

Guru's Guide to Quantum Mechanics I

Joel Gilmore

October 7, 2001

Contents

1 Preliminaries	1
1.1 Discrete Probability and Expectation Values	1
1.2 Continuous Probability, Continuous Expectation Values and Density Functions	2
1.3 Waves	4
1.4 A word on notation	5
2 Development of Quantum Mechanics	7
2.1 Blackbody radiation	7
2.2 Quantisation	11
2.3 Photoelectric effect	13
2.4 Compton Effect	14
2.5 de Broglie Wavelength	16
2.6 Bohr and the atom	17
2.7 Chapter Summary	19
2.8 Exercises	20
3 Wavefunctions	21
3.1 Introducing the wavefunction	21
3.2 Fundamentals of Wave Theory	22
3.3 Free particle momentum states	23
3.4 Superposition	25
3.5 Chapter Summary	26
3.6 Exercises	27
4 The Momentum Representation	29
4.1 The Dirac Delta Function	29
4.2 Fourier transforms	34
4.3 Momentum representation	36
4.4 Pure momentum states revisited	38
4.5 Chapter Summary	40
4.6 Exercises	41

5	Analysing the Wave function	43
5.1	Expectation values and variance	43
5.2	Operators	48
5.3	Properties of Operators	50
5.4	Expectation Values Again	52
5.5	Hermitian Operators	55
5.6	The Commutator Bracket	56
5.7	Eigenfunctions and Eigenvalues	59
5.8	Interpreting the Eigenvalues	61
5.9	The Eigenvalue Spectrum	63
5.10	Degeneracy	64
5.11	Properties of Eigenfunctions	66
5.11.1	Eigenvalues of Hermitian Operators	66
5.11.2	Orthogonality of Eigenfunctions	66
5.11.3	Completeness	69
5.12	Expanding Expectation Values	71
5.13	Interpreting the Eigenfunctions	72
5.13.1	How to Analyse a Wave Function	73
5.13.2	The Copenhagen Collapse Hypothesis	74
5.14	Bringing it together	75

Introduction

Welcome to Guru's Guide to Quantum Mechanics I. This is an introductory course into quantum mechanics, typical of the sort of course that is taught in second year, or as the introduction to a higher level subject. The course will cover

- Development of Quantum Mechanics – blackbody radiation and the Ultra-violet Catastrophe, photoelectric effect, Compton scattering, de Broglie's wavelength, Bohr Model of the atom
- Wavefunctions and their interpretation
- Dirac δ -functions, Fourier transforms and their connection to representations.
- Operators
- Eigenvalues, eigenfunctions and their physical interpretation
- Heisenberg Uncertainty Principle
- The Schrodinger equation
- The Hamiltonian, its spectrum, and its correspondence to classical dynamics
- Probability current
- Applications such as step potential, potential well and molecular lattices

(And hopefully, by the end of the course, these will make a little more sense.)

Prerequisites

This course assumes you are familiar with the basic of a range of physics areas; the sort of things that would be covered in a first year introductory course to physics. We won't require advanced equations, but you do need to understand some of the basic concepts (such as waves, diffraction, etc.)

A good working knowledge of integration and differentiation is also required, and you should understand the “physical” meaning of each (i.e. not just that integration is area under the curve, but what that area actually means e.g. area under a velocity graph gives displacement). Some familiarity with probability would be useful, but you can pick it up as you go along if need be.

Guide Through the Chapters

This course is basically linear - you need to read each chapter before proceeding to the next, and the later chapters won't make much sense without the earlier ones.

Chapter 1 is about the only chapter you could realistically skip. It reviews some of the basic concepts, particularly about probability, that need for this course. It is recommended, however, that you at least browse through it, and confirm that you do indeed know it all.

Chapter 2 covers the history and evolution of quantum mechanics, and it will help you to get a feel for why we justify the assumptions we make in later chapters. However, it *is* mainly background, so don't stress too much over it. Just convince yourself that the derivations are legitimate, and perhaps return to it later after completing the course.

Acknowledgments

Thank you to Jenny, Mark and everyone else who helped to proof read this course - your input was always helpful.

And to every abysmal, unintelligible, drab, dull and generally uneducational text book that motivated the Guru's Guide series. Thankyou.

Chapter 1

Preliminaries

Before we begin our development of quantum, there are some basic concepts that you need to understand.

1.1 Discrete Probability and Expectation Values

There are times when we wish to predict what the average value would be of a series of measurements. We can do this if we know the probability of obtaining each possible outcome. This probability might also be derived theoretically, or it could be calculated from a previous experiment. Let's say we are measuring some variable E , and that the possible measurements are E_n , and the probability of obtaining them is P_n . This is a **discrete** system - shoesize, pagenumbers and the number of points on your license are all discrete. Things like height and volume are known as **continuous** variables, because they can take any value you like.

If we took N samples, we would expect to get the value E_1 approximately $P_1 \cdot N$ times, and similarly for $n = 2$, etc. Then the average would be the sum of all these measurements, divided by N , or mathematically:

$$\begin{aligned}\bar{E} &= \frac{\sum_n E_n \cdot (P_n N)}{N} \\ &= \sum_n E_n \cdot P_n\end{aligned}$$

The average, particularly when used in physics, is also called the **expectation value** of E , written $\langle E \rangle$.

Example.

Say we were measuring the age (in whole years) of third year university students. We find there is a 60% probability of them being 20 years old, a 30% chance of them being 19 and a 10% chance of them being 21. If we sampled 10 students,

we should find $10 \times .6 = 6$ 20 year olds, three 19 year olds and one 21 year old. Taking the average we get:

$$\begin{aligned}\bar{E} &= \frac{20 + 20 + 20 + 20 + 20 + 20 + 19 + 19 + 19 + 21}{10} \\ &= \frac{20 \times 6 + 19 \times 3 + 21 \times 1}{10} \\ &= \frac{20 \times (10 \times .6) + 19 \times (10 \times .3) + 21 \times (10 \times .1)}{10} \\ &= 20 \times .6 + 19 \times .3 + 21 \times .1\end{aligned}$$

which is exactly the formula we found above. Notice that it is completely independent of the number of samples we took, as we would expect.

Normally, the sum of all the probabilities should add to give unity. But sometimes, for whatever reason, we calculate probabilities which make relative sense (the probability of A is twice that of B) but aren't normalised. Therefore, before we can use them, we need to scale them down. To do this, we say

$$P'_m = \frac{P_m}{\sum_n P_n}$$

where the P_n (which includes P_m) are the old probabilities, and P'_m is the new, normalised probability.

Example.

We find three probabilities $P_1 = 1.0$, $P_2 = 2.0$ and $P_3 = 3.0$. Clearly these don't add to unity, so:

$$\begin{aligned}P'_1 &= \frac{1}{1 + 2 + 3} \approx .167 \\ P'_2 &= \frac{2}{1 + 2 + 3} \approx .333 \\ P'_3 &= \frac{3}{1 + 2 + 3} = .5\end{aligned}$$

and clearly the new probabilities add to give 1. Notice also that P'_2 is still twice P'_1 and so on.

1.2 Continuous Probability, Continuous Expectation Values and Density Functions

If we need to calculate the expectation value of some continuous variable (such as height or length) then we can't use a sum any more, because a sum only

works for discrete systems - with continuous variables it can't take into account all possible values¹.

Instead of calculating the probability of each measurement (the P_n 's) we could make a new function $D(x)$ called the **probability density**, which gives the probability per unit x , at each point. Alternatively, we can say $D(x) \times dx$, where dx is a small width of x , gives the approximate probability of getting a measurement between x and $x + \Delta x$. So, for example, we could make the probability density for height $D(h)$, so that $D(100\text{cm}) \times 2\text{cm}$ gives the (approximate) probability of someone having a height between 100cm and 102cm . We say approximate because it is only exact when dx is very small i.e. an infinitesimal width - just like we use in integration. To find out exactly the probability of someone being between 100cm and 102cm tall, we integrate from 100 to 102, as in

$$P(100\text{cm to } 102\text{cm}) = \int_{100}^{102} D(x) dx$$

In the limit as dx goes to zero, $D(x)dx$ would be equivalent to $P(x)$. Thus, the average value of E for the discrete and continuous case becomes

$$\begin{aligned} \bar{E}_{discrete} &= \sum_n E_n \cdot P_n \\ \bar{E}_{continuous} &= \int E(x) D(x) dx \end{aligned}$$

where we integrate over the entire range of possible results (e.g. for height it would be from 0 to ∞ , although we would hope that $D(\infty)$ would be zero!)

Because we want the probability to be normalised, the total probability, and hence the integral of $D(x)$ over its range, must be unity:

$$\int_{\text{all space}} D(x) dx = 1$$

meaning the area under a graph of $D(x)$ against x must be 1.

If some of that is confusing, most statistics books will cover this in more detail, and more rigorously. For now, just accept that result - it will be used quite a bit during the course.

In addition to probability, this idea of a density function is quite common for other things in physics. Even velocity is like density - it is "distance per unit time". To find the actual distance, we multiply $v(t)$ by dt , i.e.

$$ds = v(t) dt$$

To find the total distance travelled, we integrate over time, as in

$$\Delta s = \int_{t_1}^{t_2} v(t) dt$$

¹Mathematically, a sum only works for **countable** variables - where you could match up each possible measurement with an integer. Things like height and the real numbers are **uncountable** - there's more of them than integers. Trust us!

Reversing this idea, given $ds(t)$ we could find the density function by dividing by dt , i.e.

$$D(t) = \frac{ds(t)}{dt} \equiv v(t)$$

Hence to find a density function in general, we divide by an infinitesimal (e.g. dx , dt , dv , etc) of our choice.

1.3 Waves

Because they are used so frequently throughout this course, we will briefly go over the basics of wave mechanics.

A wave is defined by three things - a frequency, a wavelength and an amplitude. The **frequency** f (or sometimes ν) is measured in **Hertz**, with units s^{-1} , and is a measure of how fast the wave is moving. The **wavelength** λ is measured in metres, and gives the distance one peak and the next; it is a measure of how spread out in space the wave is. The **amplitude** is the height of the wave - the vertical (not diagonal) distance from peak to trough. The **period** of the wave, the time taken for the wave to move one full cycle, is given by

$$T = \frac{1}{f}$$

Commonly, **angular frequency** and **wave number** are used instead of frequency and wavelength. The angular frequency ω is given by

$$\omega = 2\pi f = \frac{2\pi}{T}$$

and the wave number is

$$k = \frac{2\pi}{\lambda}$$

Why, I hear you ask? We do this so that a wave (call it Ψ) can be described by the function

$$\Psi = A \sin(kx - \omega t)$$

Imagine tracing out this function, first through time and then through space, and verify to yourself that it does indeed give a wave with period T and a wavelength λ (and an amplitude A).

The amplitude of a wave determines the amount of energy it has - bigger vibrations means bigger energy. We also talk about the **intensity** of a wave, which is proportional to the amplitude squared. When we talk about how bright light is, we are really talking about how intense it is, and hence how much energy.

We will also use three dimensional waves in this course. These waves are again specified by wavenumber and angular frequency; this time, however, the

wavenumber needs to be expressed as a vector \mathbf{k} . Hence, the formula for a three dimensional wave becomes

$$\begin{aligned}\Psi &= A \sin(\mathbf{k} \cdot \mathbf{r} - \omega t) \\ &= A \sin(k_x x + k_y y + k_z z - \omega t)\end{aligned}$$

where $\mathbf{k} = x \hat{\mathbf{i}} + y \hat{\mathbf{j}} + z \hat{\mathbf{k}}$

Unfortunately (for you!) we call both \mathbf{k} and the *length* of \mathbf{k} the wavenumber: \mathbf{k} relates to the wavelength of the wave in the x, y and z directions, while its length $k = \sqrt{k_x^2 + k_y^2 + k_z^2}$ gives the wavenumber if we were looking along the direction of the wave (i.e. if we were to move the wave around so that it ran along the x -axis, and only considered that axis, its wavenumber would be $k = \sqrt{k_x^2 + k_y^2 + k_z^2}$). Usually, it is clear from the context which we are referring to; if not, you will be told explicitly. Note that (for a set frequency and amplitude) there are many waves with a given k (e.g. both $\Psi = k \hat{\mathbf{i}}$ and $\Psi = k \hat{\mathbf{j}}$ have wavenumber k), but each \mathbf{k} defines a unique wave.

1.4 A word on notation

Physicists are slack. It's a fact of life. We (with the help of the mathematicians) develop hideously convoluted notation, just to save us a few lines of paper. And this course is no different, although we try and explain all our shortcuts before we use them. To start with, we often write

$$\int f(\mathbf{r}) d\mathbf{r}$$

where $\mathbf{r} = x \hat{\mathbf{i}} + y \hat{\mathbf{j}} + z \hat{\mathbf{k}}$. That equation is shorthand for

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y, z) dx dy dz$$

In general, when we want to say something is a function of position, we will just write it as a function of \mathbf{r} , and not of its individual components, x, y and z . The same goes for momentum, or any other similar variable.

Whenever you see an integral without limits, for *most* of this course it means to integrate over all valid space. This may mean to integrate over every position in space (for $d\mathbf{r}$), over every possible value of momentum (for $d\mathbf{p}$) or perhaps even for all positive frequencies (for $d\omega$). Usually, it will be clear from the context, but if not, there will be a note of explanation. Just don't be confused into thinking these are indefinite integrals and then wondering why we keep forgetting to add a constant!

Finally, occasionally in this course we won't use the technically correct language. Perhaps you won't even notice, but just be aware that in matters not pertaining directly to Quantum, the terminology may not all be strictly correct.

Chapter 2

Development of Quantum Mechanics

In this chapter we'll discuss some of the problems with classical theory - predictions that disagree with experiment. These are by no means all of them, it's just that these are the experiments that gave predictions *so* wrong, that there was no chance of patching up the old theories with extra terms. These results and others forced scientists to look for alternatives to the classical theories, and eventually resulted in quantum mechanics.

You'll notice that we mention a plethora of scientists in this discussion. As is the way of most of the great theories, quantum was developed over a length of time, by many theorists, experimentalists and even calligraphers! And here we list only a few - be certain that there are many more (not least of all the students doing the research) that we don't mention here.

2.1 Blackbody radiation

One of the first problems arose from considering *black body radiation*. A black body is one which absorbs all radiation that hits it - none is reflected. Thus, any radiation coming off it must be due to it being hot¹ and we find in experiments that the intensity of the various frequencies of light emitted depends only on the temperature of the black body. What we see is the black body radiation spectrum, as in Figure ***.

It turns out that a "cavity", a closed container of some shape and material, can be thought of as a black body with its walls at a temperature T . The walls will emit radiation into the centre of the cavity, and if we were sitting inside, we would see the black body spectrum.

