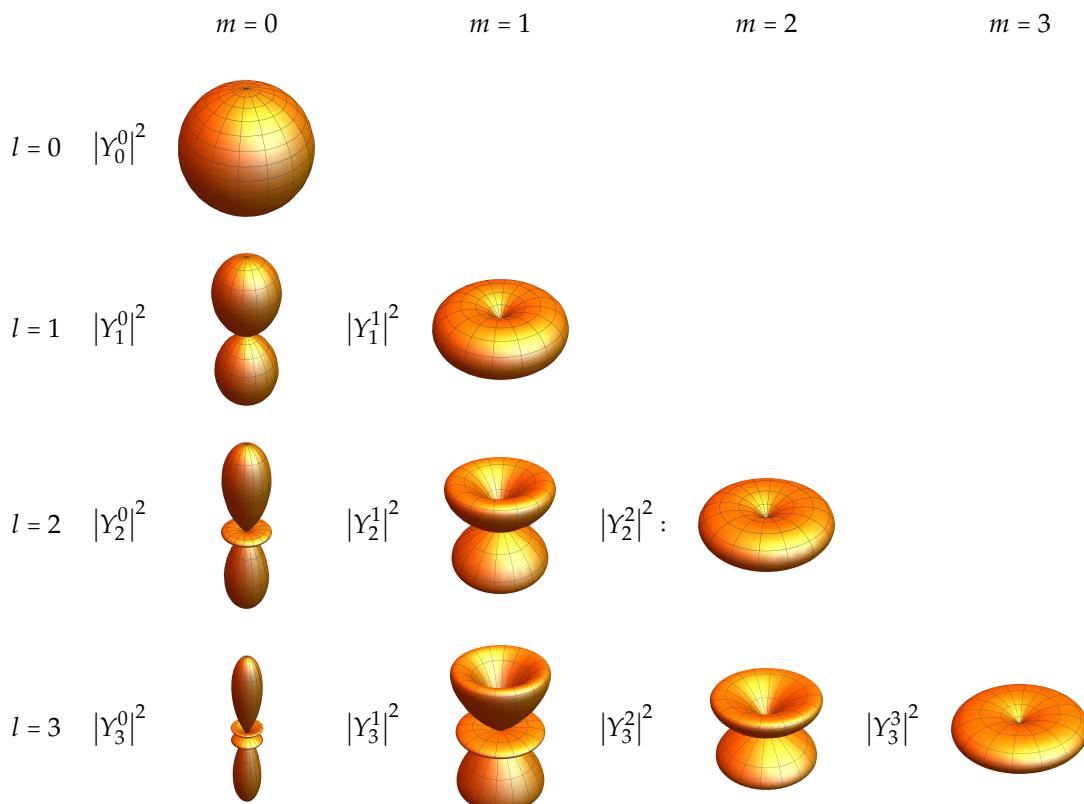


Table 34.2: Spherical Harmonic Polar Plots

Discussion 34.12

Using the expression in Table 34.1, qualitatively explain the form of the polar plot for $|Y_1^0|^2$ in Table 34.1.

❖ Discussion 34.13

Write a summary of the important points and formulae in this chapter.

35

Hydrogen atom

The hydrogen atom consists of an electron and a proton. Putting the proton at the origin, Coulomb's law tells us that the electron is subject to a potential

$$V(r) = -\frac{e^2}{4\pi\epsilon_0} \frac{1}{r}, \quad (35.1)$$

so that the radial equation, (32.38), becomes¹

$$-\frac{\hbar^2}{2m} \frac{d^2u}{dr^2} + \left(-\frac{e^2}{4\pi\epsilon_0} \frac{1}{r} + \frac{\hbar^2 l(l+1)}{2m r^2} \right) u = Eu. \quad (35.2)$$

We are interested in finding bound states² with $E < 0$. We can clean up (35.2), by defining

$$\kappa = \frac{\sqrt{-2mE}}{\hbar}, \quad \rho = \kappa r, \quad \text{and} \quad \rho_0 = \frac{me^2}{2\pi\epsilon_0\hbar^2\kappa}, \quad (35.3)$$

so that (35.2) can be written

$$\frac{d^2u}{d\rho^2} = \left(1 - \frac{\rho_0}{\rho} + \frac{l(l+1)}{\rho^2} \right) u. \quad (35.4)$$

Discussion 35.1

Show that ρ and ρ_0 are dimensionless.

Discussion 35.2

Show how we obtain (35.4) from (35.2) using (35.3).

As with the harmonic oscillator, this equation can be solved using the series solutions method, but to make the series simpler, it is convenient to first extract the asymptotic behaviour.

When ρ is large, from (35.4), we conclude that

$$\frac{d^2u}{d\rho^2} \approx u \quad (35.5)$$

which has the solutions

$$u(\rho) \approx Ae^{-\rho} + Be^\rho, \quad (35.6)$$

This reading is based on [Griffiths 2016, Section 4.2]

¹ Strictly speaking we should consider the motion of the proton as well as the electron. The net effect of doing so, is that we should replace the mass of the electron in (35.2) with the reduced mass of the system. Since the proton is much more massive than the electron, this is not a large effect. For further details see [McIntyre et al. 2012, Section 7.1].

² There are continuum states with $E > 0$ which correspond to (low-energy) electron-proton scattering. The bound states correspond to hydrogen atoms.

but since the second term diverges as $\rho \rightarrow \infty$, we just take $u(\rho) \approx Ae^{-\rho}$.

On the other hand as $\rho \rightarrow 0$, the term in (35.4) proportional to $1/\rho^2$ dominates and

$$\frac{d^2u}{d\rho^2} \approx \frac{l(l+1)}{\rho^2} u \quad (35.7)$$

which has solution

$$u(\rho) = C\rho^{l+1} + D\rho^{-l}. \quad (35.8)$$

Discussion 35.3

Show that (35.8) satisfies (35.7).

The term proportional to ρ^{-l} blows up at $\rho = 0$, so we just take $u \approx C\rho^{l+1}$, for small ρ .

Putting this all together, we can peel off the asymptotic behaviour of u by making the ansatz

$$u(\rho) = \rho^{l+1}e^{-\rho}v(\rho), \quad (35.9)$$

which, after being substituted into (35.4), gives the following equation for v ,

$$\rho \frac{d^2v}{d\rho^2} + 2(l+1-\rho) \frac{dv}{d\rho} + (\rho_0 - 2(l+1))v = 0. \quad (35.10)$$

Discussion 35.4

Show how substituting (35.9) into (35.4) leads to (35.10)

Making the series ansatz,

$$v(\rho) = \sum_{j=0}^{\infty} c_j \rho^j, \quad (35.11)$$

we find

$$\frac{dv}{d\rho} = \sum_{j=1}^{\infty} jc_j \rho^{j-1} = \sum_{j=0}^{\infty} (j+1)c_{j+1} \rho^j, \quad (35.12)$$

$$\frac{d^2v}{d\rho^2} = \sum_{j=0}^{\infty} j(j+1)c_{j+1} \rho^{j-1}. \quad (35.13)$$

Discussion 35.5

Explain the last equality in (35.12).

Then substituting (35.12) and (35.13) into (35.10) gives

$$\sum_{j=0}^{\infty} j(j+1)c_{j+1} \rho^j + 2(l+1) \sum_{j=0}^{\infty} (j+1)c_{j+1} \rho^j - 2 \sum_{j=0}^{\infty} jc_j \rho^j + (\rho_0 - 2(l+1)) \sum_{j=0}^{\infty} c_j \rho^j = 0 \quad (35.14)$$

⊕ Discussion 35.6

Show that substituting (35.12) and (35.13) into (35.10) gives (35.14).

Now equating coefficients of ρ^j in (35.14) leads to

$$j(j+1)c_{j+1} + 2(l+1)(j+1)c_{j+1} - 2jc_j + (\rho_0 - 2(l+1))c_j = 0 , \quad (35.15)$$

which in turn leads to the recursion relation

$$c_{j+1} = \left(\frac{2(j+l+1) - \rho_0}{(j+1)(j+2l+2)} \right) c_j , \quad (35.16)$$

⊕ Discussion 35.7

Show how (35.16) follows from (35.15).

From (35.16) we see that for large j ,

$$c_{j+1} \approx \frac{2}{j+1} c_j , \quad (35.17)$$

so that by induction,

$$c_j \approx A \frac{2^j}{j!} , \quad (35.18)$$

which implies

$$v(\rho) \approx A \sum_{j=0}^{\infty} \frac{(2\rho)^j}{j!} = Ae^{2\rho} , \quad (35.19)$$

and

$$u(\rho) \approx A\rho^{l+1}e^\rho , \quad (35.20)$$

which diverges when ρ is large.

⊕ Discussion 35.8

Explain the last equality in (35.20).

So, just like with the harmonic oscillator, for a normalisable solution the series must terminate at some j_{max} . For this to happen, we must have that $c_{j_{max}+1} = 0$, which from (35.16), implies that

$$2(j_{max} + l + 1) - \rho_0 = 0 . \quad (35.21)$$

If we define

$$n = j_{max} + l + 1 = \rho_0/2 , \quad (35.22)$$

we see that n is a positive integer and (35.16) becomes

$$c_{j+1} = \frac{2(j+l+1-n)}{(j+1)(j+2l+2)} c_j . \quad (35.23)$$

Discussion 35.9

How do we know n must be a positive integer?

From (35.22) and (35.3) we then find

$$E_n = - \left(\frac{m}{2\hbar^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \right) \frac{1}{n^2} = \frac{E_1}{n^2} \quad (35.24)$$

which is the famous Bohr formula. Bohr obtained it in 1913 with a lucky mish-mash of classical physics and pre-quantum mechanics.

Discussion 35.10

Show how (35.22) and (35.3) lead to (35.24).

Angular momentum quantum number

Now, for a given n , we see from (35.22) that l has a maximum value

$$l = n - j_{max} - 1 \quad (35.25)$$

so the possible values of l are

$$l = 0, 1, \dots, n - 1. \quad (35.26)$$

We also saw that m runs from $-l$ to l in integer steps. A summary is shown in (35.27).

Hydrogen atom quantum numbers

$n = 1, 2, 3, \dots, \infty$
$l = 0, 1, 2, \dots, n - 1$
$m = -l, -l + 1, \dots, 0, \dots, l - 1, l$

(35.27)

It is conventional to label states using letters³ based on the value of l as shown in (35.28)

$l = 0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9 \quad \dots$
$letter = s \quad p \quad d \quad f \quad g \quad h \quad i \quad k \quad l \quad m \quad \dots$

(35.28)

We often denote states using the value of n and a letter to indicate l . So for example we would write the $n = 2, l = 0$ states as 2s.