In 1900, *Rayleigh* attempted to derive this spectrum theoretically, calculating this *energy density of radiation* within the cavity - how much energy

¹By "hot" we mean anything above absolute zero. Therefore, everything gives off heat!

per frequency (per volume) there was inside the box, as a function of T , the temperature of the walls. Mathematically, the energy density is $u(\omega, T)$ such that

$$\begin{aligned} u(\omega, T) &\equiv \text{Energy/Volume/(unit ang. freq.) at temp. } T, \text{ such that} \\ u(\omega, T)d\omega &= \text{Energy per volume between } \omega \text{ and } \omega + d\omega \end{aligned}$$

This is an example of the density functions discussed in Section 1.2, and while it may sound a little complicated, it is just what we plot in black body spectrum diagram - the intensity of light at each frequency.

How do we go about finding this density? The approach that Rayleigh took was to find the amount of radiation (per volume) between a *range* of frequencies - so between ω and $\omega + d\omega$. Then, he could divide this energy by the width of the band (namely, $d\omega$) to find the energy per frequency.

Firstly, it is possible to show that the energy density is completely independent of the shape of the cavity, and even of the material it's made of, by use of thermodynamics. Therefore, we can simplify our analysis by choosing a cube shaped cavity, with side length L , and with walls made of metal and therefore electrically conducting. Because we're talking about radiation, there is an electric field \mathbf{E} through the cavity, so it makes sense to consider the wave equation for the electric field,

$$\frac{\partial^2 \mathbf{E}}{\partial t^2} = c^2 \left[\frac{\partial^2 \mathbf{E}}{\partial x^2} + \frac{\partial^2 \mathbf{E}}{\partial y^2} + \frac{\partial^2 \mathbf{E}}{\partial z^2} \right] \quad (2.1)$$

If you haven't seen this before, don't panic - it's a rather straight forward result that comes from a very long derivation. It relates the way a light wave changes with time to how it moves through space, and it's enough (for this course) just to accept that it's true.

You may (or may not!) recall from electromag theory that inside a cavity with conducting walls, the electromagnetic field must vanish at the sides. If it didn't, then the wave would be reflected back, interfering with itself until it bounced off the other side, then interfering with both copies of itself on the way back. Eventually, all the bounces would cancel out the wave. Therefore, the EM field would look like a standing wave fixed at zero at both ends, just like standing waves on a fixed string, and this must hold for each component of the field (See Fig ***).

Solving the wave equation for the Z component of the field yields

$$E_z = E_{z_0} \sin(k_x x + \alpha) \sin(k_y y + \beta) \sin(k_z z + \gamma) \sin(\omega t + \delta), \quad (2.2)$$

where $\mathbf{k} = k_x \hat{\mathbf{i}} + k_y \hat{\mathbf{j}} + k_z \hat{\mathbf{k}}$ is the wavenumber of the standing wave. (see Exercise ***) Note Equation 2.2 isn't the total electric field in the box - this is just the field of a single standing wave, defined by \mathbf{k} . There are an infinite number of standing waves possible and the field inside the cavity will be a combination of *all* of them.

But we're forgetting that there are some constraints on the wavenumbers of the standing waves; namely, they have to *be* standing waves! By definition, we

need the electric field to drop to zero at the edges of the box, so E_z must be zero at $x = 0$, $x = L$, $y = 0$ and $y = L$. If we solve Equation 2.2 equal to zero with each of these constraints in turn, we must have

$$\begin{aligned} \alpha = 0 & \quad , & \quad k_x L = n_x \pi \\ \beta = 0 & \quad , & \quad k_y L = n_y \pi, \end{aligned}$$

where n_x and n_y are positive integers. And because we could solve Equation 2.1 just as easily for E_x and E_y , which must also vanish at $z = 0$ and $z = L$, we can also say

$$\gamma = 0, \quad k_z L = n_z \pi$$

So the only permissible standing waves are those where

$$k_x = \frac{n_x \pi}{L} \quad k_y = \frac{n_y \pi}{L} \quad k_z = \frac{n_z \pi}{L}$$

and the electric field inside the cavity will be a combination of *all* these standing waves. The total energy between ω and $\omega + d\omega$ will be the number of standing waves with frequency in that range multiplied by the energy of each wave. It is a property of light waves that $\omega = ck$ (where this k is the length of \mathbf{k}), so we need to know the number of waves with wavenumber between $\frac{\omega}{c}$ and $\frac{\omega+d\omega}{c}$. But where do we go from here? We need to introduce a new concept - \mathbf{k} -space.

If we have a point in space (a, b, c) we can plot it on a 3D graph with x , y and z on the axis. Now, consider that each standing wave is identified by a point $(k_x = \frac{n_x \pi}{L}, k_y = \frac{n_y \pi}{L}, k_z = \frac{n_z \pi}{L})$. Imagine that we had a graph with k_x , k_y and k_z on the axis. Then we could plot each of these points on the graph, which would form a three dimensional grid, with points spaced at $\frac{\pi}{L}$ intervals. We call this graph **\mathbf{k} -space**, and just like a point in normal space refers to a particular position, a point in \mathbf{k} -space refers to a wave with that particular wavenumber.

Now, imagine that I had little cubes of side s , and I was stacking them. Then, if I plotted the position of their centres, I would get a grid of points all separated by distance s along each axis. Given some volume of space V , the total number of cubes would be approximately V/s^3 .

We see that same grid pattern when plotting allowed standing waves in \mathbf{k} -space, so we can imagine that each point in the grid taking up a volume $(\pi/L)^3$, and hence if I had a "volume" in \mathbf{k} -space (whatever that means!) the number of standing waves inside it would be $V/(\pi/L)^3$.

But how does this apply to our problem? Imagine we were looking at all waves with wavenumber less than some $k = \sqrt{k_x^2 + k_y^2 + k_z^2}$. In \mathbf{k} -space, this would look like a sphere. What if we wanted all waves with wavenumber between k and $k + dk$? This range (if we wanted any wave, not just standing waves) would look like a thin shell, with inside radius k and width dk (Figure ***), but because k_x, k_y and k_z are positive, we want only the first quadrant. We can approximate its volume, if dk is small (which it is), by multiplying its surface area ($4\pi k^2$ divided by 8) by its width (dk). Hence, its volume in k -space is

$$\frac{1}{8} 4\pi k^2 \cdot dk,$$

and the number of standing waves inside this volume would be

$$\frac{1}{8} \frac{4\pi k^2 dk}{(\pi/L)^3}$$

Thus the number of standing waves between ω and $\omega + d\omega$ should be

$$\frac{1}{8} \frac{4\pi\omega^2 d\omega}{c^3 (\pi/L)^3}$$

However, we have neglected that light has two independent planes of polarisation, so for each wavenumber, there are two distinct waves possible. Thinking of waves on a string, each wave on the string has a particular wavelength, and hence wave number, but can either be vibrated up and down, or side to side, corresponding to polarisation. (Any other direction is merely a combination of those two.) We call each of these standing waves, determined by their wavenumber and polarisation, a *mode*. Therefore, the total number of modes is

$$N = \frac{\pi\omega^2 d\omega}{c^3 (\pi/L)^3} \quad (2.3)$$

Still on our string analogy, the harder we shake the string, the bigger the waves. Therefore we expect that the energy of a mode depends on the amplitude of its oscillations, which could potentially be anything.

To find the *average* energy, Rayleigh used the *equipartition* theory, that assigns an average energy of $\frac{1}{2}k_b T$ to each *degree of freedom* of a system, where k_b is Boltzmann's constant 1.37×10^{-23} , *not* to be confused with the wavenumber k . So for example a molecule with three degrees of freedom, corresponding to the three directions of motion, should have average energy $\frac{3}{2}k_b T$. Rayleigh said that because a light wave consists of a magnetic field and an electric field, two independent properties, each mode should have average energy $k_b T$ (note that this has nothing to do with polarisation).

Alternatively, we could say that the energies of the modes are *Boltzmann distributed*, meaning the probability of a mode at temperature T having an energy E is $e^{-E/k_b T}$ (the *Boltzmann factor*) So (cf Section 1.2) we have

$$\bar{E} = \int_0^\infty E \cdot P(E) dE = \int_0^\infty E e^{-E/k_b T} dE = k_b T$$

We will use the Boltzmann distribution again in a moment, so make sure you follow it.

The energy in the box between ω and $\omega + d\omega$ is

$$\begin{aligned} \text{total } E &= \text{No. of waves} \times \text{average energy/wave} \\ &= N \times \bar{E} \\ &= \frac{\pi\omega^2 d\omega}{c^3 (\pi/L)^3} k_B T \end{aligned}$$

and finally by dividing by the volume of the cavity (L^3) and by $d\omega$ the energy density per unit angular frequency is

$$u(\omega, T) = \frac{\omega^2}{\pi^2 c^3} k_B T \quad (2.4)$$

which is known as the **Rayleigh-Jeans formula**. (Jeans' name was added because he corrected a mistake in Rayleigh's original derivation.)

This whole derivation appears to be rigorous and indeed, classically, it is. Unfortunately, there's a problem - it's wrong. The energy density depends on ω^2 , which means it increases to infinity at high frequencies no matter what the temperature. The total energy, the area under the $u(\omega, T)$ graph, is infinite, implying that our cavity contains an infinite amount of energy, most of it being ultraviolet rays, x-rays, and so on. If we were to make a small hole in the box, this energy would come rushing out and burn us to a crisp. However, this is clearly at odds with experiment, and hence the prediction of the Rayleigh-Jeans formula is known as the **Ultraviolet Catastrophe**.

So where's the problem? Figure **** shows the experimental and theoretical graphs of energy density. At low temperatures, the two match, but the theoretical function diverges. So our theory and derivation can't be completely wrong - but there must be an error either in our calculation of the average energy or the number of modes. A tip off is that other theories using the equipartition theory were also having difficulties (though nothing as serious as a death-ray shoebox!) Therefore, our average energy must be wrong - though that was hard to believe for the physicists of 1900, particularly with two independent derivations to stare at.

2.2 Quantisation

Enter **Planck**, another physicist of 1900 also interested in this problem. He decided to tackle it from a different direction - he tried to find a mathematical *fit* to the observed black body spectrum, and see if that would shed any light on the problem (pun intended). Remarkably, he discovered that

$$\frac{\omega^2}{\pi^2 c^3} \frac{\hbar \omega}{e^{\hbar \omega / k_B T} - 1} \quad (2.5)$$

could be made to fit the experimental data by changing the value of a scaling constant he called h . Planck found that a value of 6.626×10^{-34} made his curve fit exactly, no matter what material, shape or temperature the cavity was. h is in fact a universal constant, known (unsurprisingly) as **Planck's constant**.

What Planck needed to do now, was to come up with a physical derivation of that formula, work he later called the most difficult in his life. Planck eventually (in yet another ingenious move!) considered what the consequences would be if, for some reason, the allowed energies of each mode (standing wave) were not continuous; instead a mode of frequency ω could only take on energies of $\hbar \omega$ (just like in Planck's black body model), where $\hbar = \frac{h}{2\pi}$ (Although h

is formally Planck's constant, we use $\frac{h}{2\pi}$ so often that we write \hbar (read "h bar") as a shorthand.) This initially sounds like a crazy idea; as we said, classically the energy of a wave (at a given frequency) is determined by the amplitude of its oscillations, and *any* amplitude is fine. However, let's see what this **quantisation** of energy does.

If the energies are **quantised**, then the probability of having energy $n\hbar\omega$ would be

$$P(E = n\hbar\omega) = \frac{e^{-n\hbar\omega/k_B T}}{\sum_{n=0}^{\infty} e^{-n\hbar\omega/k_B T}},$$

that is, the probability of energy $n\hbar\omega$ is its classical probability, divided by the sum of the probabilities of all allowed energies. (Needed to normalise the probabilities – see Section 1.2)

Thus,

$$\begin{aligned} \bar{E} &= \sum_n E_n \cdot P(E_n) \\ &= \sum_n n\hbar\omega \frac{e^{-n\hbar\omega/k_B T}}{\sum_m e^{-m\hbar\omega/k_B T}} \\ &= \hbar\omega \frac{\sum_n n x^n}{\sum_m x^m} \quad \text{where } x = e^{-\hbar\omega/k_B T} \\ &= \hbar\omega \frac{0 + x + 2x^2 + 3x^3 + \dots}{1 + x + x^2 + x^3 + \dots} \\ &= \hbar\omega \frac{x(1 + 2x + 3x^2 + \dots)}{1 + x + x^2 + x^3 + \dots} \\ &= \hbar\omega \frac{x \frac{\partial}{\partial x} (1 + x + x^2 + x^3 + \dots)}{(1 + x + x^2 + x^3 + \dots)} \\ &= \hbar\omega x \frac{\frac{\partial}{\partial x} \frac{1}{1-x}}{\frac{1}{1-x}} \\ &= \hbar\omega x \frac{1}{(1-x)^2} \\ &= \hbar\omega \frac{x}{1-x} \\ &= \frac{\hbar\omega}{e^{\hbar\omega/k_B T} - 1} \end{aligned}$$

which is exactly what we were looking for! At low frequencies, it is possible to show that it reduces to $k_B T$ which is the classical result that Rayleigh derived (See Exercise ***). But at high frequencies, the exponential term takes over, and $\bar{E} \approx \hbar\omega e^{-\hbar\omega/k_B T}$ (Exercise ***) which goes to zero at high frequencies. This means that the area of the graph, and hence the energy, is finite, thus avoiding the UV catastrophe. Problem solved!

Before we continue, it is very important that you understand what Planck said here. He said that the energies were quantised, based on the frequency.

Any frequency is allowed, but given that frequency, we can have only certain energies. We *do* in fact have a constraint on the allowed frequencies - since $\omega = ck$, and only certain wavenumbers are allowed, we can have only certain ω 's. But this is a physical consequence of our system, and nothing to do with the quantisation that Planck suggested.

2.3 Photoelectric effect

But the question you ask now, is *why* is the problem solved? Planck's suggestion was that the walls of the cavity could only release energy in discrete lumps, and so the quantisation came from the interaction of light and matter.

Cue Einstein, who saw Planck's derivation and went one better (much to Planck's annoyance). He suggested that quantisation might be a *fundamental* property of light - radiation can only have certain energies (with the actual energies allowed dependent on the frequency of light) He then looked to see if he could apply this to the *photoelectric effect*.

Consider a metal plate onto which light is shone with varying frequencies of light. This light can dislodge electrons, which can then be detected and their energy measured. Now classically, we would expect that the brighter the light (meaning more energy), the faster electrons would be emitted, and the more kinetic energy they should have.

What we find instead is that below a certain frequency, *no* electrons are emitted, regardless of the amplitude (energy) of the light. Above this **threshold frequency**, electrons are emitted immediately (as opposed to a gradual increase in release rate). Furthermore, the kinetic energy of the electrons depends only on frequency, whereas their rate of increase depends, as we predicted classically, on the intensity of light (provided the frequency is above the threshold.) Quite a different result!

So Einstein argued (extending Planck's idea) that light only arrives in bundles of energy, with $E = \hbar\omega$. Then by conservation of energy, the kinetic energy of the electron must be

$$KE = \hbar\omega - W$$

where W is the **work function** of the metal, the minimum energy needed to dislodge the electron.

Therefore the *minimum* energy needed to dislodge a photon would be one for which $KE = 0$, and hence we would need a frequency of at least $\omega = \frac{\hbar}{W}$. This explains the threshold frequency phenomenon. The intensity of the light would have to be related to how many packets of light, with energy $\hbar\omega$, arrive (per second, say), and hence for higher intensities there would be more of these packets hitting electrons, and so we see a greater release rate. But because light *only* arrives in packets of $\hbar\omega$, the intensity doesn't effect how much energy the electrons get - their kinetic energy will be determined solely by ω .

2.4 Compton Effect

In the previous section we found that light only travels in discrete packets. *Why* would this be? Einstein suggested that light can also be thought of as a particle - which he called a **photon** - and this particle has an energy $\hbar\omega$. Note that he wasn't saying that light *is* a particle; he just meant that in some circumstances, light acts as a particle, and in others as a wave. What was needed, however, was some hard evidence that light could indeed be thought of as a particle.

Now it was time for a physicist named **Compton** to work some magic. Compton said that if light was also a particle, it should act like one, specifically in terms of collisions with other objects such as electrons.

Classically, if light (as a wave) hits an electron, then for some frequencies of light, the electron absorbs the light and then re-radiates it in *all* directions, but *at the same frequency* that it entered. What Compton asked was whether the same result applied if light was a particle, in which case we have a sort of quantum billiard ball game.

Firstly, we need review a little Special Relativity. We need to know the momentum of these photons, given their energy. For this, we consider Einstein's relativity formula

$$E^2 = p^2c^2 + m_0^2c^4 \quad (2.6)$$

(which would be covered in a first level Special Relativity subject.) Here E is the energy of the particle, p is its momentum and m_0 is the **rest mass** of the particle, i.e. its mass when it is motionless.

From this formula we can derive the **relativistic mass**, which depends on the velocity of a particle, and is defined so that the good old classical formulas $E = mc^2$ and $p = mv$ still hold. Inserting these into Equation 2.6 and solving, we find

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (2.7)$$

We know that light travels at the speed of light (!), which means photons must also travel at $v = c$. But this means (by Equation 2.7) that m is infinite unless $m_0 = 0$ (in which case the numerator goes to zero faster than the denominator - you can check it with basic limit theory.). An infinite m would imply an infinite energy (since $E = mc^2$) which is rubbish - their energy is $\hbar\omega$ - so we infer that the rest mass of a photon is zero.

Thus, from Equation 2.6 we find $E = pc$, or

$$p = \frac{E}{c} \quad (2.8)$$

Now, let's consider what happens if we collide a photon with an electron. It will be similar to a classical billiard ball game - the photon hits, and flies off at some angle θ , while the electron moves off at angle ϕ , having absorbed some momentum from the photon. Already this sounds different to the classical interpretation, because our photon only flies off in one direction.

Let ω_i and ω_f be the initial and final frequencies of the photon (whether they will be equal we'll soon see.) And let E_i , E_f and p_i , p_f be the initial and final energy and momentum of the electron.

Let's assume that the photon is moving horizontally (with momentum $\hbar\omega_i/c$) to the right, and hitting the electron which is stationary ² which means that $p_i = 0$ and $E_i = mc^2$, and let's take θ and ϕ relative to the horizontal. All this is shown in Figure ***. [Ed note: Figure coming soon!]

The algebra is rather messy, but try and bear with it. We won't give explicit working for every step here, because it really is just algebra. If you're enthusiastic, it would be good algebra practice to follow the steps below in more detail. We use conservation of momentum and energy (p_x , p_y and E respectively) to get enough simultaneous equations:

$$\frac{\hbar\omega_i}{c} + 0 = \frac{\hbar\omega_f}{c} \cos \theta + p_f \cos \phi \quad (p_x) \quad (2.9)$$

$$0 + 0 = \frac{\hbar\omega_f}{c} \sin \theta + p_f \sin \phi \quad (p_y) \quad (2.10)$$

$$\hbar\omega_i + mc^2 = \hbar\omega_f + \sqrt{p_f^2 c^2 + m^2 c^4} \quad (E) \quad (2.11)$$

where we use the relativistic formula for the energy of the electron.

We can rearrange Equations 2.9 and 2.10 to solve for p_f , and then square it because we want to sub into Equation 2.11)

$$p_f^2 = \frac{(\hbar\omega_i - \hbar\omega_f \cos \theta)^2}{c^2} + \frac{(\hbar\omega_f \sin \theta)^2}{c^2}$$

Next rearrange 2.11 so the square root is on one side, square it, and use our expression for p_f^2 :

$$\begin{aligned} \hbar\omega_i + mc^2 - \hbar\omega_f &= \sqrt{p_f^2 c^2 + m^2 c^4} \\ (\hbar\omega_i + mc^2 - \hbar\omega_f)^2 &= p_f^2 c^2 + m^2 c^4 \\ &= (\hbar\omega_i - \hbar\omega_f \cos \theta)^2 + (\hbar\omega_f \sin \theta)^2 + m^2 c^4 \end{aligned}$$

Now, we expand the brackets, and using the identity $\sin^2 \theta + \cos^2 \theta = 1$ we can simplify it to

$$mc^2(\omega_i + \omega_f) = \hbar\omega_i\omega_f(1 - \cos \theta)$$

We'll actually take this one step further, and change from angular frequency to wavelength by $\omega = \frac{2\pi c}{\lambda}$. Thus, finally, we have

$$\lambda_f - \lambda_i = \frac{h}{mc}(1 - \cos \theta) \quad (2.12)$$

(Note the use of h instead of \hbar)

Incidentally, the $\frac{h}{mc}$ out the front is called the **Compton wavelength** of an electron, and it is of interest in particle physics. Compton then took his results

²We can always change to the frame of the electron to make it stationary!

to the lab, and attempted to observe this shift in wavelength. Although the shift is very small and only noticeable in high energy x-rays, where the wavelength is comparable to the shift, it was indeed there, which provided solid evidence for photons. Compton also found that some photons bounced off without any shift in wavelength. This resulted from photons hitting the nucleus of an atom, which is so massive compared to the photon that it virtually unaffected by the impact - its Compton wavelength made the frequency shift imperceptible!

2.5 de Broglie Wavelength

The next natural step was to wonder whether the reverse was true - if light (classically a wave) sometimes acts as a particle, might not a particle sometimes act as a wave? Now for photons, we know that

$$\begin{aligned}\lambda &= \frac{c}{f} = \frac{hc}{hf} \\ &= \frac{hc}{E} \\ &= \frac{h}{p} \quad \text{by } E = pc\end{aligned}$$

The physicist *de Broglie* suggested that perhaps this held, not just for photons, but for any particle. This became known as the *de Broglie wavelength* of a particle, and is given by

$$\lambda = \frac{h}{p} \tag{2.13}$$

So, is there any evidence to support this? We return to first year optics, and *single slit diffraction*. In this well worn experiment, light waves are fired at a screen into which a small slit of width d is cut. Then the light is projected onto a second screen some distance away ³ (see Figure ***). This experiment can be found in any introductory optics text, and more popular physics books than you could imagine. The pattern of light and dark bands we see on the screen are called a *diffraction pattern*. They result because the peaks and troughs of the wave interfere, sometimes constructively (the bright bands) and other times destructively (the dark bands). The *angular half width*, i.e. the angular distance θ from the center to the first dark band is given by

$$\theta \approx \frac{\lambda}{d} \tag{2.14}$$

What we can do now is repeat this experiment, not with light waves, but with a beam of electrons. The actual experiment is done by diffracting the electrons off a crystal (similar to X-ray crystallography) - have a think about why we must use a crystal instead of a slit - but the effect is still the same. We

³This distance must be significantly larger than the slit, etc - a metre is usually sufficient

observe a diffraction pattern for where the electrons hit the screen, something which can't be explained if they are particles. This result could only be obtained if electrons, classically particles, can also be thought of as waves.

2.6 Bohr and the atom

These last two experiments provided strong evidence of *wave-particle duality*, and gave the correct results for some puzzling experiments. But perhaps the best example of the early applications of quantum theory deals with electrons around the atom.

In 1911, **Rutherford** (and his students, **Geiger** and **Marsden**, who rarely get a mention) developed a model which described atoms as a heavy, positively charged nucleus with electrons in circular orbits around it, much like planets around the sun. This was verified (and in fact inferred) from their experiment of firing alpha particles at gold foil, and observing their deflection. Some particles bounced right back at 180 degrees, which could only have happened if they had hit something heavy and solid - the nucleus.

So the theory seemed to make sense. However, it had a serious flaw - the electrons orbiting the nucleus are accelerating charges, and hence classically should emit radiation. This loss of energy would cause them to spiral into the nucleus, which means in about a tenth of a nanosecond, every atom in the universe would collapse. Clearly we're still going strong after several billion years, so again, we have theory that contradicts experiment.

In 1913, **Bohr** suggested that maybe quantisation wasn't just restricted to light - perhaps it could also apply to *orbital angular momentum*. Bohr wondered if perhaps electrons couldn't orbit at just any radius around the nucleus. He suggested that they could only exist with a radius r_n such that the angular momentum was an integer multiple of \hbar , i.e.

$$\|p \times r_n\| = n\hbar \quad (2.15)$$

(which you may be more familiar with as $L = mvr \sin \theta$)

Using the classical equation of motion for an electron in orbit (equating centripetal force with electric force, straight from introductory electromag), we have

$$m \frac{v^2}{r} = \frac{e^2}{4\pi\epsilon_0 r^2}$$

so the electron has momentum

$$p = \left[\frac{me^2}{4\pi\epsilon_0 r} \right]^{\frac{1}{2}}$$

Note also that because the orbit is circular, the electron always moves perpendicular to its position vector, so $\|p \times r_n\| = pr_n \sin 90^\circ = pr_n$, and hence our quantisation of angular momentum becomes

$$\left[\frac{me^2}{4\pi\epsilon_0 r} \right]^{\frac{1}{2}} \times r_n = n\hbar$$

Solving for the permissible radii (the *Bohr orbits*), we find

$$r_n = 4\pi\epsilon_0 \frac{n^2\hbar^2}{me^2} \quad (2.16)$$

Bohr postulated that, in addition to the radii being quantised, electron orbits are stable - the electron does not emit radiation in the classical way, and hence does not spiral in. Instead, it can only “jump” up or down to another Bohr orbit, absorbing or releasing a photon with energy equal to the energy difference between the levels.

We can calculate these energies, by summing the kinetic and potential energy of the electron (the latter we will quote straight from EM theory⁴). For an orbit of radius r (not necessarily a Bohr radius) we have

$$\begin{aligned} E &= KE + PE \\ E &= \left(\frac{p^2}{2m}\right) + \left(-\frac{e^2}{4\pi\epsilon_0 r}\right) \\ &= \frac{1}{2m} \frac{me^2}{4\pi\epsilon_0 r} - \frac{e^2}{4\pi\epsilon_0 r} \\ &= -\frac{e^2}{8\pi\epsilon_0 r} \end{aligned}$$

Then, substituting in r_n for r , we find the energy of the n th orbit is

$$E_n = -\frac{e^2}{8\pi\epsilon_0 r_n} = -\frac{me^4}{32\pi^2\epsilon_0^2\hbar^2} \times \frac{1}{n^2} \quad (2.17)$$

The difference in energy between two levels m and n , and hence the energy of the photon emitted/absorbed during a jump between those levels, is given by

$$\hbar\omega = E_m - E_n = \frac{me^4}{32\pi^2\epsilon_0^2\hbar^2} \left(\frac{1}{n^2} - \frac{1}{m^2}\right) \quad (2.18)$$

This may not exactly seem earth shattering at first, until you realise that Equation 2.17 is Balmer’s formula for the emission spectrum of hydrogen! When one observes the light emitted by hydrogen atoms, one does not find all energies - instead, only certain frequencies are observed - exactly those given by ω in the above equation! Balmer developed this formula as a fit to the data, without any physical backing for it. (His motivation for discovery is an interesting one. Balmer was heavily into numerology, and at the time taught calligraphy and mathematics at a high school for girls. One day he complained to a friend that he had “run out of things to do”. His friend suggested “Well, you are interested in numbers, why don’t you see what you can make of this set of numbers that come from the spectrum of hydrogen?” And he did.)

⁴If you’re desperate, you can derive it by from $dW = F \cdot dr$, then integrate the force given earlier.

By finding a logical, physical derivation, Bohr pretty nearly clinched the case for quantisation, and help to make quantum a mainstream theory. Although it made some astoundingly accurate predictions, it also had some flaws, for example in applying Bohr theory to more complex atoms, which met mixed success. And there was still no real justification for stopping electrons from radiating classically, nor even for *why* we should have quantisation of angular momentum. Despite this, extensions to the theory were tacked on as needed, and the theory proved workable. This version of quantum mechanics is now called the ***Old Quantum Theory***. In the next chapter, we'll go back to basics and have a go at developing a fully fledged quantum mechanical theory from the bottom up.

2.7 Chapter Summary

- Rayleigh-Jeans discovered that a classical approach to the energy density in a cavity led to a nonsense result, called the Ultraviolet Catastrophe (Section 2.1)
- Planck resolved this dilemma, by quantizing the allowed energies of light in the cavity to integer multiples of $\hbar\omega$, introducing in the process h , known as Planck's constant (Section 2.2).
- Einstein generalised this quantisation to light anywhere, and verified that this was consistent with the Photoelectric Effect, a previously unexplained experiment. He suggested that light could also be considered a particle, called a photon. (Section 2.3)
- Compton then showed that if light was a photon, there should be a shift in the wavelength of light after a collision with an electron. He then observed this successfully in the laboratory. (Section 2.4)
- de Broglie then suggested that if waves could act as particles, then particles might be able to act as waves. He proposed what is now known as the de Broglie wavelength for particles

$$\lambda = \frac{h}{p}$$

Experiments have shown that electrons do diffract just like waves of the de Broglie wavelength.