³ If you are curious, these labels apparently come from the characteristics of spectral lines in certain alkali metals (sharp, principal, diffuse and fundamental) – see [Jensen 2007].

Wave functions

It is useful to define the Bohr radius

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{me^2} = 0.529 \times 10^{-10} \text{ m}, \quad (35.29)$$

which allows us to write

$$\kappa = \frac{1}{a_0 n}. \quad (35.30)$$

and

$$\rho = \frac{r}{a_0 n} , \quad (35.31)$$

The Bohr radius, a_0 , is the relevant length scale for the hydrogen atom.

We now have all the ingredients to write down the spacial wave function for hydrogen. The wave functions are labelled by three quantum numbers

$$\psi_{nlm}(r, \theta, \phi) = R_{nl} Y_l^m(\theta, \phi) , \quad (35.32)$$

with

$$R_{nl} = \frac{1}{r} \rho^{l+1} e^{-\rho} v(\rho) , \quad (35.33)$$

where $v(\rho)$ is a polynomial in ρ of degree, $j_{max} = n - l - 1$, whose coefficients are determined by (35.16).

Let's first look at the ground state. With $n = 1$,

$$E_1 = - \left(\frac{m}{2\hbar^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \right) = -13.6 \text{ eV} . \quad (35.34)$$

This tell us to ionise the ground state of hydrogen, that is, to convert a bound state to a scattering state, since $V(\infty) = 0$, we need 13.6 eV.

Now, given that $n = 1 = j_{max} + l + 1$, with $j_{max} \geq 0$, we conclude that $l = j_{max} = 0$ which in turn implies $m = 0$. So the ground state of hydrogen is

$$\psi_{100}(r, \theta, \phi) = R_{10}(r) Y_0^0(\theta, \phi) . \quad (35.35)$$

Then, with $l = 0$, $n = 1$, the recursion formula (35.23) tells us that $c_1 = 0$, so,

$$R_{10} = \frac{1}{r} \rho^1 e^{-\rho} c_0 = \frac{c_0}{a_0} e^{-r/a_0} . \quad (35.36)$$

Discussion 35.11

Check that with $l = 0$ and $n = 1$, (35.23) tells us that $c_1 = 0$.

To fix c_0 , we need to normalise $R(r)$,

$$\int_0^\infty |R_{10}|^2 r^2 dr = \frac{|c_0|^2}{a_0^2} \int_0^\infty e^{-2r/a_0} r^2 dr = |c_0|^2 \frac{a_0}{4} , \quad (35.37)$$

which implies that $c_0 = 2/\sqrt{a_0}$.

Discussion 35.12

Reproduce the last step in (35.37). You may use [McIntyre et al. 2012, Appendix F].

Finally using, $Y_0^0 = 1/\sqrt{4\pi}$, we get

$$\psi_{100}(r, \theta, \phi) = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0} . \quad (35.38)$$

The first excited state of hydrogen corresponds to $n = 2$, so that

$$E_2 \stackrel{(35.24)}{=} -\frac{13.6 \text{ eV}}{2^2} = -3.4 \text{ eV}. \quad (35.39)$$

With $n = 2$, we can have, $l = 0$ and $m = 0$, or $l = 1$ and $m = -1, 0, +1$.

We see that there are four states⁴ with E_2 .

Now with $n = 2$, $l = 0$, using (35.23) gives

$$\begin{aligned} c_1 &= \frac{2(1-2)}{(1)(2)}c_0 = -c_0, \\ c_2 &= \frac{2(1+0+1-2)}{(2)(3)}(-c_0) = 0, \end{aligned} \quad (35.40)$$

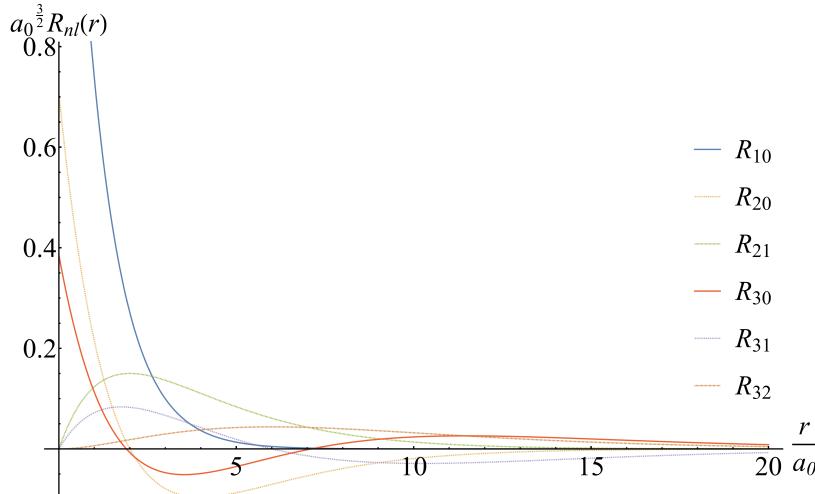
so $v(\rho) = c_0(1-\rho)$ and

$$R_{20}(r) = \frac{c_0}{2a_0} \left(1 - \frac{r}{2a_0}\right) e^{-r/2a_0}. \quad (35.41)$$

With $n = 2$ and $l = 1$, one finds that $c_1 = 0$, so

$$R_{21}(r) = \frac{c_0}{4a^2} r e^{-r/2a}. \quad (35.42)$$

To determine c_0 in (35.41) and (35.42) one needs to use the



⁴ The four energy eigenstates states in position space are $\psi_{200} = R_{20}Y_0^0$, $\psi_{21,-1} = R_{21}Y_1^{-1}$, $\psi_{210} = R_{21}Y_1^0$ and $\psi_{211} = R_{21}Y_1^1$. In ket notation we would write $|200\rangle$, $|21, -1\rangle$, $|210\rangle$ and $|211\rangle$.

normalisation condition. For a plot of the first few radial wavefunctions see fig. 35.1. The first few radial wave functions are tabulated in table 35.1.

Now it turns out that the $v(\rho)$'s are well known functions, which, apart from their normalisation, can be written,

$$v(\rho) = L_{n-l-1}^{2l+1}(2\rho), \quad (35.43)$$

where

$$L_{q-p}^p(x) := (-1)^p \left(\frac{d}{dx}\right)^p L_q(x), \quad (35.44)$$

is the associated Laguerre polynomial and

$$L_q(x) = e^x \left(\frac{d}{dx}\right)^q (e^{-x} x^q), \quad (35.45)$$

$R_{10}(r) = 2\left(\frac{1}{a_0}\right)^{3/2} e^{-r/a_0}$
$R_{20}(r) = 2\left(\frac{1}{2a_0}\right)^{3/2} \left(1 - \frac{r}{2a_0}\right) e^{-r/2a_0}$
$R_{21}(r) = \frac{1}{\sqrt{3}} \left(\frac{1}{2a_0}\right)^{3/2} \frac{r}{a_0} e^{-r/2a_0}$
$R_{31}(r) = \frac{4\sqrt{2}}{9} \left(\frac{1}{3a_0}\right)^{3/2} \frac{r}{a_0} \left(1 - \frac{r}{6a_0}\right) e^{-r/3a_0}$
$R_{32}(r) = \frac{2\sqrt{2}}{27\sqrt{5}} \left(\frac{1}{3a_0}\right)^{3/2} \left(\frac{r}{a_0}\right)^2 e^{-r/3a_0}$

Table 35.1: Selection of hydrogen radial wave functions

is the q th Laguerre polynomial.

We can now write down the form of the normalised wave function, ψ_{lmn} ,

$$\psi_{nlm}(r, \theta, \phi) = \sqrt{\left(\frac{2}{na_0}\right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3}} e^{-r/na_0} \left(\frac{2r}{na_0}\right)^l L_{n-l-1}^{2l+1}(2r/na_0) Y_l^m(\theta, \phi). \quad (35.46)$$

The hydrogen atom is one of the few realistic quantum mechanical systems which can be solved exactly. Notice that the wavefunction depends on, n, l, m , but the energy only depends on n . This feature is a peculiarity of the Coulomb potential and is not a general feature of spherically symmetric potentials.

The wave functions ψ_{nlm} are orthogonal:

$$\int \psi_{nlm}^* \psi_{n'l'm'} r^2 \sin \theta dr d\theta d\phi = \delta_{nn'} \delta_{ll'} \delta_{mm'}. \quad (35.47)$$

Hydrogen spectrum

The principle quantum number, n , is sometimes called the shell number. We see from (35.24) that E does not depend on l and m . A plot of the hydrogen atom spectrum is shown in Figure 35.2. Unlike the finite square well, hydrogen has an infinite number of bound states.

For each E_n , there are l states $l = 0, \dots, n-1$ and for each l , there are in turn, $2l+1$ states, $m = -l$ to $m = +l$. This means that the degeneracy of E_n is

$$\sum_{l=0}^{n-1} (2l+1) = \sum_{l=0}^{n-1} 2l + \sum_{l=0}^{n-1} 1 = 2 \frac{n(n-1)}{2} + n = n^2. \quad (35.48)$$

In addition, an electron can either be spin-up or spin-down so the degeneracy of E_n is $2n^2$.