- Bohr applied Quantum Theory to the atom, suggesting that the angular momentum be quantised to multiples of \hbar , and then calculated the permissible orbit radii. He successfully reproduced Balmer's formula for the hydrogen lines

$$E_m - E_n = \frac{me^4}{32\pi^2\epsilon_0^2\hbar^2} \left(\frac{1}{n^2} - \frac{1}{m^2} \right)$$

2.8 Exercises

1. Given Planck's formula for the energy density,

$$u(\omega, T) = \frac{\omega^2}{\pi^2 c^3} \frac{\hbar \omega}{e^{\hbar \omega / k_B T} - 1},$$

find the total energy in a cavity $U(T)$. You may find

$$\int_0^\infty \frac{x^3}{e^x - 1} dx = \frac{\pi^4}{15}$$

a helpful rule.

Calculate the total energy at temperatures 1000K, 500K and 298K (room temperature).

2. Consider making a hole of unit area in the box. The energy should then be emitted from the box at the speed c . However, only half the energy (radiation) in the cavity is moving towards the opening, and we reduce that by half again to account for radiation not being at normal incidence to the hole. Thus we find that the power emitted is

$$P(T) = \frac{c}{4} U(T) = c \frac{\pi^2}{60} \frac{(k_B T)^4}{(c\hbar)^3} = \sigma T^4 \text{ W m}^{-2}$$

Find the Power radiated, and hence the energy emitted from the box at each of the temperatures in the previous question. Comment on your results.

Chapter 3

Wavefunctions

3.1 Introducing the wavefunction

Let's again think about light waves and single slit diffraction (see Figure ***). A graph of *intensity* vs. position on the screen would look like Figure ***. The intensity is a measure of how bright the light is, and from wave theory we find it is proportional to the square of the amplitude of the electromagnetic field. (So if the EM field has amplitude -0.5, the intensity is 0.25) We could now make a function $\Psi(x)$ which gives the amplitude of light along the screen, and hence $|\Psi(x)|^2$ would give the intensity. [Show Figures *** here] Note that $\Psi(x)$ could be negative, but the intensity would still be positive.

If, however, we think of light as being made up of discrete particles, then this interpretation of intensity, based on waves, doesn't have any physical meaning – photons don't have amplitudes that can be split. Instead, the intensity would have to be related to the *number* of photons hitting each point per second; a spot receiving 50 photons per second would be half as bright as a spot with 100 hits per second.

Then as each photon goes through the slit, in order to explain the banded pattern, there must be a *probability distribution* for where the photon hits the screen – some photons fly off to the left, some to the right, and high probability spots would have more photons hit, and would hence be brighter than the low probability spots. This is an important result – if we have discrete particles, the intensity at each point is directly related to the probability of a photon hitting it.

Therefore a plot of probability against position along the screen would have to look qualitatively the same as a plot of intensity, so $|\Psi(x)|^2$ can also be interpreted as the probability of finding a particle around x . But what then of the function $\Psi(x)$? For light waves, $|\Psi(x)|^2$ gave the intensity, and $\Psi(x)$ gave the amplitude, which was related to intensity. Logically, therefore, the function $\Psi(x)$ – which we call the *wave function* – must then be some sort of *probability amplitude* – related to probability, but possibly being negative,

or even complex. The absolute square, however, will always be positive and real, and we interpret that as the probability of a photon being found around x . (This is just saying the amplitude of light can be negative, but the square of it, the intensity, must always be positive.) Read this paragraph again – make sure you follow what we’ve said here.

Because other particles (like electrons, protons, etc) also act like waves (See Section ??), and give similar diffraction patterns, this concept of a wave function should apply equally well to them, or indeed to any body. This idea of the wave function being the probability amplitude is the interpretation first suggested by Born, and forms a major part of quantum physics.

3.2 Fundamentals of Wave Theory

The wave function introduced above describes the probability of a diffracted particle being found somewhere along the screen. We now extend this idea to a generalised wave function that completely describes a particle anywhere in space, and this forms the first of what are known as the *Postulates of Quantum Mechanics*

Postulate 1.

*In quantum mechanics, the dynamical state of a particle is described by a complex number field $\Psi(\mathbf{r}, t)$, called the wave function. The generally accepted physical interpretation of $\Psi(\mathbf{r}, t)$, due to Born, is that it is the **probability amplitude (density)** such that*

$$\rho(\mathbf{r}, t) = |\Psi(\mathbf{r}, t)|^2 = \Psi^*(\mathbf{r}, t)\Psi(\mathbf{r}, t) \quad (3.1)$$

is the probability per unit volume of finding the particle at position \mathbf{r} at time t .

Put a little more simply, the wave function $\Psi(\mathbf{r}, t)$ is a function of x, y, z and t (so is defined at every point in space and time) and can take any complex value. If we take the absolute square of one of these numbers, we arrive at the **probability density** at that point in space. Note that this is slightly different to what we suggested before. $|\Psi(\mathbf{r}, t)|^2$ is probability *per unit volume* (see Section 1.2) so we must multiply it by the volume we’re interested in before we have an actual measure of probability (in fact, usually we integrate as we will see in just a moment.)

Notice how the wave function representation is fundamentally different from classical mechanics. Classically, if we were to take a particle and measure its momentum, then if we always prepared it the same way, we would always get the same result. But in quantum mechanics, if we try and measure the position (or momentum) of a particle, with the particle being prepared in the same state before each measurement ¹, we could potentially obtain different values

¹We will see in Section ?? why this is important. For now, just understand that “preparing the particle in the same state” simply means resetting it after each measurement, to make sure nothing has changed.

each time. The particle has, in some sense, many values for its momentum or position *at the same time*. We will see soon that this is an example of the **Superposition Principle**.

Now would be a good time to re-view Section 1.2 if you haven't already done so. To find the actual probability of finding the particle in some position, we must integrate over the range we are interested in. Thus, to find the probability of finding the particle between $\mathbf{r} = (1m, 1m, 1m)$ and $\mathbf{r} = (1.01m, 1.01m, 1.01m)$ (so in a cube of side $1cm$ with its corner at $(1, 1, 1)$) we would say

$$P = \int_1^{1.01} \int_1^{1.01} \int_1^{1.01} \rho(\mathbf{r}, t) dx dy dz$$

We would hope that the total probability of finding a particle anywhere in space would be equal to unity (there's a 100% chance of finding it *somewhere!*). Therefore, the total area under the probability density graph should be 1. Mathematically, we write this as

$$\int \rho(\mathbf{r}, t) d\mathbf{r}$$

where the integration is over all space.

To sum up, in quantum mechanics it is not possible (in general) to obtain exact information about a particle. Instead, we can only obtain information about the *probability* of a particle being at different positions, and this information is stored in the wave function $\Psi(\mathbf{r}, t)$ such that $|\Psi(\mathbf{r}, t)|^2$ gives the probability density. This applies equally well to momentum, although this may not be obvious yet; the next chapter will talk more about it.

3.3 Free particle momentum states

So, we know that we can describe a particle by this wave function. But what does the wave function actually look like? Recall that particles can also be thought of as waves, with Broglie wavelength

$$\lambda = \frac{h}{p} \tag{3.2}$$

We also have Einstein's relation that

$$E = hf \Rightarrow f = \frac{E}{h} \tag{3.3}$$

Therefore, we have both a wavelength and a frequency, and these are the same everywhere in time and space. A wave with these properties is called a **plane wave** - it oscillates uniformly, and is about the simplest wave you can make. The formula given in Section ?? $A \sin(\mathbf{k} \cdot \mathbf{r} - \omega t)$ is an example of a plane wave, but one that oscillates only through the real numbers. Because we're saying that the wave function is complex, we need a plane wave that has an oscillating imaginary part as well. Hence,

Postulate 2.

The wave function for a free particle, i.e., a particle not acted on by any forces, with a well defined momentum \mathbf{p} is given by

$$\Psi(\mathbf{r}, t) = Ae^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \quad (3.4)$$

where $\mathbf{k} = \frac{\mathbf{p}}{\hbar}$ (the de Broglie relation) and $\omega = \frac{E}{\hbar} = \frac{p^2}{2m\hbar}$ (the Einstein relation).

Note that we are now using $\mathbf{k} = \frac{\lambda}{2\pi}$ and $\omega = 2\pi f$, where \mathbf{k} is the *wave number* and ω is the *angular frequency*. This wave is equivalent to

$$\Psi(\mathbf{r}, t) = A \cos(\mathbf{k} \cdot \mathbf{r} - \omega t) + iA \sin(\mathbf{k} \cdot \mathbf{r} - \omega t)$$

which you can see oscillates in both the real and imaginary planes, as needed.

This state is one for which the momentum is known to be exactly \mathbf{p} . What can we say about the particle's position? The probability density is

$$\begin{aligned} \rho(\mathbf{r}, t) &= |\Psi(\mathbf{r}, t)|^2 = \Psi^*(\mathbf{r}, t)\Psi(\mathbf{r}, t) \\ &= A^* A e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t)} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \\ &= |A|^2 \end{aligned} \quad (3.5)$$

which does not depend on either time or position - hence there is equal probability of finding the particle anywhere in space! This result - that a particle with a definite momentum has its position completely unspecified - is due to the **Heisenberg Uncertainty Principle** which we will learn more about later in the course.

Perhaps at this point you're beginning to have doubts about the reality of quantum physics, and this is understandable - it's a very counter-intuitive theory! The very idea of a particle with a known momentum having the same probability of being everywhere in space is very strange, and you may have noticed that the area under the probability graph is not unity - it is infinite! Bear with it, however, and eventually it will make a little more sense.

Example 3.3.1.

Consider another wave function, known as a 1-D **Gaussian**, one which we will use frequently for examples. It represents a particle whose position is known to be approximately $x = 0$, but may be found a little to either side of that. It is of great importance to applied quantum mechanics; in particular, it is the wave function of a particle bouncing on a spring. While this may not sound particularly useful, there are many situations where we can model a particle as such - atoms in crystals as they vibrate is one example. (In practice, the wave function would be much more complicated, but the Gaussian provides a basis to work from.)

Its wave function (without time dependence) is

$$\Psi(x) = e^{-\beta x^2/2} \quad (3.6)$$

Calculate the probability density.

Answer.

$$\begin{aligned}\rho(x) &= \Psi^*(x)\Psi(x) \\ &= e^{-\beta x^2}\end{aligned}$$

Notice the probability density looks very much like the wave function - they are both Gaussian distributions. This will not always be the case, that the two match up so closely. In particular, in a pure momentum state the wave function looks nothing like the probability density at all!

3.4 Superposition

Clearly, a pure momentum state is not a very “real life” situation; in fact it is one of the best examples of the differences between quantum and classical mechanics. So how can we make a state that has its position better specified? Recall that earlier we mentioned that a particle can be thought about as having more than one value for its momentum. It may then be thought of as being in two momentum states *simultaneously*, which we could write like:

$$\Psi(\mathbf{r}, t) = \Psi_1(\mathbf{r}, t) + \Psi_2(\mathbf{r}, t) \quad (3.7)$$

where

$$\Psi_1(\mathbf{r}, t) = A_1 e^{i(\mathbf{k}_1 \cdot \mathbf{r} - \omega t)}$$

and

$$\Psi_2(\mathbf{r}, t) = A_2 e^{i(\mathbf{k}_2 \cdot \mathbf{r} - \omega t)}$$

Here, we have combined two states with wave numbers \mathbf{k}_1 and \mathbf{k}_2 to give a new state that has *both* values of momentum². This is an example of the Superposition Principle, which says that if $\Psi_1(\mathbf{r}, t)$ and $\Psi_2(\mathbf{r}, t)$ are both valid states, then

$$\Psi(\mathbf{r}, t) = a\Psi_1(\mathbf{r}, t) + b\Psi_2(\mathbf{r}, t) \quad (3.8)$$

must also be a valid state, where a and b are complex numbers.

Perhaps it seems a little odd to add two wave functions together like that, and just assume that it represents a particle with two values of momentum, and you’d be right. Certainly it would make no sense in classical physics! But even the idea that this is the correct way to represent it mathematically may be a little odd. However, the superposition principle is supported by diffraction experiments (with particles) where in order to explain the results it is necessary to superimpose wave functions on the screen, just the way we do with light waves. Or, even simpler, because it holds for light waves, by our analogy between waves and the wave function, there’s a good chance it holds for quantum.

²Be aware that we sometimes use the terms momentum and wavenumber interchangeably, since they are just scaled by \hbar . Confusing, but that’s the way it is!

In terms of quantum theory, if superposition holds, then the wave function must be linear, and superposition should hold for all states, not just free particle states. (If this doesn't mean much to you, don't worry - just accept you can add wave functions.)

We could now add more states to this superimposed state, and still have a valid state, one in which we can no longer say for certain what the momentum of the particle is. In fact, the most general state would be one where we add *all* the momentum states together, each with some scaling co-efficient (like a and b in Equation 3.8 above), some of which could be zero. But because there are an infinite number of momentum states, and because momentum is continuous, we have to write this sum as an integration, over every possible state. (See Section 1.2) We write this general state as:

$$\Psi(\mathbf{r}, t) = \int_{\text{all } \mathbf{k}\text{-space}} A(\mathbf{k}) e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)} d\mathbf{k} \quad (3.9)$$

where we now use $A(\mathbf{k})$ as our coefficients.

If you have done some linear algebra, you might notice this is a similar idea to creating a general vector by summing the basis vectors, only here our basis "vectors" are the pure momentum states. In fact, this is exactly what we are doing, and it plays an important role that we look at later.

This new state is much more complex than the pure momentum states we made before. In fact, to analyse this state, and find the probability of different values of momentum and position requires the introduction of some more advanced mathematics - Fourier transforms and δ -functions - and this is what we cover in the next chapter.

3.5 Chapter Summary

- Particles sometimes act as waves, and visa versa. The wavelength of a particle is given by the de Broglie wavelength $\lambda = \frac{h}{\mathbf{p}}$ and its frequency $f = \frac{E}{h}$
- In quantum mechanics, the *wave function* is a complex field, and contains all the dynamical information about a particle. (Section 3.2)
- The wave function $\Psi(\mathbf{r}, t)$ is the probability amplitude of finding the particle at a certain position in space, and its absolute square $|\Psi(\mathbf{r}, t)|^2$ gives the probability density, i.e. probability per unit volume. (Section 3.2) This is the interpretation first suggested by Born.
- For a particle with well defined momentum, the wave function looks like

$$\Psi(\mathbf{r}, t) = A e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)} \quad (3.10)$$

where we define \mathbf{k} and ω by $\mathbf{k} = \frac{2\pi}{\lambda} = \frac{2\pi\mathbf{p}}{h} = \mathbf{p}/\hbar$ and $\omega = 2\pi f = \frac{2\pi E}{h} = \frac{E}{\hbar}$. (Section 3.3)

- Quantum theory is linear, and hence the superposition principle holds. Therefore, we can add pure momentum states together to generate new states, with the momentum less precisely specified. (Section 3.4)

You should know:

- the definition of the wave function, and the reasoning behind it
- what the de Broglie wavelength is, and its definition
- the form of the wave function for a free particle

3.6 Exercises

Chapter 4

The Momentum Representation

This chapter is rather more mathematical than the last. We need to introduce two concepts – the Dirac Delta Function and Fourier Transforms – which you may not have seen before.

4.1 The Dirac Delta Function

For the moment we are going to simplify our maths by just considering a 1D particle – it can only move along the x -axis – and we will assume that the wave function does not change with time. Hence $\Psi(\mathbf{r}, t)$ simplifies to $\Psi(x)$ and $|\Psi(x)|^2$ is the probability density as a function of position. As we said earlier, the probability of finding a particle *anywhere* in space must equal 1 therefore the area under $|\Psi(x)|^2$, which will be the total probability, must equal 1 also.

We've already discussed what the wave function looks like when the momentum is known exactly - it's a plane wave. But what about when the *position* is known exactly? Let's say we know the position to be precisely x_0 . Anywhere else in space, the wave function has to be zero – there's no probability of it being there. Therefore the wave function would have to look like a single spike at x_0 , and be zero everywhere else. But there's a problem - how could a spike, with zero width, have an area equal to 1?!

Dirac solved this problem by introducing what is now known as the ***Dirac delta function***. In truth, it's not strictly a function, but it acts just like it for everything we do.

The delta function is denoted $\delta(x - x')$, and gives a spike at exactly x' (e.g. $\delta(x - 2)$ gives a spike at 2, but is zero elsewhere) but still has unit area. A δ -function must satisfy the following properties:

1. $\delta(x - x') = 0$ for $x \neq x'$

That is, everywhere except at x' the function is equal to zero.

2. As $x \rightarrow x'$, $\delta(x - x') \rightarrow \infty$ in such a way that

$$\int_{-\infty}^{\infty} \delta(x - x') dx = 1 \quad (4.1)$$

That is, at x' , the δ -function shoots up to infinity, but in such a way that the area of the whole function is unity. At the moment, we don't know how we could make such a function; all we are doing is listing the properties we need it to have.

Additionally,

$$\int_R \delta(x - x') = 1 \quad \text{if } x' \in R \quad (4.2)$$

that is, when integrating over a region R , the integral is 1 if R contains x' and zero otherwise.

3. For a function $f(x)$ that is well behaved¹ near x' , we find

$$\int_{-\infty}^{\infty} f(x) \delta(x - x') dx = f(x') \quad (4.3)$$

which is called the **filtering property** of the δ -function.

4. Again if f is well behaved, we have

$$\int_{-\infty}^{\infty} f(x) \delta(x - x') dx = -f'(x') \quad (4.4)$$

The last two may at first seem a little unintuitive; but the Filtering property, at least, can be loosely explained as follows: The δ -function is zero everywhere except at x' , so $\delta(x - x')f(x)$ would be zero everywhere except at x' as well (zero times anything is still zero!) So, when integrating this over x , the only value that actually contributes to the integral is $f(x')$. Of course, why this doesn't get scaled to infinity is a little harder to understand, but this perhaps gives you a general idea. The rigorous proof of the last two properties is left as an Exercise.

Now that we know what properties the δ -function must satisfy, what does it look like mathematically - i.e. how can we make one? What sort of graph could possibly satisfy our requirements? First, we need some new definitions.

An **even function** is one where $f(x) = f(-x)$. For example, $\cos(x)$ is an even function. And an **odd function** is one where $f(x) = -f(-x)$ (for example, $\sin(x)$). Even and odd functions satisfy the following properties:

1. If $f(x)$ and $g(x)$ are both even or both odd functions, then $f(x)g(x)$ is an even function.
2. If $f(x)$ is even and $g(x)$ is odd, then $f(x)g(x)$ is odd.

¹“Well behaved” just means it doesn't do anything strange - it's not discontinuous, it doesn't change direction suddenly, and so on. Everything we do will involve well behaved functions!

3. If $f(x)$ is odd, then

$$\int_{-\infty}^{\infty} f(x)dx = 0 \quad (4.5)$$

This is because the area on either side will cancel out, leaving zero (on one side it is positive, on the other negative.)

4. If $f(x)$ is even, then

$$\int_{-\infty}^{\infty} f(x)dx = 0 \quad (4.6)$$

Because the function is symmetrical around the y -axis, we need only find the area on one side, and double it.

If any of those properties don't make sense, try drawing some graphs of sin and cos functions, and check the properties against them.

Example 4.1.1.

$y = x^2$ and $y = x^4$ are both even functions, while $y = x$ and $y = 1/x$ are both odd functions. Because an even function times an odd function is odd, we would hope that $x^2 \times 1/x = x$ would be odd, and it is.

Example 4.1.2.

$y = e^x$ is neither even nor odd, because $e^{-x} \neq e^x$ and $e^{-x} \neq -e^x$.

$y = e^{|x|}$, however, is even, because the sign of x doesn't matter. Note that multiplying an even function but a function that is neither even nor odd is still neither even nor odd.

In a leap of intuition, we're going to suggest that

$$f(x - x') = \int_{-\infty}^{\infty} e^{ik(x-x')} dx \quad (4.7)$$

might be a good place to start in finding a functions that satisfies the requirements of a δ -function. This probably won't be immediately obvious, so let's look at more closely.

When $x = x'$ the exponential becomes unity, and so the integral reduces to

$$\begin{aligned} f(x' - x') &= \int_{-\infty}^{\infty} e^{ik(x'-x')} dk \\ &= \int_{-\infty}^{\infty} dk \\ &\rightarrow \infty \end{aligned}$$

as required by a δ -function. When $x \neq x'$ it's a lot harder to see what the integral does, so what we do is to introduce a "converging factor" $e^{-w|k|}$ into

the equation. As $w \rightarrow 0$, this goes to 1, so we can write:

$$\begin{aligned} \int_{-\infty}^{\infty} e^{ik(x-x')} dk &= \lim_{w \rightarrow 0} e^{-w|k|} \int_{-\infty}^{\infty} e^{ik(x-x')} dk \\ &= \lim_{w \rightarrow 0} \int_{-\infty}^{\infty} e^{-w|k|} e^{ik(x-x')} dk \end{aligned} \quad (4.8)$$

All we've done is to stick a $\lim_{w \rightarrow 0} e^{-w|k|}$ out the front, which changes nothing in the equation, because it equals 1, and w doesn't appear anywhere else. So why bother? Here's the gimmick - for the moment, we forget about the limit. We'll evaluate the RHS, and then take the limit at the end - all perfectly legitimate, because as we said, w appears nowhere else.

So, dropping the limit, we can evaluate the RHS by splitting up the exponential into trig functions, and then use the properties of odd and even functions.

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-w|k|} e^{ik(x-x')} dk &= \int_{-\infty}^{\infty} e^{-w|k|} \cos k(x-x') dk + i \int_{-\infty}^{\infty} e^{-w|k|} \sin k(x-x') dk \\ &= 2 \int_0^{\infty} e^{-w|k|} \cos k(x-x') dk \\ &= 2 \operatorname{Re} \left(\int_0^{\infty} e^{ik(x-x'-iw)} dk \right) \\ &= 2 \operatorname{Re} \left(\frac{-1}{i(x-x'-iw)} \right) \\ &= \frac{2w}{(x-x')^2 + w^2} \end{aligned} \quad (4.9)$$

The very last step (if you didn't follow it) can be obtained by multiplying by the complex conjugate, thus getting a real denominator. Then we just discard the imaginary part of the numerator.

By looking at its derivatives, we can show that this function has a maximum at $x = x'$ with a height of $2/w$, then drops off to either side, tapering off to zero. We can get an idea of the width of this peak by finding when it drops to half its maximum height, and we find this to be when $x = x' \pm w$, so its approximate width is $2w$.

What about its area?

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{2w}{(x-x')^2 + w^2} dx &= 2 \int_{-\infty}^{\infty} \frac{du}{1+u^2}, \text{ with } u = \frac{x-x'}{w} \\ &= 2 \tan^{-1} u \Big|_{-\infty}^{\infty} \\ &= 2\pi \end{aligned} \quad (4.10)$$

So its area is always 2π , no matter what we choose for w . So, as w approaches zero (what we're interested in!) this peak (of height $2/w$ and width $2w$) gets squeezed higher but narrower, so keeping the same area. Eventually (i.e. in the limit as w goes to zero), it has zero width and infinite height - but still has area

2π ! This is almost exactly like a genuine δ -function, except we need unit area, so we simply divide by 2π .

Therefore, we suggest that (putting the limit back)

$$\delta(x - x') = \lim_{w \rightarrow 0} \frac{1}{\pi} \frac{2w}{(x - x')^2 + w^2} \quad (4.11)$$

$$= (2\pi)^{-1} \int_{-\infty}^{\infty} e^{ik(x-x')} dk \quad (4.12)$$

are two possible representations of the δ -function, in fact equal to each other. There are more possible ways to make a δ -function, but these are all we need for this course.)

We've already proven that the area of Equation 4.11 (and hence Equation 4.12) is unity. Previously, we also showed that for $x = x'$ our function goes to infinity, but let's prove that again, and Property 1, with our new expression in Equation 4.11. First, if $x = x'$:

$$\begin{aligned} \delta(x - x') &= \lim_{w \rightarrow 0} \frac{2w}{w^2} \\ &= \lim_{w \rightarrow 0} \frac{2}{w} \\ &\rightarrow \infty, \end{aligned}$$

as required.

If $x \neq x'$:

$$\begin{aligned} \delta(x - x') &= \lim_{w \rightarrow 0} \frac{2w}{(x - x')^2 + w^2} \\ &= \frac{0}{(x - x')^2} \quad (\text{No problems with } w = 0) \\ &= 0 \end{aligned}$$

Hence, our representations satisfy the first two properties. As mention earlier, it is left as an exercise at the end of the chapter to show that the final two also hold.

Exercise 4.1.1.

Confirm that the function give above by

$$\frac{2w}{(x - x')^2 + w^2} \quad (4.13)$$

does indeed have a peak at $x = x'$ (and only at $x = x'$) with height $2/w$, half-height width of $2w$ and show that it vanishes at $x = \pm\infty$.

Notice two important things about our δ -functions. Firstly, because of the squared term $(x - x')^2$ in Equation 4.11, the δ -function is symmetric around x' . This means that

$$\delta(x - x') = \delta(x' - x)$$

(check if you don't believe this!) We will use this identity in the next section.

Secondly, its complex conjugate $\delta^*(x - x')$ is just the δ -function again, because

$$\begin{aligned}\delta^*(x - x') &= \left(\int_{-\infty}^{\infty} e^{ik(x-x')} dk \right)^* \\ &= \int_{-\infty}^{\infty} e^{-ik(x-x')} dk \\ &= \delta(x' - x) \\ &= \delta(x - x') \quad (\text{From above})\end{aligned}$$

as required.

4.2 Fourier transforms

We're now going to introduce a very powerful mathematical concept, known as the **Fourier Transform**. It's a little hard to explain, so let's start off with an example.

Sound waves are often described by their amplitude, and how it changes with time – square waves and sine waves are common examples that you may have seen before. A sine wave is particularly important, because it corresponds to a pure note – a sound made of just one frequency. We could make a more complex sound by adding two, three or many pure sounds together. In fact, any sound, even the human voice, is just a combination of pure notes.

Enter the Fourier Transform – given the amplitude of a wave (as a function of t , say) it tells us what frequencies, and in what proportions, make up that wave. This applies not only to sound waves – we can show that any function can be rewritten as the sum of sine curves (but perhaps both complex and imaginary ones, if the function is complex). Exactly what “frequency” means depends on the situation; a function of time is changed into a function of $\omega = \frac{2\pi}{T}$ (where T is the period of the component waves), while the transform of a function of position is a function of wave number $k = \frac{2\pi}{\lambda}$ (with λ being the wavelength of the waves.) Notice the similarity between these cases – wave number is just the space analogy of angular frequency.

It's important to note that whether we know the amplitude as a function of time, or the strengths of the frequencies that make up the wave, it's enough to completely describe the wave. Therefore, one description is as good as another, and using the Fourier Transform (and its inverse, which we will see in a moment) we can get from one to the other without losing any information.

A full description of the Fourier Transform and its applications would require a Guide of its own, so we will leave our general discussion here. We'll now derive the Fourier transform, and show how it applies to quantum. Already, you might be picking up on its connection – a sound wave is made up of pure, oscillating notes, and its transform gives the amplitude of each of those waves. A wave function is made up of pure momentum states (which are oscillating

plane waves), so what does this suggest about its transform? We will come back to this shortly.

Firstly, using Property 3 of a δ -function, we can "decompose" a function as follows:

$$\begin{aligned}
 f(x') &= \int_{-\infty}^{\infty} f(x)\delta(x-x')dx && \text{(Property 3)} \\
 &= \int_{-\infty}^{\infty} f(x) \left[(2\pi)^{-1} \int_{-\infty}^{\infty} e^{ik(x'-x)} dk \right] dx && \text{(Swapping } x \text{ and } x') \\
 &= (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} e^{ikx'} \left[(2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(x)e^{-ikx} dx \right] dk \\
 &= (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} e^{ikx'} F(k) dk && (4.14)
 \end{aligned}$$

where we have defined

$$F(k) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} e^{-ikx} f(x) dx \quad (4.15)$$

$F(k)$ is called the **Fourier Transform** of the function $f(x)$, and if we write Equation 4.14 with x instead of x' , then we have

$$f(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} e^{ikx} F(k) dk \quad (4.16)$$

called the **Fourier Inversion Theorem**

It's probably not clear yet that $F(k)$ gives the amplitude of each wave making up $f(x)$, but in a moment we will apply this theory to quantum mechanics, and show that it does indeed work. First, however, we need to find out how to take the Fourier transform in three dimensions. Let's extend our idea of a δ -function to three dimensions. Such a function, written $\delta(\mathbf{r} - \mathbf{r}')$ would need to satisfy the corresponding 3-D δ -function properties:

1.
$$\delta(\mathbf{r} - \mathbf{r}') = 0 \text{ for } \mathbf{r} \neq \mathbf{r}' \quad (4.17)$$

2.
$$\delta(\mathbf{r} - \mathbf{r}') \rightarrow \infty \text{ as } \mathbf{r} \rightarrow \mathbf{r}' \quad (4.18)$$

such that

$$\int \delta(\mathbf{r} - \mathbf{r}') d\mathbf{r} = 1 \quad (4.19)$$

3.
$$\int f(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}') d\mathbf{r} = f(\mathbf{r}') \quad (4.20)$$

4.
$$\int f(\mathbf{r})\delta'(\mathbf{r} - \mathbf{r}') d\mathbf{r} = f'(\mathbf{r}') \quad (4.21)$$

These properties are almost identical to the one dimensional ones given previously, except now we consider \mathbf{r} instead of just x , and we integrate over all of the 3D space.