Hydrogen atoms absorb or emit photons when electrons make transitions between energy levels. Transition from a higher to a lower state leads to photon emission, while transition from lower to higher is a result of photon absorption. When a photon is absorbed or emitted, its energy matches the difference in energy levels

$$E_{\text{photon}} = \Delta E_{fi} = |E_f - E_i| = 13.6 \text{ eV} \left| \frac{1}{n_i^2} - \frac{1}{n_f^2} \right| \quad (35.49)$$

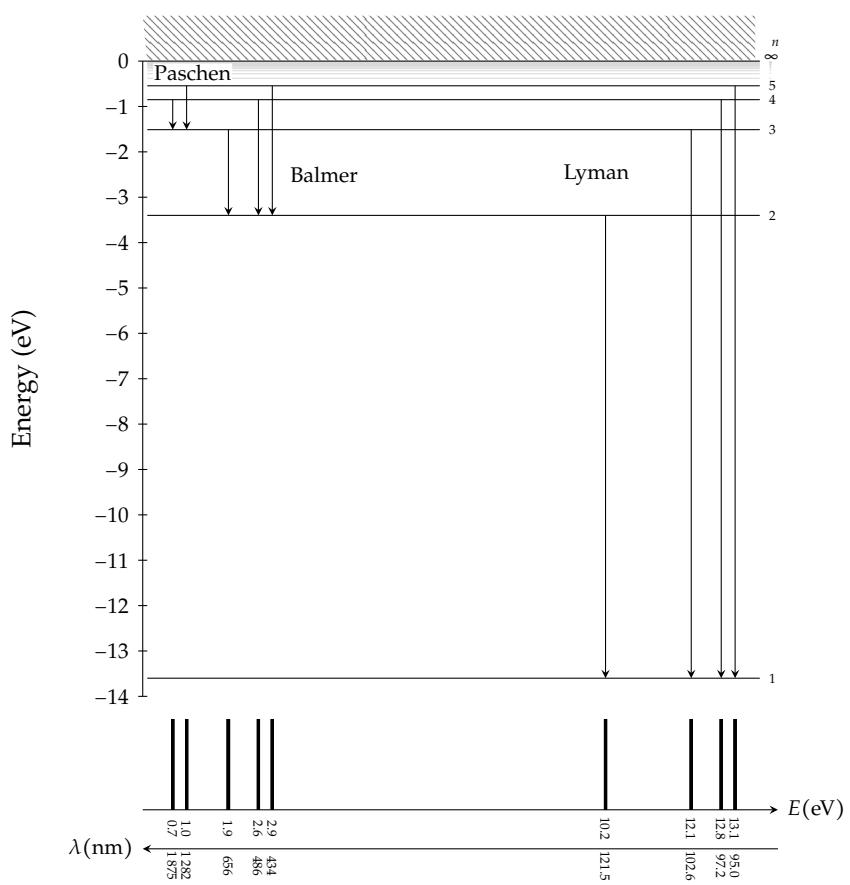


Figure 35.2: Hydrogen atom spectrum

Enrichment

Transitions to $n = 1$ are called the **Lyman series**. Transitions to $n = 2$ are called the **Balmer series** and transitions to $n = 3$ are called the **Paschen series**.

Discussion 35.13

Write a summary of the important points and formulae in this chapter.

36

Hydrogen atom II

We saw that the hydrogen energy eigenstates can be written

This reading is based on [McIntyre et al. 2012, Section 8.4]

$$|nlm\rangle = \psi_{nlm}(r, \theta, \phi) = R_{nl}(r)Y_l^m(\theta, \phi), \quad (36.1)$$

with ψ_{nlm} given by (35.46). The first few wave functions are

$$\begin{aligned} \psi_{100} &= \frac{1}{\sqrt{\pi}} \left(\frac{1}{a_0} \right)^{\frac{3}{2}} e^{-r/a_0} \\ \psi_{200} &= \frac{1}{\sqrt{\pi}} \left(\frac{1}{2a_0} \right)^{\frac{3}{2}} \left(1 - \frac{r}{2a_0} \right) e^{-r/2a_0} \\ \psi_{210} &= \frac{1}{2\sqrt{\pi}} \left(\frac{1}{2a_0} \right)^{\frac{3}{2}} \left(\frac{r}{a_0} \right) e^{-r/2a_0} \cos \theta \\ \psi_{21\pm 1} &= \mp \frac{1}{2\sqrt{2\pi}} \left(\frac{1}{2a_0} \right)^{\frac{3}{2}} \left(\frac{r}{a_0} \right) e^{-r/2a_0} \sin \theta e^{\pm i\phi} \\ \psi_{300} &= \frac{1}{\sqrt{\pi}} \left(\frac{1}{3a_0} \right)^{\frac{3}{2}} \left(1 - \left(\frac{2r}{3a_0} \right) + \frac{2}{27} \left(\frac{r}{a_0} \right)^2 \right) e^{-r/3a_0} \\ \psi_{310} &= \frac{2\sqrt{2}}{3\sqrt{3\pi}} \left(\frac{1}{3a_0} \right)^{\frac{3}{2}} \left(\frac{r}{a_0} \right) \left(1 - \frac{r}{6a_0} \right) e^{-r/3a_0} \cos \theta \\ \psi_{31,\pm 1} &= \mp \frac{2}{3\sqrt{3\pi}} \left(\frac{1}{3a_0} \right)^{\frac{3}{2}} \left(\frac{r}{a_0} \right) \left(1 - \frac{r}{6a_0} \right) e^{-r/3a_0} \sin \theta e^{\pm i\phi} \\ \psi_{320} &= \frac{1}{27\sqrt{2\pi}} \left(\frac{1}{3a_0} \right)^{\frac{3}{2}} \left(\frac{r}{a_0} \right)^2 e^{-r/3a_0} (3 \cos^2 \theta - 1) \\ \psi_{32,\pm 1} &= \mp \frac{\sqrt{3}}{27\sqrt{\pi}} \left(\frac{1}{3a_0} \right)^{\frac{3}{2}} \left(\frac{r}{a_0} \right)^2 e^{-r/3a_0} \cos \theta \sin \theta e^{\pm i\phi} \\ \psi_{32,\pm 2} &= \frac{\sqrt{3}}{54\sqrt{\pi}} \left(\frac{1}{3a_0} \right)^{\frac{3}{2}} \left(\frac{r}{a_0} \right)^2 e^{-r/3a_0} \sin^2 \theta e^{\pm 2i\phi} \end{aligned}$$

Hydrogen Energy Eigenstates

The energy eigenfunctions of \hat{H} are also eigenfunctions of \hat{L}^2 and \hat{L}_z :

$$\hat{H}\psi_{nlm}(r, \theta, \phi) = -\frac{13.6 \text{ eV}}{n^2} \psi_{nlm}(r, \theta, \phi) \quad (36.2)$$

$$\hat{L}^2\psi_{nlm}(r, \theta, \phi) = l(l+1)\hbar^2\psi_{nlm}(r, \theta, \phi) \quad (36.3)$$

$$\hat{L}_z\psi_{nlm}(r, \theta, \phi) = m\hbar\psi_{nlm}(r, \theta, \phi) \quad (36.4)$$

Probability density

The probability density is the modulus squared of the wave function

$$\begin{aligned}\mathcal{P}(r, \theta, \phi) &= |\psi_{nlm}(r, \theta, \phi)|^2 \\ &= |R_{nl}(r)Y_l^m(\theta, \phi)|^2\end{aligned}$$

Multiplying the probability density by dV gives the probability of finding a particle in that volume:

$$\begin{aligned}\mathcal{P}(r, \theta, \phi) dV &= |\psi_{nlm}(r, \theta, \phi)|^2 r^2 dr \sin \theta d\theta d\phi \\ &= (|R_{nl}(r)|^2 r^2 dr)(|Y_l^m(\theta, \phi)|^2 \sin \theta d\theta d\phi)\end{aligned}\quad (36.5)$$

One needs to integrate (36.5) to find \mathcal{P} for a finite region.

Plotting Probability density

We will look at different ways to represent a three dimensional probability density graphically. The simplest case is the ground state $|100\rangle$:

$$\psi_{100}(r, \theta, \phi) = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0} \quad (36.6)$$

with the probability density

$$\mathcal{P}_{100}(r, \theta, \phi) = \frac{1}{\pi a_0^3} e^{-2r/a_0} \quad (36.7)$$

Note that $[\mathcal{P}] = 1/L^3$, and in this case, \mathcal{P} is independent of θ and ϕ . In Figure 36.1, we suppress the z coordinate, and plot \mathcal{P} as a height above the xy -plane.

One can also plot \mathcal{P} only as a function of r as in Figure 36.2. This will obviously give a very incomplete picture for wave functions that have angular dependence. Can you see the relationship between the two plots?

We can also plot probability densities using density plots where the value of the probability density is represented by a colour or shade of grey. The probability density in the $x - y$ plane is shown in Figure 36.3. The white area corresponds to the most probable location of electron.

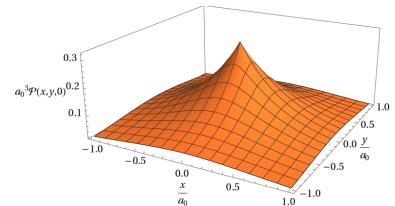


Figure 36.1: Plot of \mathcal{P}_{100} as a function of x and y .

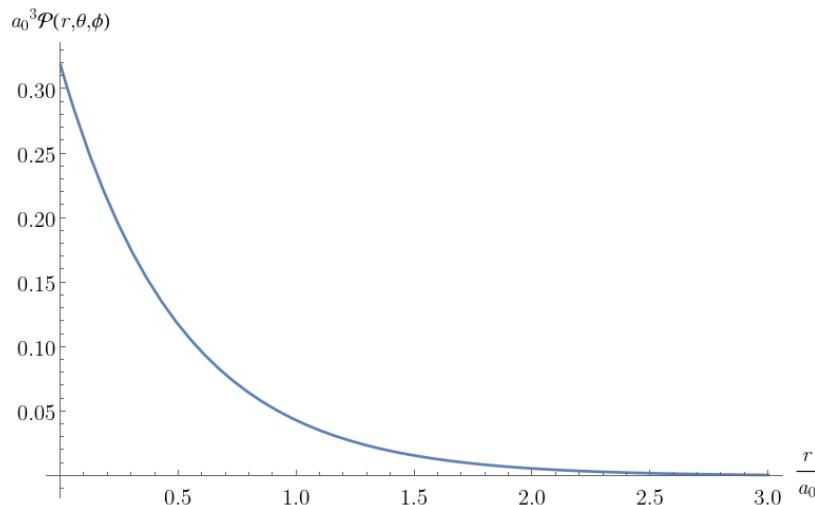


Figure 36.2: Plot of \mathcal{P}_{100} as a function of r .

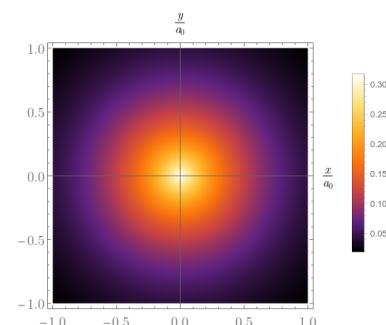


Figure 36.3: Plot of $|\psi_{100}|^2$ on the xy -plane.

To try to capture more three dimensional information about the probability density in three dimensions one can plot the density on various surfaces. The Figure 36.4 (a) shows the probability density on the planes $\frac{x}{a_0} = 0, \frac{1}{2}$ and $-\frac{1}{2}$, while The Figure 36.4 (b) shows the density on the planes $x = 0, y = 0$ and $z = 0$.