We can make a three dimensional δ -function from our one dimensional representation as follows:

$$\begin{aligned}
 \delta(\mathbf{r} - \mathbf{r}') &= \delta(x - x')\delta(y - y')\delta(z - z') \\
 &= (2\pi)^{-1} \int_{-\infty}^{\infty} e^{ik(x-x')} dk_x (2\pi)^{-1} \int_{-\infty}^{\infty} e^{ik(y-y')} dk_y (2\pi)^{-1} \int_{-\infty}^{\infty} e^{ik(z-z')} dk_z \\
 &= (2\pi)^{-3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i[k_x(x-x') + k_y(y-y') + k_z(z-z')]} dk_x dk_y dk_z \\
 &= (2\pi)^{-3} \int e^{i\mathbf{k} \cdot (\mathbf{r} - \mathbf{r}')} d\mathbf{r} \tag{4.22}
 \end{aligned}$$

Again, very little has changed except that we have $(2\pi)^{-3}$ instead of $(2\pi)^{-1}$ (guess what we'd have in 2D?) We could also repeat this whole procedure, except swapping variables \mathbf{r} and \mathbf{k} , which would give us

$$\delta(\mathbf{k} - \mathbf{k}') = (2\pi)^{-3} \int_{\text{allspace}} e^{i\mathbf{k} \cdot (\mathbf{r} - \mathbf{r}')} , \tag{4.23}$$

a δ -function in k -space.

So, we can now derive the three dimensional Fourier Transform, which is left as an Exercise:

$$F(\mathbf{k}) = (2\pi)^{-\frac{3}{2}} \int_{\text{allspace}} f(\mathbf{r}) e^{-i\mathbf{k} \cdot \mathbf{r}} d\mathbf{r} \tag{4.24}$$

and

$$f(\mathbf{r}) = (2\pi)^{-\frac{3}{2}} \int_{\text{allk-space}} F(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{r}} d\mathbf{k} \tag{4.25}$$

Exercise 4.2.1.

Manually derive Equations 4.24 and 4.25

4.3 Momentum representation

Now (finally!) we can explore the relevance of the Fourier transform to quantum mechanics. Recall that the wave function for a free particle in its most general state is given by Equation 3.9 to be

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

Let's write this in a slightly different way. Let's let $A(\mathbf{k}, t) = A(\mathbf{k}) e^{-i\omega t}$ so that we can write

$$\Psi(\mathbf{r}, t) = \int A(\mathbf{k}, t) e^{i\mathbf{k} \cdot \mathbf{r}} d\mathbf{k} \tag{4.26}$$

and let's also change variables from wavenumber to momentum by the substitution $\mathbf{p} = \hbar\mathbf{k}$ (which you'll recall from Section 3.3.) Notice that because $d\mathbf{k} = dk_x dk_y dk_z$, and $dp_i = \hbar dk_i$ we must replace $d\mathbf{k}$ with $d\mathbf{p}/\hbar^3 = d\mathbf{p}(2\pi)^3\hbar^{-3}$. Changing to some new variable $A(\mathbf{p}, t)$ (which we could presumably calculate somehow, just by replacing any dependence on \mathbf{k} by \mathbf{p}/\hbar), this finally gives us:

$$\Psi(\mathbf{r}, t) = \hbar^{-3} \int A(\mathbf{p}, t) e^{i\mathbf{p}\cdot\mathbf{r}/\hbar} d\mathbf{p} \quad (4.27)$$

Next, let's look at what would happen if we decompose this function using the Fourier Inversion Theorem. We'll do this manually, just to avoid any confusion because there is the slight difference that $\Psi(\mathbf{r}, t)$ depends on time as well as \mathbf{r} . But we just carry this time dependence through into the Fourier transform, and also change to momentum space instead of \mathbf{k} -space:

$$\begin{aligned} \Psi(\mathbf{r}, t) &= \int \Psi(\mathbf{r}', t) \delta(\mathbf{r} - \mathbf{r}') d\mathbf{r}' \\ &= \int \Psi(\mathbf{r}', t) \left[(2\pi)^{-3} \int e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} d\mathbf{k} \right] d\mathbf{r}' \\ &= \int \Psi(\mathbf{r}', t) \left[(\hbar)^{-3} \int e^{i\mathbf{p}\cdot(\mathbf{r}-\mathbf{r}')/\hbar} d\mathbf{p} \right] d\mathbf{r}' \\ &= \hbar^{-\frac{3}{2}} \int e^{i\mathbf{p}\cdot\mathbf{r}/\hbar} \left[\hbar^{-\frac{3}{2}} \int \Psi(\mathbf{r}, t) \mathbf{r}' \cdot t e^{-i\mathbf{p}\cdot\mathbf{r}'/\hbar} d\mathbf{r}' \right] d\mathbf{p} \\ &= \hbar^{-\frac{3}{2}} \int \Phi(\mathbf{p}, t) e^{i\mathbf{p}\cdot\mathbf{r}/\hbar} d\mathbf{p} \end{aligned} \quad (4.28)$$

where (with a quick relabel of \mathbf{r}' to \mathbf{r})

$$\Phi(\mathbf{p}, t) = \hbar^{-\frac{3}{2}} \int \Psi(\mathbf{r}, t) e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar} d\mathbf{r} \quad (4.29)$$

is the Fourier transform of the wave function, and Equation 4.28 is the Fourier inversion theorem.

Now look closely at Equation 4.28. You'll see that it looks very much like our free particle wave function, given in the form of Equation 4.27. But instead of $A(\mathbf{p}, t)$ we have $\Phi(\mathbf{p}, t)$.

What is the significance of $A(\mathbf{p}, t)$? When we add all the momentum states together (to get a general state), we multiply each of them by a coefficient $A(\mathbf{k})$. The waves with higher coefficients are going to contribute more to the shape of the final wave function – their peaks will be a lot higher. This suggests that $A(\mathbf{k})$ might be related to the probability of finding the particle with wave number \mathbf{k} – the higher the coefficient the more the particle is “in” that momentum state. Based on that, we would expect $A(\mathbf{p}, t)$ to be related to probability also.

Therefore, it is logical to assume that $\Phi(\mathbf{p}, t)$ might also be related to the probability amplitude of the momentum of the particle. This is what we were

hinting at at the start of this section - a general state is made up of a sum² of oscillating pure momentum states. Then, just like for a sound wave, the Fourier transform returns the amplitude of each of these states. That is, $\Phi(\mathbf{p}', t)$ is a measure of “how much” the particle has momentum \mathbf{p}' .

It differs from $A(k)$ because it is valid for *any* wave function, not just a free particle state, and can be calculated directly from $\Psi(\mathbf{r}, t)$ by Equation 4.29 – we don’t always make our wave function by integrating over the pure momentum states, and hence don’t always know $A(\mathbf{k})$.

This discussion leads us to Postulate 3:

Postulate 3.

If the wave function for a particle is $\Psi(\mathbf{r}, t)$ then $\Phi^(\mathbf{p}, t)\Phi(\mathbf{p}, t)$ is the probability density in momentum-space of finding the particle to have a momentum \mathbf{p} at time t , where*

$$\Phi(\mathbf{p}, t) = h^{-\frac{3}{2}} \int \Psi(\mathbf{r}, t) e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar} d\mathbf{r}$$

$\Psi(\mathbf{r}, t)$ is called the coordinate or \mathbf{r} -representation, and $\Phi(\mathbf{p}, t)$ is called the momentum or \mathbf{p} -representation. They are Fourier transform pairs, and can be derived from each other by:

$$\Phi(\mathbf{p}, t) = h^{-\frac{3}{2}} \int \Psi(\mathbf{r}, t) e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar} d\mathbf{r} \quad (4.30)$$

$$\Psi(\mathbf{r}, t) = h^{-\frac{3}{2}} \int \Phi(\mathbf{p}, t) e^{i\mathbf{p}\cdot\mathbf{r}/\hbar} d\mathbf{p} \quad (4.31)$$

What’s important about this is that knowing $\Psi(\mathbf{r}, t)$ we can calculate $\Phi(\mathbf{p}, t)$ and visa versa - hence, both must contain *exactly* the same amount of information, just presented in a different way. Thus $\Phi(\mathbf{p}, t)$ is just as good description of the state of the particle as $\Psi(\mathbf{r}, t)$. This is why we call them **representations**, because they are just mathematically different ways of viewing the information about the particle.

4.4 Pure momentum states revisited

Okay, so we can represent a particle as a function of momentum – but why do we care? Well, given a wave function (in the position representation) we can calculate the probability of a particle being found in a certain place, the probability density. Then, by taking the Fourier transform, we can use our new momentum representation to find the probability of the particle having a particular momentum; we find a probability density for momentum now, instead of position. This was something we couldn’t do before - when we superimposed all the momentum states, it was no longer clear what the momentum of the particle was.

²An integration, to be precise.

Let's think about what the momentum representation should be for a particle whose momentum is known definitely to be \mathbf{p}' . Don't go any further until you've worked it out!

Hopefully, you realised that it would have to be a δ -function, but of momentum and centred at \mathbf{p}' . Let's check this now.

In Postulate 1 (Section 3.3) we said that the wave function for a free particle with a well defined wavenumber \mathbf{k} was

$$\Psi(\mathbf{r}, t) = Ae^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} = Ae^{i(\mathbf{p}'\cdot\mathbf{r}/\hbar-\omega t)}$$

where we have used \mathbf{p}' to make it clear that it is an actual known value, and not a variable.

Let's convert this to momentum representation:

$$\begin{aligned}\Phi(\mathbf{p}, t) &= h^{-\frac{3}{2}} \int \Psi(\mathbf{r}, t) e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar} d\mathbf{r} \\ &= h^{-\frac{3}{2}} \int \left[Ae^{i(\mathbf{p}'\cdot\mathbf{r}/\hbar-\omega t)} \right] e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar} d\mathbf{r} \\ &= h^{-\frac{3}{2}} Ae^{-i\omega t} \int e^{i\mathbf{r}\cdot(\mathbf{p}'-\mathbf{p})} \\ &= h^{-\frac{3}{2}} Ae^{-i\omega t} \delta(\mathbf{p}-\mathbf{p}')\end{aligned}$$

So for a particle whose momentum is known exactly to be \mathbf{p}' , its momentum representation is just a δ -function, but scaled so that its height changes with time. This might sound a little strange, but remember that waves usually oscillate, and this often the case for the probability amplitude. More physical information can be obtained if we look at the probability density:

$$\begin{aligned}\Phi^*(\mathbf{p}, t)\Phi(\mathbf{p}, t) &= \left[h^{-\frac{3}{2}} Ae^{i\omega t} \delta^*(\mathbf{p}-\mathbf{p}') \right] \left[h^{-\frac{3}{2}} Ae^{-i\omega t} \delta(\mathbf{p}-\mathbf{p}') \right] \\ &= h^{-3} |A|^2 \delta^*(\mathbf{p}-\mathbf{p}') \delta(\mathbf{p}-\mathbf{p}') \\ &= h^{-3} |A|^2 \delta^2(\mathbf{p}-\mathbf{p}') \\ &= h^{-3} |A|^2 \delta(\mathbf{p}-\mathbf{p}')\end{aligned}$$

This lets us see clearer what we already figured out - when the momentum is known exactly, the momentum probability density is just a δ -function (scaled in height), centered on \mathbf{p}' . So the only possible measurement for momentum can be \mathbf{p}' , as we would hope! In making a realistic wave function, this scaling would be adjusted so that the area under the $\Phi^*(\mathbf{p}, t)\Phi(\mathbf{p}, t)$ graph is unity (which was the whole point of defining the δ -function!)

For a more general wave function, our answer would obviously be more complicated. A combination of a finite number of pure momentum states would give that many δ -function peaks, centred at the possibly momentum values. Their height would be determined by the relative probabilities of those values, and by normalisation. Exercise *** explores this further. An infinite combination of pure states may result in a continuous momentum wave function.

Example 4.4.1.

Consider the Gaussian wave function from Example 3.3.1:

$$\Psi(x) = e^{-\beta x^2/2} \quad (4.32)$$

Find its momentum representation.

Answer.

Using the one dimensional Fourier Transform, we have

$$\Phi(p) = h^{-\frac{1}{2}} \int \Psi(\mathbf{r}, t) e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar} d\mathbf{r}$$

And (working will be added soon!) we can show that this is another Gaussian!

Thus, we can see now how a wave function could represent a “realistic” particle – both its position and its momentum are known approximately to be some given values, but there is some uncertainty in each.

Now, you may begin to ask yourself whether we could change to an “angular momentum” representation, and obtain a similar probability density function. You might also wonder whether quantisation falls out naturally from these wave functions. The next chapter deals with obtaining even more information from the wave function, and introduces the next most important concept (after wave functions!) in quantum – operators.

4.5 Chapter Summary

- The Dirac δ -function denoted $\delta(x-x')$ is a special function that has a peak at x' , is zero everywhere else, but yet still has unit area. This is useful, for example, in making a position wave function representing a particle whose position is known exactly. (Section 4.1)
- Two possible, and in fact equal, representations of the δ -function are

$$\delta(x-x') = \lim_{w \rightarrow 0} \frac{1}{\pi} \frac{2w}{(x-x')^2 + w^2}$$

and

$$\delta(x-x') = (2\pi)^{-1} \int_{-\infty}^{\infty} e^{ik(x-x')} dk$$

These are not the only possible ways of representing it, but they are the most useful in quantum physics. (Section 4.1)

- In three dimensions, we can write our δ -function as

$$(2\pi)^{-3} \int e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} d\mathbf{r}$$

- Any real function can be composed of a series of sine waves, and the **Fourier Transform** can be used to determine which, and in what quantity, the composite waves are needed. The **Fourier Inversion Theorem** takes a series of sine waves and returns the result of adding them all together, the opposite of the Fourier Transform. Therefore, applying the fourier transform and then the inversion theorem, just returns the original function. (Section 4.2)
- A complex function can be decomposed in just the same way, except now we need waves oscillating in the complex plane (for example, e^{ikx}). Therefore, the general Fourier Transform and Inversion can be written

$$F(k) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} e^{-ikx} f(x) dx$$

$$f(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} e^{ikx} F(k) dk$$

(Section 4.2)

- In quantum mechanics, the Fourier Transform of a (position) wave function is related to the probability amplitude of finding the particle with particular wave number (or momentum). We refer to this as the **momentum representation**. This follows from the fact that we can construct a wave function from a series of pure momentum states (basically oscillating sine waves), and the Fourier Transform gives the “frequency” (in this case wave number) of these waves. (Section 4.3)
- These two representations contain exactly the same amount of information, and we can switch between them by

$$\Phi(\mathbf{p}, t) = h^{-\frac{3}{2}} \int \Psi(\mathbf{r}, t) e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar} d\mathbf{r}$$

$$\Psi(\mathbf{r}, t) = h^{-\frac{3}{2}} \int \Phi(\mathbf{p}, t) e^{i\mathbf{p}\cdot\mathbf{r}/\hbar} d\mathbf{p}$$

(Section 4.3)

- In the momentum representation, a pure momentum state is represented by a δ -function, as we would hope.