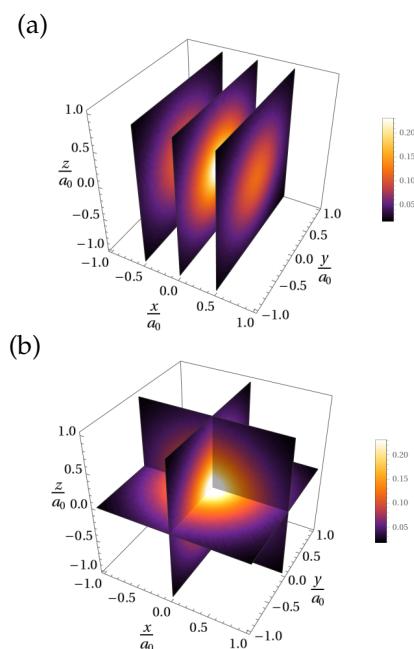


Figure 36.4: Plot of $|\psi_{100}|^2$ on (a) a series of parallel planes and (b) a series of intersecting planes.

Probability plots for more hydrogen wave functions (labelled by the (n, l, m) quantum) are shown in Figure 36.5.

Example 36.1: Finding an electron

Problem Find the probability that the electron in the ground state of hydrogen is measured to be within one Bohr radius of the nucleus and calculate the expectation value of the radial position r .

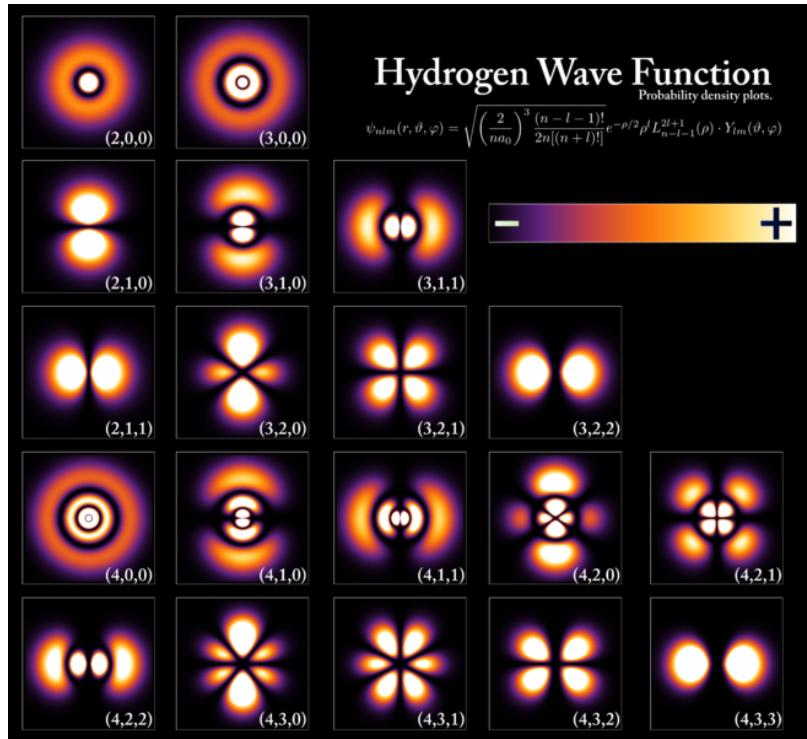


Figure 36.5: Atomic orbitals of the electron in a hydrogen atom at different energy levels. The probability of finding the electron is given by the color, as shown in the key at upper right.

image credit: PoorLeno
Image source

Solution To find the probability we need to integrate the probability density over the region of interest.

$$\begin{aligned} \mathcal{P}_{r < a_0} &= \int_0^{a_0} \int_0^\pi \int_0^{2\pi} |\psi_{100}|^2(r, \theta, \phi) dV \\ &= \int_0^{a_0} |R_{10}|^2 r^2 dr \int_0^\pi \int_0^{2\pi} |Y_1^0|^2 \sin \theta d\theta d\phi \end{aligned}$$

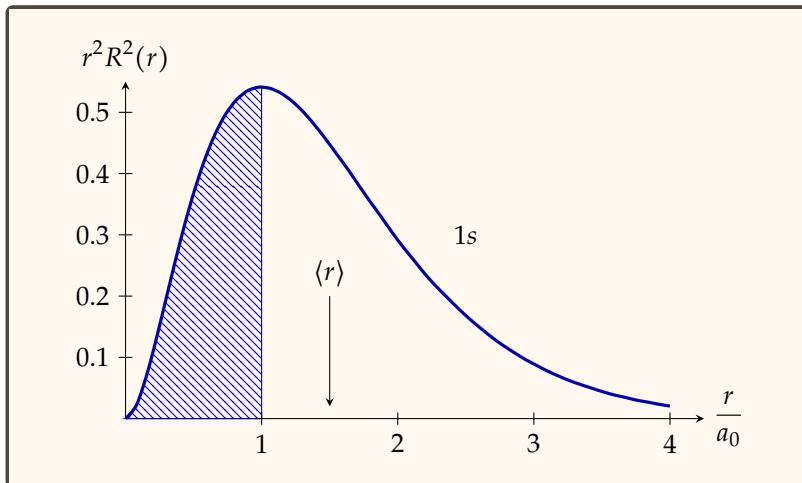
Then using the fact that the spherical harmonics, Y_l^m are normalised, we get

$$\mathcal{P}_{r < a_0} = \int_0^{a_0} |R_{10}|^2 r^2 dr = \frac{1}{a_0^3} \int_0^{a_0} 4e^{-2r/a_0} r^2 dr \quad (36.8)$$

Letting, $u = 2r/a_0$, so that $r = \frac{1}{2}a_0 u$ and

$$\mathcal{P}_{r < a_0} = \frac{1}{2} \int_0^2 e^{-u} u^2 du = \frac{1}{2} [(-u^2 - 2u - 2)e^{-u}]_0^2 = 1 - 5e^{-2} \approx 0.323 \quad (36.9)$$

which tells us there is about a 32% chance of finding the electron within a distance of a_0 from the atom. The radial probability integrand for the hydrogen 1s ground state is shown below with the shaded area corresponding to $\mathcal{P}_{r < a_0}$.



Superposition states

The most general bound state is a superposition of the bound energy eigenstates

$$|\psi(t)\rangle \doteq \psi(r, \theta, \phi, t) = \sum_{n,l,m} c_{nlm} R_{nl}(r) Y_l^m(\theta, \phi) e^{-iE_n t/\hbar} \quad (36.10)$$

and we can get the coefficients c_{nlm} in the usual way

$$c_{nlm} = \langle nlm | \psi(0) \rangle = \int_0^\infty \int_0^\pi \int_0^{2\pi} R_{nl}^*(r) Y_l^{m*}(\theta, \phi) \psi(r, \theta, \phi, 0) r^2 \sin \theta dr d\theta d\phi \quad (36.11)$$

Example 36.2: A hydrogen superposition state

Problem Find the time evolution of an equal superposition of the 1s ground state and the $2p_0$ ($m = 0$) excited state:

$$|\psi(0)\rangle = \frac{1}{\sqrt{2}}|100\rangle + \frac{1}{\sqrt{2}}|210\rangle. \quad (36.12)$$

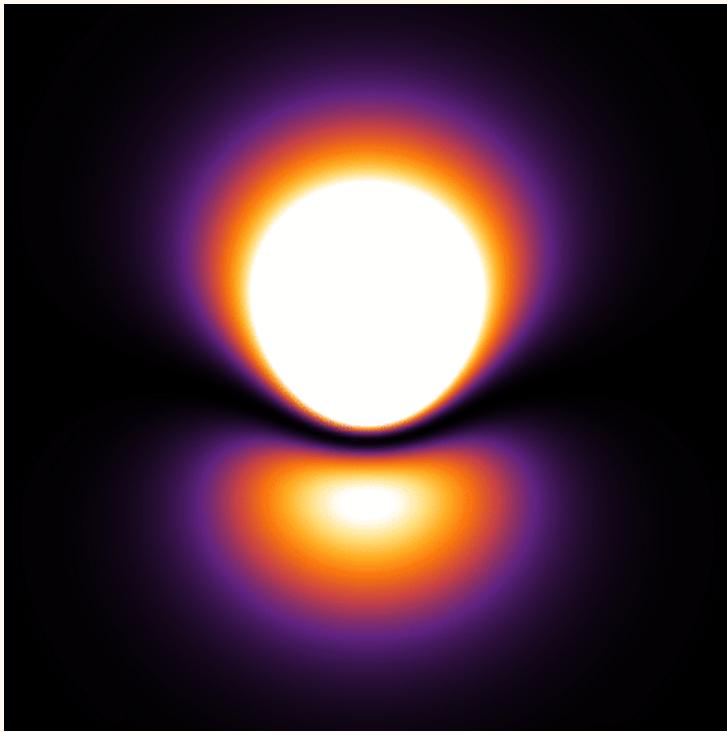
Solution We put in the time-dependence in the usual way

$$\begin{aligned} \psi(r, \theta, \phi, t) &= \frac{1}{\sqrt{2}}\psi_{100}e^{-iE_1 t/\hbar} + \frac{1}{\sqrt{2}}\psi_{210}e^{-iE_2 t/\hbar} \\ &= \frac{1}{\sqrt{2\pi a_0^3}}e^{-r/a_0}e^{-iE_1 t/\hbar} + \frac{1}{\sqrt{\pi a_0^3}}\left(\frac{r \cos \theta}{8a_0}\right)e^{-r/2a_0}e^{-iE_2 t/\hbar} \\ &= \frac{1}{\sqrt{2\pi a_0^3}}e^{-iE_1 t/\hbar}\left(e^{-r/a_0} + \frac{r \cos \theta}{4\sqrt{2}a_0}e^{-r/2a_0}e^{-i\omega_{12}t}\right), \end{aligned}$$

where ω_{12} is the Bohr frequency $\omega_{12} = (E_2 - E_1)/\hbar$. Using the fact that $z = r \cos \theta$ we can write the wave function as

$$\psi(r, \theta, \phi, t) = \frac{1}{\sqrt{2\pi a_0^3}}e^{-iE_1 t/\hbar}\left(e^{-r/a_0} + \frac{z}{4\sqrt{2}a_0}e^{-r/2a_0}e^{-i\omega_{12}t}\right) \quad (36.13)$$

You can click on the image below to see an animation of the probability density for the $1s-2p$ superposition.



We can see the probability density moves up and down along the z -axis. This is a model of an oscillating electric dipole moment.

Example 36.3: Another hydrogen superposition state

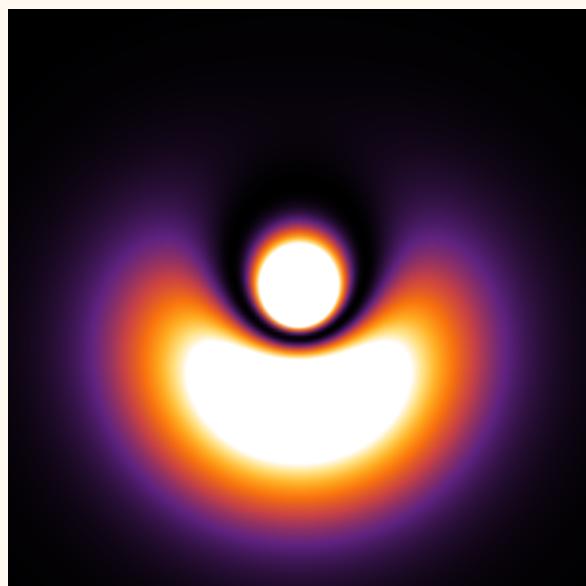
Problem Find the time evolution of an equal superposition of the $2s$ excited state and the $2p_0(m = 0)$ excited state:

$$|\psi(0)\rangle = \frac{1}{\sqrt{2}}|200\rangle + \frac{1}{\sqrt{2}}|210\rangle . \quad (36.14)$$

Solution We put in the time-dependence in the usual way

$$\begin{aligned} \psi(r, \theta, \phi, t) &= \frac{1}{\sqrt{2}}\psi_{200}e^{-iE_2t/\hbar} + \frac{1}{\sqrt{2}}\psi_{210}e^{-iE_2t/\hbar} \\ &= \frac{1}{4\sqrt{\pi a_0^3}}\left(1 - \frac{r}{2a_0}\right)e^{-r/2a_0}e^{-iE_2t/\hbar} + \frac{1}{\sqrt{\pi a_0^3}}\left(\frac{r \cos \theta}{8a_0}\right)e^{-r/2a_0}e^{-iE_2t/\hbar} \\ &= \frac{1}{4\sqrt{\pi a_0^3}}e^{-iE_2t/\hbar}e^{-r/2a_0}\left(1 - \frac{r}{2a_0} + \frac{z}{2a_0}\right), \end{aligned}$$

In this example, the two states have the same energy so that there is not a time-dependant relative phase. This in turn means that the probability amplitude, shown below is time independent. The electron cloud is displaced in the negative z direction and this is a model of a *static* electric dipole. This kind of state can be used to explain molecular bonding. Chemists often refer to superpositions of different orbitals as **hybrids** – so this state would be called an *s-p* hybridization.



❖ **Discussion 36.1**

Write a summary of the important points and formulae in this chapter.

A

Review of linear algebra I

Linear algebra generalises the arithmetic of ordinary vectors allowing scalars to be complex and going beyond 3 dimensions. This chapter serves as both a review and introduction to notation we will be using for quantum mechanics.

A.1 Vector spaces

A vector space consists of a set of vectors, $\{|\alpha\rangle, |\beta\rangle, |\gamma\rangle, \dots\}$ and a set of scalars $\{a, b, c, \dots\}$. Vector spaces are closed under two operations

1. vector addition

If $|\alpha\rangle$ and $|\beta\rangle$ are vectors then $|\alpha\rangle + |\beta\rangle$ is a vector too. To express this fact we write

$$|\alpha\rangle + |\beta\rangle = |\alpha + \beta\rangle . \quad (\text{A.1})$$

Vector addition has the following properties

- (a) commutativity : $|\alpha\rangle + |\beta\rangle = |\beta\rangle + |\alpha\rangle$
- (b) associativity : $|\alpha\rangle + (|\beta\rangle + |\gamma\rangle) = (|\alpha\rangle + |\beta\rangle) + |\gamma\rangle$
- (c) \exists a zero vector $|0\rangle$: $\forall |\alpha\rangle; |0\rangle + |\alpha\rangle = |\alpha\rangle$
- (d) every vector has an additive inverse: $\forall |\alpha\rangle, \exists$ an $|- \alpha\rangle$ such that $|\alpha\rangle + |- \alpha\rangle = |0\rangle$.

2. scalar multiplication

If a is a scalar and $|\alpha\rangle$ is a vectors then $a|\alpha\rangle$ is a vector too. To express this fact we write

$$a|\alpha\rangle = |a\alpha\rangle . \quad (\text{A.2})$$

scalar multiplication has the following properties

- (a) distributive over vector addition : $a(|\alpha\rangle + |\beta\rangle) = a|\alpha\rangle + a|\beta\rangle$
- (b) distributive over scalar addition $(a + b)|\alpha\rangle = a|\alpha\rangle + b|\alpha\rangle$
- (c) associative: $a(b|\alpha\rangle) = (ab)|\alpha\rangle$
- (d) $0|\alpha\rangle = |0\rangle$
- (e) $1|\alpha\rangle = |\alpha\rangle$
- (f) $|- \alpha\rangle = -1|\alpha\rangle = -|\alpha\rangle$

A **linear combination** of vectors $|\alpha\rangle, |\beta\rangle, |\gamma\rangle \dots$ is an expression of the form

$$a|\alpha\rangle + b|\beta\rangle + c|\gamma\rangle + \dots , \quad (\text{A.3})$$

where a, b, c, \dots are scalars.

We say that a vector $|\lambda\rangle$ is **linearly independent** of vectors, $|\alpha\rangle, |\beta\rangle, |\gamma\rangle, \dots$, if it can not be written in the form (A.3). For example, \hat{k} is linearly independent of \hat{i} and \hat{j} but any vectors in the xy -plane are linearly dependent on \hat{i} and \hat{j} .

We say a set of vectors is linearly independent if they are all linearly independent of the rest. For example the set $\{\hat{i}, \hat{j}, \hat{k}\}$ is linearly independent.

We say a set of vectors **spans** the space if every vector in the space can be written as a linear combination of elements of the set. For example the set $\{\hat{i}, \hat{i} + \hat{j}, \hat{i} + \hat{k}, \hat{j} + \hat{k}\}$ spans 3D vectors.

A linearly independent set that spans the space is called a **basis**. For example the set $\{\hat{i} + \hat{j}, \hat{i} - \hat{j}, \hat{k}\}$ is a basis for 3D vectors.

The number of vectors in a basis is called the dimension of the space, which we assume to be finite for now.

For a given basis, $\mathcal{O} = \{|e_1\rangle, |e_2\rangle, \dots, |e_n\rangle\}$, we can write any vector as a linear combination of the basis vectors

$$|\alpha\rangle = a_1|e_1\rangle + a_2|e_2\rangle + \dots + a_n|e_n\rangle = \sum_{i=1}^n a_i|e_i\rangle . \quad (\text{A.4})$$

The coefficients a_i are called the **components** of the vector (with respect to the basis \mathcal{O}). For a given basis, a vector is uniquely represented by its components and we write

$$|\alpha\rangle \xrightarrow{\mathcal{O}} (a_1, a_2, \dots, a_n) . \quad (\text{A.5})$$

The notation, $\xrightarrow{\mathcal{O}}$, is short hand for “components with respect to basis \mathcal{O} ”. As you should be aware, it is often easier to work with components:

$$\begin{aligned} |\alpha\rangle + |\beta\rangle &\xrightarrow{\mathcal{O}} (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n) , \\ c|\alpha\rangle &\xrightarrow{\mathcal{O}} (ca_1, ca_2, \dots, ca_n) , \\ |-\alpha\rangle &\xrightarrow{\mathcal{O}} (-a_1, -a_2, \dots, -a_n) . \end{aligned} \quad (\text{A.6})$$

A.2 Inner product

The inner product is a generalisation of the dot product¹.

The inner product of two vectors $|\alpha\rangle$ and $|\beta\rangle$, is complex and we write it as $\langle\alpha|\beta\rangle$.

The inner product satisfies four important properties

1. $\langle\alpha|\beta\rangle = \langle\beta|\alpha\rangle^*$
2. $\langle\alpha|\alpha\rangle \geq 0$
3. $\langle\alpha|\alpha\rangle = 0 \Leftrightarrow |\alpha\rangle = |0\rangle$
4. $\langle\alpha|(b|\beta\rangle + c|\gamma\rangle) = b\langle\alpha|\beta\rangle + c\langle\alpha|\gamma\rangle$

¹ The cross product does not generalise in any natural way when $d \neq 3$ – one needs to go beyond vectors in other dimension to generalise the cross product.

Properties 1 and 4 of the inner product imply that

$$\langle a\alpha + b\beta|\gamma \rangle = a^* \langle \alpha|\gamma \rangle + b^* \langle \beta|\gamma \rangle . \quad (\text{A.7})$$

A vector space with an inner product is called an **inner product space**.

Definition 1.2.1: The **norm** of a vector is defined as $\|\alpha\| = \sqrt{\langle \alpha|\alpha \rangle}$.

Definition 1.2.2: A **normalised vector**, $|\alpha\rangle$, satisfies $\|\alpha\| = 1$.

Definition 1.2.3: We say two vectors are **orthogonal** if $\langle \alpha|\beta \rangle = 0$.

Definition 1.2.4: A set of mutually orthogonal normalised vectors – ie. which satisfy $\langle \alpha_i|\alpha_j \rangle = \delta_{ij}$ – is called an **orthonormal set**.

It is very convenient to choose an orthonormal basis. In an orthonormal basis, $\{|e_1\rangle, |e_2\rangle, \dots, |e_n\rangle\}$, it is easy to write the inner product in terms of components:

$$\begin{aligned} \langle \alpha|\beta \rangle &= a_1^* b_1 + a_2^* b_2 + \dots + a_n^* b_n , \\ \langle \alpha|\alpha \rangle &= |a_1|^2 + |a_2|^2 + \dots + |a_n|^2 , \\ a_i &= \langle e_i|\alpha \rangle , \end{aligned} \quad (\text{A.8})$$

which reduce to familiar formula for real 3D vectors

$$\begin{aligned} \vec{a} \cdot \vec{b} &= a_x b_x + a_y b_y + a_z b_z , \\ \vec{a} \cdot \vec{a} &= (a_x)^2 + (a_y)^2 + (a_z)^2 , \\ a_i &= \hat{e}_i \cdot \vec{a} , \end{aligned} \quad (\text{A.9})$$

Using the inner product, we can generalise the concept of an angle between vectors. In 3D, the angle between two vectors is given by

$$\cos \theta = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}| |\vec{b}|} . \quad (\text{A.10})$$

In a general inner product space we can use the Schwarz inequality,

$$|\langle \alpha|\beta \rangle|^2 \leq \langle \alpha|\alpha \rangle \langle \beta|\beta \rangle , \quad (\text{A.11})$$

which implies that,

$$0 \leq \frac{|\langle \alpha|\beta \rangle|^2}{\langle \alpha|\alpha \rangle \langle \beta|\beta \rangle} \leq 1 , \quad (\text{A.12})$$

so that we can define the angle between two vectors to be

$$\cos \theta = \sqrt{\frac{\langle \alpha|\beta \rangle \langle \beta|\alpha \rangle}{\langle \alpha|\alpha \rangle \langle \beta|\beta \rangle}} . \quad (\text{A.13})$$

A.3 Matrices

An **operator** takes a vector and “transforms” it into another vector

$$|\alpha\rangle \rightarrow \hat{T}|\alpha\rangle . \quad (\text{A.14})$$

If an operator satisfies the property

$$\hat{T}(a|\alpha\rangle + b|\beta\rangle) = a\hat{T}|\alpha\rangle + b\hat{T}|\beta\rangle , \quad (\text{A.15})$$

we call it a **linear operator** or **linear transformation**. Rotation of a vector is a linear transformation.