4.6 Exercises

TBA

Chapter 5

Analysing the Wave function

So far, given a wave function, we've been able to determine what the probabilities are of measuring a certain value of position or momentum for the particle. However, we often want to find out other things – like what the probability is of measuring a *function* of position, say. We might also be interested in finding average values and so forth.

This chapter explores how to squeeze out all the information contained in the wave function. We will also discuss *measurement* in quantum mechanics, and why it is so important (and confusing!)

5.1 Expectation values and variance

We know that for a given state, we can potentially obtain different results for each measurement of position (with the particle being prepared in the same state before each measurement). So, it makes sense to ask what the average value of all these measurements would be, which we call the *expectation value* of position. Can we predict this theoretically from our wave function?

In statistics, the expectation value is defined to be the sum of the possible values multiplied by their probabilities, just like when we found the average energy in Section 2.1. Because position is continuous, however, the sum is actually an integral.

What is the probability of obtaining measuring the particle to be at $\mathbf{r}_0 = (x_0, y_0, z_0)$? This is the probability density we've already discussed in detail: the probability of a particle's position being between \mathbf{r}_0 and $\mathbf{r}_0 + d\mathbf{r}$ is $\rho(\mathbf{r}_0, t)d\mathbf{r}$.

Thus, the expectation value of \mathbf{r} would be

$$\langle \mathbf{r} \rangle = \int \mathbf{r} \cdot \rho(\mathbf{r}, t) d\mathbf{r} \quad (5.1)$$

$$= \int \Psi^*(\mathbf{r}, t) \mathbf{r} \Psi(\mathbf{r}, t) d\mathbf{r} \quad (5.2)$$

(The order of the terms in the integral don't matter, of course, but we'll write them this way, for reasons that will become clear soon!) Here, we integrate \mathbf{r} (representing the possible measurements) times $\rho(\mathbf{r}, t)$ (effectively the probability of obtaining each measurement) over all \mathbf{r} -space. Refer back to Section 1.2 on continuous probability if this isn't clear. (The idea of probability density and how to use it is a difficult concept, but it will become clearer as you are more familiar with it.)

Similarly, the expectation values of x, y and z can be calculated by

$$\langle x \rangle = \int x \cdot \rho(\mathbf{r}, t) d\mathbf{r} \quad (5.3)$$

You may find this fairly intuitive, but just to make this absolutely clear, we can expand out Equation 5.1 as follows:

$$\begin{aligned} \langle \mathbf{r} \rangle &= \int \mathbf{r} \cdot \rho(\mathbf{r}, t) d\mathbf{r} \\ \langle x \hat{\mathbf{i}} + y \hat{\mathbf{j}} + z \hat{\mathbf{k}} \rangle &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \\ (x \hat{\mathbf{i}} + y \hat{\mathbf{j}} + z \hat{\mathbf{k}}) \cdot \rho(x, y, z, t) dx dy dz \\ \langle x \rangle \hat{\mathbf{i}} + \langle y \rangle \hat{\mathbf{j}} + \langle z \rangle \hat{\mathbf{k}} &= \hat{\mathbf{i}} \int_{-\infty}^{\infty} x \cdot \rho(x, y, z, t) dx + \hat{\mathbf{j}} \int_{-\infty}^{\infty} y \cdot \rho(x, y, z, t) dy + \hat{\mathbf{k}} \int_{-\infty}^{\infty} z \cdot \rho(x, y, z, t) dz \end{aligned}$$

and then by equating the $\hat{\mathbf{i}}, \hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ vectors on each side, we get the expectation values of each component as suggested above. It's important to note then that the expectation value of a vector is just a vector of the expectation values of its components, i.e. $\langle \mathbf{r} \rangle = \langle (x, y, z) \rangle = (\langle x \rangle, \langle y \rangle, \langle z \rangle)$.

Can we also compute the expectation value of a function of \mathbf{r} such as the potential energy of a spring $\frac{1}{2}k|\mathbf{r}|^2$? Let's consider a general function of position $f(\mathbf{r})$. Firstly, the possible values of $f(\mathbf{r})$ will depend only on the values that \mathbf{r} takes. So in a state where the position is known to be $2\hat{\mathbf{i}}$, the energy of the spring could only be $\frac{1}{2}k \cdot 4$, while if the position is either $2\hat{\mathbf{i}}$ or $3\hat{\mathbf{i}}$, then the energy could be either $\frac{1}{2}k \cdot 4$ or $\frac{1}{2}k \cdot 9$. The probability of finding each value of $f(\mathbf{r})$ just depends on the the probability of the corresponding \mathbf{r} value. So, thinking of the spring again, if we were twice as likely to find the spring out at $3\hat{\mathbf{i}}$ than $2\hat{\mathbf{i}}$, then we would expect to measure its energy to be $\frac{9}{2}k$ twice as often as $2k$.

Therefore, in our expectation formula above, the only thing that changes are the possible values; we still integrate over $d\mathbf{r}$, and the probabilities are still the

Position	Count
1	10
2	35
3	45
6	20

same, but we replace the \mathbf{r} at the front (representing the possible measurements) with $f(\mathbf{r})$ as in

$$\langle f(\mathbf{r}) \rangle = \int f(\mathbf{r}) \cdot \rho(\mathbf{r}, t) d\mathbf{r} \quad (5.4)$$

In addition to knowing the expectation value, we might want to know what the *spread* of the measurements are - are most of the measurements to be found near the average? Or are they spread out over the whole space? A measure of this is the **variance**, which is defined to be

$$\langle (f(\mathbf{r}) - \langle f(\mathbf{r}) \rangle)^2 \rangle = \langle f(\mathbf{r})^2 \rangle - \langle f(\mathbf{r}) \rangle^2 \quad (5.5)$$

A full explanation is really more statistics, but mathematically, the variance is the average of the square of the separation of all measurements from the mean. (Phew!) For the rest of us, while not exactly correct, it may be simpler just to think of it as the average distance of the measurements from the mean. Therefore, a large variance means that the measurements are very spread out, while a zero variance means that we obtain no measurements other than the expectation value.

Exercise 5.1.1.

Prove the equality in 5.5.

Hint: Expand the brackets, and ask your self what the expectation value of an expectation value is.

Exercise 5.1.2.

What would a graph of probability vs. position look like as the variance approaches zero? Does it look familiar?

Exercise 5.1.3.

An experimenter makes 100 measurements of a particle's position and finds

What is the expectation value of her measurements? What is the variance?

Qualitatively describe the spread of data.

What about the expectation value of a function of momentum? (So either \mathbf{p} itself, or a function like energy, which is $\frac{\mathbf{p}^2}{2m}$) This time, we can use the momentum representation, which gives us the probability distribution of momentum.

So,

$$\begin{aligned}\langle g(\mathbf{p}) \rangle &= \int g(\mathbf{p})\rho(\mathbf{p}, t)d\mathbf{p} \\ &= \int \Phi^*(\mathbf{p}, t)\mathbf{p}\Phi(\mathbf{p}, t)\end{aligned}$$

And the variance will be simply

$$\langle (g(\mathbf{p})) - \langle g(\mathbf{p}) \rangle \rangle^2 \quad (5.6)$$

But what happens when we require the expectation value of a function that depends on both \mathbf{r} and \mathbf{p} ? An example would be angular momentum which is given by $\mathbf{r} \times \mathbf{p}$. Now, using just $\Psi(\mathbf{r}, t)$ or $\Phi(\mathbf{p}, t)$ doesn't seem to be sufficient, because we need to consider the probability of both \mathbf{r} and \mathbf{p} simultaneously.

An initial idea might be to try and use two integrals, and sum over both position and momentum probability densities, but unfortunately, this soon leads to a dead end. Instead we ask: do we really need the momentum representation here at all? Think back to Postulate 3, and how we said that both representations carry exactly the same information – there shouldn't be anything we can do with one but not the other.

The most basic way to get rid of the momentum wave functions would be just to replace every $\Phi(\mathbf{p}, t)$ with $h^{-\frac{3}{2}} \int \Psi(\mathbf{r}, t)e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar}d\mathbf{r}$. As is, that doesn't help much with our two variable problem. But let's see if we can do some simplifications on that idea that might help.

Let's start with the expectation value of the x -component of momentum and replace the $\Phi(\mathbf{p}, t)$'s.

$$\begin{aligned}\langle p_x \rangle &= \int \Phi^*(\mathbf{p}, t)p_x\Phi(\mathbf{p}, t)d\mathbf{p} \\ &= \int \left[h^{-\frac{3}{2}} \int \Psi^*(\mathbf{r}', t)e^{i\mathbf{p}\cdot\mathbf{r}'/\hbar}d\mathbf{r}' \right] p_x \left[h^{-\frac{3}{2}} \int \Psi(\mathbf{r}, t)e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar}d\mathbf{r} \right] d\mathbf{p} \\ &= h^{-3} \int \int \int \Psi^*(\mathbf{r}', t)\Psi(\mathbf{r}, t)p_x e^{-i\mathbf{p}\cdot(\mathbf{r}-\mathbf{r}')/\hbar}d\mathbf{r}'d\mathbf{r}d\mathbf{p}\end{aligned}$$

Because $\mathbf{p} \cdot \mathbf{r} = p_x x + p_y y + p_z z$, we can say that

$$p_x e^{-i\mathbf{p}\cdot(\mathbf{r}-\mathbf{r}')/\hbar} = i\hbar \frac{\partial}{\partial x} e^{-i\mathbf{p}\cdot(\mathbf{r}-\mathbf{r}')/\hbar}$$

This is just maths and the simplest way to check it is to evaluate the RHS, and show they are equal. Hence, substituting this into the integral, and then re-arranging the terms,

$$\begin{aligned}\langle p_x \rangle &= h^{-3} \int \int \int \Psi^*(\mathbf{r}', t)\Psi(\mathbf{r}, t)i\hbar \frac{\partial}{\partial x} e^{-i\mathbf{p}\cdot(\mathbf{r}-\mathbf{r}')/\hbar}d\mathbf{r}d\mathbf{r}'d\mathbf{p} \\ &= h^{-3} \int \int \Psi^*(\mathbf{r}', t) \left[\int \Psi(\mathbf{r}, t)i\hbar \frac{\partial}{\partial x} e^{-i\mathbf{p}\cdot(\mathbf{r}-\mathbf{r}')/\hbar}d\mathbf{r} \right] d\mathbf{r}'d\mathbf{p}\end{aligned}$$

Now, let's just look at those terms inside [...] and expand out our integration shorthand:

$$\begin{aligned} & \int \Psi(\mathbf{r}, t) i\hbar \frac{\partial}{\partial x} e^{-i\mathbf{p}\cdot(\mathbf{r}-\mathbf{r}')/\hbar} d\mathbf{r} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} \Psi(\mathbf{r}, t) i\hbar \frac{\partial}{\partial x} e^{-i\mathbf{p}\cdot(\mathbf{r}-\mathbf{r}')/\hbar} dx \right] dy dz \end{aligned}$$

Now, again just consider the inner integration. The derivative suggests trying to use integration by parts (if it doesn't suggest it, have a go using it anyway...):

$$\begin{aligned} [\dots] &= \int_{-\infty}^{\infty} \Psi(\mathbf{r}, t) i\hbar \frac{\partial}{\partial x} e^{-i\mathbf{p}\cdot(\mathbf{r}-\mathbf{r}')/\hbar} dx \\ &= \Psi(\mathbf{r}, t) i\hbar \frac{\partial}{\partial x} e^{-i\mathbf{p}\cdot(\mathbf{r}-\mathbf{r}')/\hbar} \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} i\hbar e^{-i\mathbf{p}\cdot(\mathbf{r}-\mathbf{r}')/\hbar} \frac{\partial}{\partial x} \Psi(\mathbf{r}, t) dx \\ &= \int_{-\infty}^{\infty} e^{-i\mathbf{p}\cdot(\mathbf{r}-\mathbf{r}')/\hbar} \left(i\hbar \frac{\partial}{\partial x} \right) \Psi(\mathbf{r}, t) dx \end{aligned}$$

The first term after the integration by parts vanishes because we assume the wave function vanishes at infinity¹.

Now we can sub this expression back into all the previous equations, and arrive at

$$\begin{aligned} \langle p_x \rangle &= \hbar^{-3} \int \int \int \Psi^*(\mathbf{r}', t) e^{i\mathbf{p}\cdot(\mathbf{r}'-\mathbf{r})/\hbar} \left(i\hbar \frac{\partial}{\partial x} \right) \Psi(\mathbf{r}, t) d\mathbf{r}' d\mathbf{r} d\mathbf{p} \\ &= \int \int \Psi^*(\mathbf{r}', t) \left[\hbar^{-3} \int e^{i\mathbf{p}\cdot(\mathbf{r}'-\mathbf{r})/\hbar} d\mathbf{p} \right] \left(i\hbar \frac{\partial}{\partial x} \right) \Psi(\mathbf{r}, t) d\mathbf{r}' d\mathbf{r} \\ &= \int (\Psi^*(\mathbf{r}', t) [\delta(\mathbf{r}' - \mathbf{r}) d\mathbf{r}']) \left(i\hbar \frac{\partial}{\partial x} \right) \Psi(\mathbf{r}, t) d\mathbf{r} \\ &= \int \Psi^*(\mathbf{r}, t) \left(i\hbar \frac{\partial}{\partial x} \right) \Psi(\mathbf{r}, t) d\mathbf{r} \\ &= \int \Psi^*(\mathbf{r}, t) \hat{p}_x \Psi(\mathbf{r}, t) d\mathbf{r} \end{aligned} \tag{5.7}$$

where we define $\hat{p}_x = -i\hbar \frac{\partial}{\partial x}$. We give it a hat to distinguish it from the variable p_x , but we don't give it a completely different name, so that we can see that it still relates to the x coordinate of momentum.