Linearity is a very useful property, since if we know how an operator acts on our basis, we know how it acts on any vectors.

Suppose

$$\begin{aligned}\hat{T}|e_1\rangle &= T_{11}|e_1\rangle + T_{21}|e_2\rangle + \dots + T_{n1}|e_n\rangle , \\ \hat{T}|e_2\rangle &= T_{12}|e_1\rangle + T_{22}|e_2\rangle + \dots + T_{n2}|e_n\rangle , \\ &\vdots && \vdots \\ \hat{T}|e_n\rangle &= T_{1n}|e_1\rangle + T_{2n}|e_2\rangle + \dots + T_{nn}|e_n\rangle ,\end{aligned}\quad (\text{A.16})$$

which can be compactly written

$$\hat{T}|e_j\rangle = \sum_{i=1}^N T_{ij}|e_i\rangle . \quad (\text{A.17})$$

Now given some vector, $|\alpha\rangle$,

$$\begin{aligned}|\alpha\rangle &= a_1|e_1\rangle + a_2|e_2\rangle + \dots + a_n|e_n\rangle \\ &= \sum_{k=1}^N a_k|e_k\rangle ,\end{aligned}\quad (\text{A.18})$$

so

$$\begin{aligned}\hat{T}|\alpha\rangle &\stackrel{(\text{A.15})(\text{A.18})}{=} \sum_{k=1}^N a_k(\hat{T}|e_k\rangle) \\ &\stackrel{(\text{A.17})}{=} \sum_{k=1}^N \sum_{i=1}^N a_k T_{ik}|e_i\rangle \\ &= \sum_{i=1}^N \left(\sum_{k=1}^N T_{ik} a_k \right) |e_i\rangle \\ &= |\alpha'\rangle = \sum_{i=1}^N a'_i|e_i\rangle ,\end{aligned}\quad (\text{A.19})$$

so that evidently

$$a'_i = \sum_{k=1}^N T_{ik} a_k . \quad (\text{A.20})$$

Now (A.20) should look familiar – it is the rule for multiplying a column by a matrix:

$$\vec{a}' = \mathbf{T}\vec{a} \Leftrightarrow \begin{pmatrix} a'_1 \\ a'_2 \\ \vdots \\ a'_n \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} & \cdots & T_{1n} \\ T_{21} & T_{22} & \cdots & T_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ T_{n1} & T_{n2} & \cdots & T_{nn} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} . \quad (\text{A.21})$$

So, with respect to a given basis, \mathcal{O} , we can write a linear operator, \hat{T} , as a matrix

$$\hat{T} \xrightarrow{\mathcal{O}} \begin{pmatrix} T_{11} & T_{12} & \cdots & T_{1n} \\ T_{21} & T_{22} & \cdots & T_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ T_{n1} & T_{n2} & \cdots & T_{nn} \end{pmatrix} \quad (\text{A.22})$$

In particular, in an orthonormal basis we can find the elements of the matrix representation of a linear operator as follows:

$$T_{ij} = \langle e_i | \hat{T} | e_j \rangle . \quad (\text{A.23})$$

We see that the study of linear transformations in a finite dimensional vector space, reduces to the study of matrices. For example since

$$(\hat{S} + \hat{T}) |\alpha\rangle = \hat{S} |\alpha\rangle + \hat{T} |\alpha\rangle , \quad (\text{A.24})$$

in terms of components

$$\hat{U} = \hat{S} + \hat{T} \xrightarrow{\mathcal{O}} U_{ij} = S_{ij} + T_{ij} . \quad (\text{A.25})$$

What about matrix multiplication? What are the components of $\hat{U} = \hat{S}\hat{T}$?

Take $|\alpha'\rangle = \hat{T}|\alpha\rangle$ and $|\alpha''\rangle = \hat{S}|\alpha'\rangle = \hat{S}\hat{T}|\alpha\rangle$, so

$$a''_i = \sum_{j=1}^N S_{ij} a'_j = \sum_{j=1}^N \sum_{k=1}^N S_{ij} T_{jk} a_k , \quad (\text{A.26})$$

which we can express in terms of components as

$$\hat{U} = \hat{S}\hat{T} \xrightarrow{\mathcal{O}} U_{ik} = \sum_{j=1}^N S_{ij} T_{jk} . \quad (\text{A.27})$$

A.3.1 Operations on matrices

Definition 1.3.5: To obtain the **transpose** of a matrix, we interchange rows and columns. The transpose is denoted with a superscript T .

For example

$$\vec{a}^T = (a_1 \ a_2 \ \dots \ a_n) , \quad (\text{A.28})$$

and

$$(\mathbf{A})^T \xrightarrow{\mathcal{O}} \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix} . \quad (\text{A.29})$$

The transpose has the property

$$(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T . \quad (\text{A.30})$$

Definition 1.3.6: We say a matrix is **symmetric** if $\mathbf{A}^T = \mathbf{A}$ (ie. $A_{ij} = A_{ji}$).

Definition 1.3.7: We say a matrix is **anti-symmetric** if $\mathbf{A}^T = -\mathbf{A}$ (ie. $A_{ij} = -A_{ji}$).

Definition 1.3.8: To obtain the complex conjugate of a matrix, take the complex conjugate of each element.

For example

$$\mathbf{T}^* \xrightarrow{\mathcal{O}} \begin{pmatrix} T_{11}^* & T_{12}^* & \cdots & T_{1n}^* \\ T_{21}^* & T_{22}^* & \cdots & T_{2n}^* \\ \vdots & \vdots & \ddots & \vdots \\ T_{n1}^* & T_{n2}^* & \cdots & T_{nn}^* \end{pmatrix} \quad (\text{A.31})$$

Definition 1.3.9: A matrix is real if all its elements are real and imaginary if all its elements are imaginary.

ie. real: $\mathbf{T} = \mathbf{T}^*$

imaginary: $\mathbf{T} = -\mathbf{T}^*$

Definition 1.3.10: The **hermitian conjugate** (or adjoint) \mathbf{T}^\dagger of a matrix is the transpose of the complex conjugate.

For example

$$\mathbf{T}^\dagger \xrightarrow{\mathcal{O}} \begin{pmatrix} T_{11}^* & T_{21}^* & \cdots & T_{n1}^* \\ T_{12}^* & T_{22}^* & \cdots & T_{n2}^* \\ \vdots & \vdots & \ddots & \vdots \\ T_{1n}^* & T_{2n}^* & \cdots & T_{nn}^* \end{pmatrix} \quad (\text{A.32})$$

and

$$\vec{a}^\dagger = (\vec{a}^T)^* = (a_1^* \ a_2^* \ \dots \ a_n^*) , \quad (\text{A.33})$$

The hermitian conjugate has the property

$$(\mathbf{AB})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger . \quad (\text{A.34})$$

Definition 1.3.11:

We say a matrix is **hermitian** (or self-adjoint) if

$$\mathbf{A}^\dagger = \mathbf{A} . \quad (\text{A.35})$$

Definition 1.3.12: We say a matrix is **anti-hermitian** (or skew-hermitian) if $\mathbf{A}^\dagger = -\mathbf{A}$.

With this notation, we can write the inner product as

$$\langle \alpha | \beta \rangle = \vec{a}^\dagger \vec{b} \quad (\text{A.36})$$

In general, operators do not commutative, ie $\hat{S}\hat{T} \neq \hat{T}\hat{S}$.

Definition 1.3.13: The **commutator** is

$$[\hat{S}, \hat{T}] = \hat{S}\hat{T} - \hat{T}\hat{S}$$

The commutators tells us about the failure of two operators to commute.

Definition 1.3.14: The identity operator, \hat{I} , satisfies $\hat{A}\hat{I} = \hat{I}\hat{A} = \hat{A}$ for any operator \hat{A} . In any basis its components are

$$\hat{I} \xrightarrow{\mathcal{O}} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} , \quad (\text{A.37})$$

so that in terms of components we have $\hat{I}_{ij} = \delta_{ij}$.

Definition 1.3.15: The **inverse** of an operator is defined by

$$\hat{A}^{-1}\hat{A} = \hat{A}\hat{A}^{-1} = \hat{I} \quad (\text{A.38})$$

A matrix has an inverse iff its determinant is non-zero. In fact

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} (\mathbf{C}_\mathbf{A})^T , \quad (\text{A.39})$$

where $\mathbf{C}_\mathbf{A}$ is the matrix of co-factors

Definition 1.3.16: The cofactor of \mathbf{A}_{ij} is $(-1)^{i+j}$ times the determinant of the submatrix we get by removing the i -th row and j -th column.

For example if

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

then, $\det \mathbf{A} = ad - bc$, and

$$\mathbf{C}_\mathbf{A} = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$$

and

$$\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Definition 1.3.17: A matrix without an inverse is said to be **singular**.

For a product of operators, the inverse, if it exists, is given by

$$(\hat{A}\hat{B})^{-1} = \hat{B}^{-1}\hat{A}^{-1}. \quad (\text{A.40})$$

Definition 1.3.18: A matrix is **unitary** if its inverse is equal to its hermitian conjugate

$$\mathbf{U}^\dagger = \mathbf{U}^{-1}. \quad (\text{A.41})$$

Unitary operators have the important property that they preserve the inner product. Suppose, $|\alpha'\rangle = \hat{U}|\alpha\rangle$, then

$$\begin{aligned} \langle \alpha' | \beta' \rangle &= (\vec{a}')^\dagger \vec{b}' \\ &= (\mathbf{U}\vec{a})^\dagger (\mathbf{U}\vec{b}) \\ &= \vec{a}^\dagger \underbrace{\mathbf{U}^\dagger \mathbf{U}}_{\hat{I}} \vec{b} \\ &= \vec{a}^\dagger \vec{b} \\ &= \langle \alpha | \beta \rangle \end{aligned} \quad (\text{A.42})$$

A.4 Boxes

BOX 1.1

Show that properties 1 and 4 of the inner product imply (A.7).

BOX 1.2

Explain why the norm of a vector is always real.

BOX 1.3

Derive (A.8)

BOX 1.4

Prove the Schwarz inequality (A.11).

Hint: Let $|\gamma\rangle = |\beta\rangle - (\langle\alpha|\beta\rangle / \langle\alpha|\alpha\rangle) |\alpha\rangle$ and use $\langle\gamma|\gamma\rangle \geq 0$.

BOX 1.5

Find the angle (in the sense of (A.13)) between the vectors $|\alpha\rangle = (i)\hat{i} + (i)\hat{j} + \hat{k}$ and $|\beta\rangle = (4)\hat{i} + (2 - 2i)\hat{k}$.

BOX 1.6

Draw a few diagrams to sketch the outline of a proof that rotation acts linearly on vectors. Briefly explain your diagrams.

BOX 1.7

Show how we get (A.23).

BOX 1.8

Show that $(\mathbf{A}\vec{a})^T = \vec{a}^T \mathbf{A}^T$ and $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$.

Hint: use components

BOX 1.9

$$(\mathbf{AB})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger.$$

Hint: use components

BOX 1.10

Show (A.36).

BOX 1.11

Prove (A.40). Hint: multiply $\hat{A}\hat{B}$ by $\hat{B}^{-1}\hat{A}^{-1}$ on the right and on the left.

A.5 Exercises

Ex 1.1: Consider the collection of all polynomials (with complex coefficients) of degree less than N in x .

- (a) Does this set constitute a vector space (with the polynomials as “vectors”)? If so, suggest a convenient basis, and give the dimension of the space. If not, which of the defining properties does it lack?
- (b) What if we require that the polynomials be even functions?
- (c) What if we require that the leading coefficient (i.e., the number multiplying x^{N-1}) be 1?
- (d) What if we require that the polynomials have the value 0 at $x = 1$?
- (e) What if we require that the polynomials have the value 1 at $x = 0$?

Ex 1.2: Gram-Schmidt procedure

Suppose you start out with a basis $\{|e_1\rangle, |e_2\rangle, \dots, |e_n\rangle\}$ that is not orthonormal. The Gram-Schmidt procedure is a systematic ritual for generating from it an orthonormal basis $\{|e'_1\rangle, |e'_2\rangle, \dots, |e'_n\rangle\}$. It goes like this:

- (a) Normalize the first basis vector (divide by its norm):

$$|e'_1\rangle = \frac{|e_1\rangle}{\|e_1\|}$$

- (b) Find the projection of the second vector along the first, and subtract it off:

$$|f_2\rangle = |e_2\rangle - \langle e'_1 | e_2 \rangle |e'_1\rangle$$

This vector is orthogonal to $|e'_1\rangle$. Normalise $|f_2\rangle$ to get $|e'_2\rangle$.

- (c) Subtract from $|e_3\rangle$ its projections along $|e'_1\rangle$ and $|e'_2\rangle$:

$$|f_3\rangle = |e_3\rangle - \langle e'_2 | e_3 \rangle |e'_2\rangle - \langle e'_1 | e_3 \rangle |e'_1\rangle$$

This vector is orthogonal to $|e'_1\rangle$ and $|e'_2\rangle$. Normalise $|f_3\rangle$ to get $|e'_3\rangle$. And so on.

Use the Gram-Schmidt procedure to orthonormalize the 3-space basis

$$|e_1\rangle = (i)\hat{i} + (2-i)\hat{j} + (1)\hat{k}, |e_2\rangle = (1+i)\hat{i} + (3)\hat{j} + (1)\hat{k}, |e_3\rangle = (0)\hat{i} + (2)\hat{j} + (0)\hat{k}$$

Ex 1.3: Given the matrices

$$\mathbf{A} = \begin{pmatrix} 1 & i & 0 \\ 0 & i & 2 \\ 0 & 0 & -1 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 2 \\ i & i & 3 \end{pmatrix}$$

Calculate : (a) $\mathbf{A} + \mathbf{B}$ (b) \mathbf{AB} (c) $[\mathbf{A}, \mathbf{B}]$ (d) \mathbf{A}^T (e) \mathbf{A}^* (f) \mathbf{A}^\dagger (g) $\det \mathbf{A}$ (h) \mathbf{A}^{-1} Check $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. Does \mathbf{B} have an inverse?

Ex 1.4: By explicit construction of the matrices in question, show that any matrix \mathbf{T} can be written

- (a) as the sum of a symmetric matrix \mathbf{S} and an antisymmetric matrix \mathbf{A} ;
- (b) as the sum of a real matrix \mathbf{R} and an imaginary matrix \mathbf{M} ;
- (c) as the sum of a hermitian matrix \mathbf{H} and a skew-hermitian matrix \mathbf{K} .

Ex 1.5: Show that the rows and columns of a unitary matrix constitute orthonormal sets.

B

Review of linear algebra II

B.1 Changing basis

We've seen that the components of vectors and matrices depend on our choice of basis – we'd now like to investigate the relationship between components in different bases.

Suppose we switch from an old basis, $|e_i\rangle$, to new basis, $|f_i\rangle$. Since all vectors are expressible in terms of a basis, the $|e_i\rangle$'s can be written as a linear combination of the $|f_i\rangle$'s, we must have

$$\begin{aligned} |e_1\rangle &= S_{11}|f_1\rangle + S_{21}|f_2\rangle + \dots S_{n1}|f_n\rangle , \\ |e_2\rangle &= S_{12}|f_1\rangle + S_{22}|f_2\rangle + \dots S_{n2}|f_n\rangle , \\ &\vdots && \vdots \\ |e_n\rangle &= S_{1n}|f_1\rangle + S_{2n}|f_2\rangle + \dots S_{nn}|f_n\rangle , \end{aligned} \quad (\text{B.1})$$

for some numbers S_{ij} . This can be compactly written as

$$|e_j\rangle = \sum_{i=1}^N S_{ij}|f_i\rangle , \quad (\text{B.2})$$

which is a linear transformation. Comparing this with (A.17), following the same argument that lead to (A.20), leads us to

$$a_{(F)i} = \sum_{j=1}^N S_{ij}a_{(E)j} , \quad (\text{B.3})$$

where the capitalised subscripts in brackets denote the basis. In matrix form, we can write (B.3) as

$$\vec{a}_{(F)} = \mathbf{S}\vec{a}_{(E)} . \quad (\text{B.4})$$

How do the components of a matrix compare in different bases? In the old basis, for some matrix $\mathbf{T}_{(E)}$,

$$\vec{a}'_{(E)} = \mathbf{T}_{(E)}\vec{a}_{(E)} , \quad (\text{B.5})$$

but using (B.3) for \vec{a}' gives

$$\vec{a}'_{(F)} = \mathbf{S}\vec{a}'_{(E)} , \quad (\text{B.6})$$

so

$$\vec{a}'_{(F)} = \mathbf{S}\vec{a}'_{(E)} \stackrel{(\text{B.5})}{=} \mathbf{S}\mathbf{T}\vec{a}_{(E)} = \mathbf{S}\mathbf{T}_{(E)}\mathbf{I}\vec{a}_{(E)} = \mathbf{S}\mathbf{T}_{(E)}\mathbf{S}^{-1}\mathbf{S}\vec{a}_{(E)} \stackrel{(\text{B.4})}{=} \mathbf{S}\mathbf{T}_{(E)}\mathbf{S}^{-1}\vec{a}_{(F)} , \quad (\text{B.7})$$

so that comparing (B.5) and (B.7) we see that

$$\mathbf{T}_{(F)} = \mathbf{S}\mathbf{T}_{(E)}\mathbf{S}^{-1}. \quad (\text{B.8})$$

Definition 2.1.1: We say that two matrices, \mathbf{T}_1 and \mathbf{T}_2 , are **similar**, if $\mathbf{T}_2 = \mathbf{S}\mathbf{T}_1\mathbf{S}^{-1}$ for some non-singular matrix \mathbf{S} .

From (B.8) we see that matrices representing the same linear transformation with respect to different bases are similar. Although the components of a matrix with respect to different bases are different, two important numbers are unchanged – the determinant and the trace. For the determinant,

$$\det(\mathbf{T}_{(F)}) = \det(\mathbf{S}\mathbf{T}_{(E)}\mathbf{S}^{-1}) = \det(\mathbf{S})\det(\mathbf{T}_{(E)})\det(\mathbf{S}^{-1}) = \det(\mathbf{T}_{(E)}). \quad (\text{B.9})$$

Definition 2.1.2: The **trace** is the sum of the diagonal elements of a matrix:

$$\text{tr}(\mathbf{T}) = \sum_{i=1}^N T_{ii} \quad (\text{B.10})$$

The trace satisfies the important property

$$\text{tr}(\mathbf{T}_1\mathbf{T}_2) = \text{tr}(\mathbf{T}_2\mathbf{T}_1), \quad (\text{B.11})$$

so

$$\text{tr}(\mathbf{T}_{(F)}) = \text{tr}(\mathbf{S}\mathbf{T}_{(E)}\mathbf{S}^{-1}) = \text{tr}(\mathbf{T}_{(E)} \underbrace{\mathbf{S}^{-1}\mathbf{S}}_I) = \text{tr}(\mathbf{T}_{(E)}) \quad (\text{B.12})$$

B.2 Eigenvectors and eigenvalues

For a given linear transformation \hat{T} , the special vectors which satisfy

$$\hat{T}|\alpha\rangle = \lambda|\alpha\rangle, \quad (\text{B.13})$$

are called **eigenvectors** of \hat{T} and the numbers λ are called eigenvalues. In matrix form, (B.13), becomes

$$\mathbf{T}\vec{a} = \lambda\vec{a} \implies (\mathbf{T} - \lambda\mathbf{I})\vec{a} = 0. \quad (\text{B.14})$$

Now if $(\mathbf{T} - \lambda\mathbf{I})^{-1}$ exists, we can multiply both sides of (B.14) with it, which would lead to $\vec{a} = 0$. So, for non-trivial eigenvectors, we require that $(\mathbf{T} - \lambda\mathbf{I})$ is singular which in turn implies that $\det(\mathbf{T} - \lambda\mathbf{I}) = 0$. Expanding

$$\det(\mathbf{T} - \lambda\mathbf{I}) = \begin{vmatrix} (T_{11} - \lambda) & T_{12} & \dots & T_{1n} \\ T_{21} & (T_{22} - \lambda) & \dots & T_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ T_{n1} & T_{n2} & \dots & (T_{nn} - \lambda) \end{vmatrix} = 0, \quad (\text{B.15})$$

leads to some algebraic equation for λ

$$C_n\lambda^n + C_{n-1}\lambda^{n-1} + \dots + C_1\lambda + C_0 = 0. \quad (\text{B.16})$$

(B.16) is called the **characteristic equation**. Since, (B.16) is an n -th order polynomial equation, in general it has n complex roots. These root are not necessarily unique, so we conclude that an $n \times n$ matrix has at least 1 and at most n distinct eigenvalues.

For a given operator, the collection of eigenvalues is called the **spectrum** of the operator. If two or more linearly independent eigenvectors have the same eigenvalue we say they are **degenerate**.

To find the eigenvectors, it is easiest to solve (B.13) for a given λ directly (see BOX B.1).

If the eigenvectors of span the space, we can use them to write a special basis. Suppose

$$\begin{aligned}\hat{T}|f_1\rangle &= \lambda_1|f_1\rangle \\ \hat{T}|f_2\rangle &= \lambda_2|f_2\rangle \\ &\vdots \\ \hat{T}|f_n\rangle &= \lambda_n|f_n\rangle ,\end{aligned}\quad (\text{B.17})$$

then in the basis $|f_i\rangle$, \hat{T} has a very simple form

$$\mathbf{T}_{(F)} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \quad (\text{B.18})$$

A matrix which can be brought into the form (B.18) is called **diagonalisable**. The similarity matrix, \mathbf{S} , takes us to this new basis can be constructed by using the normalised eigenvectors (in the old basis) as the columns of \mathbf{S}^{-1} :

$$(\mathbf{S}^{-1})_{ij} = (\vec{f}_j)_i , \quad (\text{B.19})$$

where \vec{f}_j is the j -th eigenvector.

Diagonal matrices are much easier to work with but, unfortunately, not all matrices can be diagonalised. A sufficient (but not necessary) condition for a matrix to be diagonalisable is that it is **normal**.

Definition 2.2.3: A matrix, \mathbf{N} , is **normal** if it commutes with its hermitian conjugate – ie. if $[\mathbf{N}, \mathbf{N}^\dagger] = 0$

In particular we see that, since they are normal, all hermitian and unitary matrices are diagonalisable.

In quantum mechanics, given two diagonalisable matrices, we are often interested in whether they can be simultaneously diagonalised – ie. is there a basis in which both matrices are diagonal. It turns out that two matrices can be simultaneously diagonalised if they commute.

B.3 Hermitian transformations

Previously we defined the hermitian conjugate (or adjoint) of a matrix as the transpose of its conjugate: $\mathbf{M}^\dagger = (\mathbf{M}^T)^*$. We can more generally define the hermitian conjugate, \hat{T}^\dagger , of a linear

transformation, \hat{T} , as the operator such that for all vectors, $|\alpha\rangle$ and $|\beta\rangle$,

$$\langle \hat{T}^\dagger \alpha | \beta \rangle = \langle \alpha | \hat{T} \beta \rangle , \quad (\text{B.20})$$

where $\langle \hat{T}^\dagger \alpha | \beta \rangle$ is the inner product of $\hat{T}^\dagger |\alpha\rangle$ and $|\beta\rangle$, while $\langle \alpha | \hat{T} \beta \rangle$ is the inner product of $|\alpha\rangle$ and $\hat{T} |\beta\rangle$. In an orthonormal basis, the hermitian conjugate of a linear operator is the hermitian conjugate of the matrix it represents:

$$\langle \alpha | \hat{T} \beta \rangle = \vec{a}^\dagger \mathbf{T} \vec{b} = (\mathbf{T}^\dagger \vec{a})^\dagger \vec{b} = \langle \hat{T}^\dagger \alpha | \beta \rangle . \quad (\text{B.21})$$

Theorem 2.3.1: The eigenvalues of a hermitian operator are real.

Proof. Let λ be the eigenvalue of \hat{T} , with non-zero eigenvector $|\alpha\rangle$, so that $\hat{T} |\alpha\rangle = \lambda |\alpha\rangle$, and

$$\langle \alpha | \hat{T} \alpha \rangle = \langle \alpha | \lambda \alpha \rangle = \lambda \langle \alpha | \alpha \rangle , \quad (\text{B.22})$$

but on the other hand if \hat{T} is hermitian,

$$\langle \alpha | \hat{T} \alpha \rangle = \langle \hat{T}^\dagger \alpha | \alpha \rangle = \langle \hat{T} \alpha | \alpha \rangle = \langle \lambda \alpha | \alpha \rangle = \lambda^* \langle \alpha | \alpha \rangle , \quad (\text{B.23})$$

and comparing (B.22) and (B.23), we conclude that $\lambda = \lambda^*$ which implies that λ is real. \square

Theorem 2.3.2: The eigenvectors of hermitian operators with distinct eigenvalues are orthogonal.

Proof. Suppose we have a hermitian operator, \hat{T} , with $\hat{T} |\alpha\rangle = \lambda |\alpha\rangle$, $\hat{T} |\beta\rangle = \mu |\beta\rangle$ with $\mu \neq \lambda$. Then

$$\langle \alpha | \hat{T} \beta \rangle = \langle \alpha | \mu \beta \rangle = \mu \langle \alpha | \beta \rangle , \quad (\text{B.24})$$

but on the other hand

$$\langle \alpha | \hat{T} \beta \rangle = \langle \hat{T}^\dagger \alpha | \beta \rangle = \langle \hat{T} \alpha | \beta \rangle = \langle \lambda \alpha | \beta \rangle = \lambda^* \langle \alpha | \beta \rangle , \quad (\text{B.25})$$

and since $\mu \neq \lambda = \lambda^*$, comparing (B.24) and (B.25), we conclude that $\langle \alpha | \beta \rangle = 0$. \square

Theorem 2.3.3: The eigenvectors of a hermitian transformation span the space.

Proof. The statement that eigenvectors span a space is true if the operator is diagonalisable which we know to be true for hermitian operators. \square

B.4 Boxes

BOX 2.1

In the derivation of (B.8) we assumed that \mathbf{S} has an inverse. Explain how we know that this must be true.

BOX 2.2

Explain the last step of (B.9) .

BOX 2.3

Prove (B.11).

BOX 2.4

Prove (B.18).

BOX 2.5

Show that unitary and hermitian matrices are normal.

BOX 2.6

Show that if two matrices commute in one basis, then they commute in any basis. That is:

$$[\mathbf{T}_{(E)1}, \mathbf{T}_{(E)2}] = 0 \implies [\mathbf{T}_{(F)1}, \mathbf{T}_{(F)2}] = 0 \quad (\text{B.26})$$

B.5 Exercises

Ex 2.1: Find the eigenvalues and eigenvectors of

$$\mathbf{M} = \begin{pmatrix} 1 & i & -1 \\ 0 & -i & 1 \\ -1 & 1 & 0 \end{pmatrix} \quad (\text{B.27})$$

Ex 2.2: Construct the similarity matrix \mathbf{S} which diagonalises (B.27) in Ex B.1 and check that $\mathbf{S}\mathbf{M}\mathbf{S}^{-1}$ is diagonal with the eigenvalues of \mathbf{M} along the diagonal.

Ex 2.3: Show that the first, second, and last coefficients in the characteristic equation (B.16) are:

$$C_n = (-1)^n, \quad C_{n-1} = (-1)^{n-1} \operatorname{tr}(\mathbf{T}), \quad \text{and } C_0 = \det(\mathbf{T}). \quad (\text{B.28})$$

Ex 2.4: It's obvious that the trace of a diagonal matrix is the sum of its eigenvalues, and its determinant is their product. It follows that the same holds for any diagonalizable matrix (using (B.9) and (B.12)). Prove that in fact

$$\det(\mathbf{M}) = \lambda_1 \lambda_2 \dots \lambda_n, \quad \operatorname{tr}(\mathbf{M}) = \lambda_1 + \lambda_2 + \dots + \lambda_n \quad (\text{B.29})$$

for any matrix. (The λ_i are the n solutions to the characteristic equation – in the case of multiple roots, there may be fewer linearly independent eigenvectors than there are solutions, but we still count each λ as many times as it occurs.)

Hint: Write the characteristic equation in the form $(\lambda_1 - \lambda)(\lambda_2 - \lambda) \dots (\lambda_n - \lambda) = 0$. and use the result of Ex B.3.

Ex 2.5: Let

$$\mathbf{T} = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$$

- (a) Verify that \mathbf{T} is hermitian.
- (b) Find its eigenvalues (note that they are real).
- (c) Find and normalize the eigenvectors (note that they are orthogonal).
- (d) Construct the unitary diagonalizing matrix \mathbf{S} , and check explicitly that it diagonalizes \mathbf{T} .
- (e) Check that $\det(\mathbf{T})$ and $\operatorname{tr}(\mathbf{T})$ are the same for \mathbf{T} as they are for its diagonalised form.

Ex 2.6: A unitary transformation is one for which $\hat{U}^\dagger \hat{U} = 1$.

- (a) Show that unitary transformations preserve inner products, in the sense that $\langle \hat{U}\alpha | \hat{U}\beta \rangle = \langle \alpha | \beta \rangle$, for all vectors $|\alpha\rangle$ and $|\beta\rangle$.
- (b) Show that the eigenvalues of a unitary transformation have modulus 1.
- (c) Show that the eigenvectors of a unitary transformation belonging to distinct eigenvalues are orthogonal.

Ex 2.7: Show that if two matrices are simultaneously diagonalisable, they commute.

Bibliography

M. Brune, E. Hagley, J. Dreyer, X. Maître, A. Maali, C. Wunderlich, J. M. Raimond, and S. Haroche. Observing the progressive decoherence of the “meter” in a quantum measurement. *Phys. Rev. Lett.*, 77:4887–4890, Dec 1996. doi: 10.1103/PhysRevLett.77.4887. URL <https://link.aps.org/doi/10.1103/PhysRevLett.77.4887>.

W. Gerlach and O. Stern. Der experimentelle nachweis der richtungsquantelung im magnetfeld. *Zeitschrift für Physik*, 9(1):349–352, 1922.

D. J. Griffiths. *Introduction to quantum mechanics*. Cambridge University Press, 2016.

W. B. Jensen. The origin of the s, p, d, f orbital labels. *Journal of Chemical Education*, 84(5):757, 2007. doi: 10.1021/ed084p757. URL <https://doi.org/10.1021/ed084p757>.

D. H. McIntyre, C. A. Manogue, and J. Tate. *Quantum mechanics: a paradigms approach*, volume 192. Pearson Boston, 2012.

A. L. Pérez-Martínez, M. d. P. Aguilar-Del-Valle, and A. Rodríguez-Gómez. Do it by yourself: An instructional derivation of the laplacian operator in spherical polar coordinates. *Mathematics*, 9(22):2943, 2021. doi: 10.3390/math9222943. URL <https://www.mdpi.com/2227-7390/9/22/2943>.