You can see that Equation 5.7 looks just the expectation value of a function of x , namely $\hat{p}_x = -i\hbar \frac{\partial}{\partial x}$.

Exercise 5.1.4.

Show that we can calculate $\langle p_y \rangle$ and $\langle p_z \rangle$ by replacing p_y and p_z by $\hat{p}_y = -i\hbar \frac{\partial}{\partial y}$ and $\hat{p}_z = -i\hbar \frac{\partial}{\partial z}$ respectively.

¹If it *didn't* vanish, there would be some probability of finding our particle *at* infinity, which doesn't make physical sense.

We can also show that

$$\langle p_x^n \rangle = \int \Psi^*(\mathbf{r}, t) \left(i\hbar \frac{\partial}{\partial x} \right)^n \Psi(\mathbf{r}, t) d\mathbf{r} \quad (5.8)$$

using a similar procedure to above (See Exercise (***)). It seems, then, that when working in the position representation, we can replace \mathbf{p} by $-i\hbar \left(\frac{\partial}{\partial x} \hat{\mathbf{i}} + \frac{\partial}{\partial y} \hat{\mathbf{j}} + \frac{\partial}{\partial z} \hat{\mathbf{k}} \right)$ and then proceed as if we were working with a function of \mathbf{r} . Therefore, we could find the expectation value of any polynomial of p_x, p_y and p_z by

$$\langle f(p_x, p_y, p_z) \rangle = \int \Psi^*(\mathbf{r}, t) f(\hat{p}_x, \hat{p}_y, \hat{p}_z) \Psi(\mathbf{r}, t) d\mathbf{r}$$

We just replace every p_x in the function by $-i\hbar \frac{\partial}{\partial x}$ and likewise for p_y and p_z . Hence for any (polynomial) function of x, y, z, p_x, p_y, p_z , we can say

$$\langle F(x, y, z, p_x, p_y, p_z) \rangle = \int \Psi^*(\mathbf{r}, t) F(x, y, z, \hat{p}_x, \hat{p}_y, \hat{p}_z) \Psi(\mathbf{r}, t) d\mathbf{r} \quad (5.9)$$

If you have done some multivariable calculus, you will recognise $\frac{\partial}{\partial x} \hat{\mathbf{i}} + \frac{\partial}{\partial y} \hat{\mathbf{j}} + \frac{\partial}{\partial z} \hat{\mathbf{k}}$ as the gradient function or ∇ (read as “del” or sometimes “nabla”). If you haven’t, then not to worry – just accept that we will use ∇ as the symbol for the 3D momentum operator in the position representation. Then, as mentioned in Chapter ??, we often shorthand $F(x, y, z, p_x, p_y, p_z)$ to $F(\mathbf{r}, \mathbf{p})$. So, we can write

$$\langle F(\mathbf{r}, \hat{\mathbf{p}}) \rangle = \int \Psi^*(\mathbf{r}, t) F(\mathbf{r}, \nabla) \Psi(\mathbf{r}, t) d\mathbf{r}$$

Notice that now the order of terms in the integral *is* important - changing the position of $F(\mathbf{r}, \hat{\mathbf{p}})$, or swapping $\Psi^*(\mathbf{r}, t)$ and $\Psi(\mathbf{r}, t)$ around would change what the partial derivative acts on, and would (in general) change the integration.

And so we have solved our problem - we can now calculate the expectation value of a physical quantity that involves both momentum and position. Better yet, we don’t need to calculate the momentum representation to do it; as we expected, the position representation is all we need. What we have derived, \hat{p}_x , is called the momentum **operator** in the position representation.

5.2 Operators

Like with the example of angular momentum above, we often want to calculate the expectation values of functions of more than one type of variable, such as position or momentum, but also other quantities like charge or spin (which you may not have heard of!). We can do this completely from the comfort of the position representation, by calculating the expectation value of the **operator** corresponding to that physical quantity.

But what exactly do we mean by an “operator”? An operator, in general, is very much like a function - it acts on something to give an answer that is of the

same "type". So, for example, matrices "operate" on vectors by normal matrix multiplication to give a new vector:

$$MATRIX \times VECTOR = VECTOR$$

This is an excellent example of operators because it has nothing to do with functions, like most of the others that we consider do. However

In quantum, operators are denoted by placing a hat over a letter, like \hat{A} . Then $\hat{A}x$ denotes the operator A acting on x . Operators in quantum mechanics are **linear operators**. These are operators such that $f(x_1+x_2) = f(x_1)+f(x_2)$. An example would be $f(x) = 4x$, because

$$f(x_1 + x_2) = 4(x_1 + x_2) = 4x_1 + 4x_2 = f(x_1) + f(x_2)$$

But functions like $f(x) = x^2$ or even $f(x) = 2x+3$ are not linear, which is easily checked.

Perhaps the simplest types of operators are numbers themselves – they operate on nearly anything, with their action being multiplication. So $\hat{2}f(x) = 2f(x)$. Fairly obvious? The negative sign is also an operator – it takes a function and returns the negative of it. (Of course, the negative sign is equivalent to -1)

Another class of operators, that you have already met in the preceding section, is the **differential operators**. These are functions like $\frac{\partial}{\partial x}$ or $\frac{\partial}{\partial t}$ which give the rate of change. They act on other *functions*, and so return functions also. The momentum operator defined in the last section as $\hat{p}_x = -i\hbar\frac{\partial}{\partial x}$ is a differential operator. So too is the three dimensional momentum operator, $\hat{\mathbf{p}} = \nabla$. A simple example is acting $\hat{\mathbf{p}}$ on a function $\Psi(\mathbf{r}) = x^2 + \sin(y)$ (but note that this wave function is just a mathematical example – it is very unrealistic!):

$$\begin{aligned} \hat{\mathbf{p}}\Psi(\mathbf{r}) &= \nabla\Psi(\mathbf{r}) \\ &= \frac{\partial}{\partial x} (x^2 + \sin(y)) \hat{\mathbf{i}} + \frac{\partial}{\partial y} (x^2 + \sin(y)) \hat{\mathbf{j}} + \frac{\partial}{\partial z} (x^2 + \sin(y)) \hat{\mathbf{k}} \\ &= 2x \hat{\mathbf{i}} + \cos(y) \hat{\mathbf{j}} \end{aligned}$$

which is just another function of \mathbf{r} .

In quantum we represent all physical quantities by operators – momentum, position, and anything else we might measure. Why do we do this? Imagine we decided to use the momentum representation instead of position for all our expectation value work. Then to calculate the average momentum we would use $\int \Phi^*(\mathbf{p}, t) p_x \Phi(\mathbf{p}, t) d\mathbf{p}$, whereas one can show that the average x -position can be given by $\int \Phi^*(\mathbf{p}, t) (i\hbar\frac{\partial}{\partial x}) \Phi(\mathbf{p}, t) d\mathbf{p}$ by using a similar argument to before. Notice the symmetry between the two?

The point is that in both representations we're finding the same thing – the average value of momentum or position. And this doesn't depend on what mathematical representation we use! So when we talk about the momentum operator, we're talking about the physical concept of momentum, without worrying about what representation we choose to work in. Only once we've picked

either the position, momentum or some other representation do we have to consider what form the operator takes.

Table ?? shows the form of various different operators in the position representation. That's the rep that we usually use, although you could equally well use momentum.

INSERT TABLE HERE - CONTAINS BASIC STUFF DERIVED SO FAR + HAMILTONIAN + ANGULAR MOMENTUM

Actually working out the form of an operator in a given representation requires a combination of clever mathematics (like calculating \hat{p}_x above!) and intuitive guess work. The guess work is needed particularly working with concepts such as spin which has no classical counterpart.

One final point before we move on. You might ask what the returned wave function means when we act an operator on it. One option is that the operator leaves the wave function almost unchanged, except by a constant. Wave functions where this happens are called *eigenfunctions* of that operator, and we'll be studying these hard shortly. In some cases, the new wave function represents a related, yet different, state of the particle. An example is the *raising* operator, which is related to energy levels (like you get for electrons in an atom – think Bohr!). If you take the wave function for a particle in the first energy level and act the raising operator on it, you get back the wave function of the second level, and so on.² In most other cases, the resulting wave function is not of any real physical interest, but still works nicely for our expectation values.

5.3 Properties of Operators

All operators obey several properties which are listed below. Many of these will seem blatantly obvious, but it's still good to list them all so that you see exactly what you can do with them.

- Operators are linear.

This means that if \hat{A} is an operator (that operates on functions) and $f_1(\mathbf{r}, \mathbf{p})$, $f_2(\mathbf{r}, \mathbf{p})$, $g_1(\mathbf{r}, \mathbf{p})$ and $g_2(\mathbf{r}, \mathbf{p})$ are functions such that

$$\hat{A}f_1(\mathbf{r}, \mathbf{p}) = g_1(\mathbf{r}, \mathbf{p}) \quad \hat{A}f_2(\mathbf{r}, \mathbf{p}) = g_2(\mathbf{r}, \mathbf{p})$$

then we can say that

$$\begin{aligned} & \hat{A}(cf_1(\mathbf{r}, \mathbf{p}) + df_2(\mathbf{r}, \mathbf{p})) \\ = & c\hat{A}f_1(\mathbf{r}, \mathbf{p}) + d\hat{A}f_2(\mathbf{r}, \mathbf{p}) \\ = & cg_1(\mathbf{r}, \mathbf{p}) + dg_2(\mathbf{r}, \mathbf{p}) \end{aligned}$$

Basically, this property means that you can use the distributive law when using operators.

²This is a rather simplified description, but it should give you some idea.

- The sum and product of two operators \hat{A} and \hat{B} are defined by

$$(\hat{A} + \hat{B})f(\mathbf{r}, \mathbf{p}) = \hat{A}f(\mathbf{r}, \mathbf{p}) + \hat{B}f(\mathbf{r}, \mathbf{p}) \quad (5.10)$$

and

$$(\hat{A}\hat{B})f(\mathbf{r}, \mathbf{p}) = \hat{A}(\hat{B}f(\mathbf{r}, \mathbf{p})) \quad (5.11)$$

Operators also satisfy the following basic algebra rules:

- $\hat{A} + \hat{B} = \hat{B} + \hat{A}$
- $(\hat{A} + \hat{B}) + \hat{C} = \hat{A} + (\hat{B} + \hat{C}) = \hat{A} + \hat{B} + \hat{C}$
- $(\hat{A}\hat{B})\hat{C} = \hat{A}(\hat{B}\hat{C}) = \hat{A}\hat{B}\hat{C}$
- $\hat{A}(\hat{B} + \hat{C}) = \hat{A}\hat{B} + \hat{A}\hat{C}$
- and $(\hat{A} + \hat{B})\hat{C} = \hat{A}\hat{C} + \hat{B}\hat{C}$

- The *identity operator* $\hat{1}$ is defined by

$$\hat{1} \cdot f(\mathbf{r}, \mathbf{p}) = f(\mathbf{r}, \mathbf{p})$$

- The inverse of an operator \hat{A} , if it exists, is such that

$$\hat{A}\hat{A}^{-1} = \hat{A}^{-1}\hat{A} = \hat{1}$$

- The complex conjugate of an operator is just the same as in normal maths. So, the complex conjugate of $\hat{p}_x = -i\hbar\frac{\partial}{\partial x}$ is $\hat{p}_x^* = i\hbar\frac{\partial}{\partial x}$, just by changing the sign on all i 's. We also say that

$$(\hat{F}\Psi(\mathbf{r}, t))^* = \hat{F}^*\Psi(\mathbf{r}, t)^* \quad (5.12)$$

This extends to as many terms as needed, so that $(\Psi^*(\mathbf{r}, t)\hat{F}\Psi(\mathbf{r}, t))^* = \Psi(\mathbf{r}, t)\hat{F}^*\Psi^*(\mathbf{r}, t)$ and so forth. Also, the complex conjugate isn't affected by integration; you can integrate and then take the complex conjugate, or the other way round, whichever makes the maths easier.

- All operators in quantum are designed to act from the left onto a wave function, as in

$$\hat{F}(\Psi(\mathbf{r}, t))$$

The action of an operator when it acts from the right is to act its complex conjugate from the left. This no doubt seems bizarre, and we won't go into the reason why (just yet anyway!) but it works. (Also don't ask what the physical meaning of the complex conjugate of position is ...) Mathematically, we write this as

$$\Psi(\mathbf{r}, t)\hat{F} = \hat{F}^*\Psi(\mathbf{r}, t)$$

At the moment this property doesn't affect us much, although we will use it in a couple of proofs now and then.

Although most of these are what you'd expect, just on your experience with algebra, one property that is significant in its absence is that of commutativity, i.e. in general $\hat{A}\hat{B} \neq \hat{B}\hat{A}$. And this is more than a mathematical curiosity – it plays an important role in the development of our theories that we need to explore further. First, let's just formalise a couple of things.

5.4 Expectation Values Again

We've already done the expectation value for momentum and position. Let's now state the expectation value and variance for some general operator \hat{F} , which could be position, momentum or something else entirely. We can define the expectation value of this operator just as before, by

$$\langle \hat{F} \rangle = \int \Psi^*(\mathbf{r}, t) \hat{F} \Psi(\mathbf{r}, t) d\mathbf{r} \quad (5.13)$$

So, you'd operate \hat{F} on the wave function $\Psi(\mathbf{r}, t)$, then multiply by $\Psi^*(\mathbf{r}, t)$ and integrate. Exactly how you operate the operator would be something you would need to determine, mathematically or otherwise. It's worth noting that once you've acted the operator on the wave function, the order of terms in the integration no longer matters, but until then the order must be $\Psi^*(\mathbf{r}, t) \hat{F} \Psi(\mathbf{r}, t)$.

To calculate the variance, we will introduce a new operator $\Delta \hat{F}$ such that

$$\Delta \hat{F} = \hat{F} - \langle \hat{F} \rangle \quad (5.14)$$

so that, similar to Equation 5.5,

$$\text{Variance} = \langle (\hat{F} - \langle \hat{F} \rangle)^2 \rangle \quad (5.15)$$

$$= \langle \hat{F}^2 \rangle - \langle \hat{F} \rangle^2 \quad (5.16)$$

$$= \langle \Delta \hat{F}^2 \rangle \quad (5.17)$$

This gives us three ways of looking at the variance of an operator. Convention is to label the variance by the last of those, and the second is the most useful to work with. So we can say

$$\langle \Delta \hat{F}^2 \rangle = \text{expec} \hat{F}^2 - \langle \hat{F} \rangle^2 \quad (5.18)$$

We'll just mention here a few common confusions about our definition of variance. First of all, the $\Delta \hat{F}$ operator is nothing new – it's really just a mathematical convenience, and we could do without it if we had to. However, it does have a physical meaning. Whereas the operator \hat{F} represents an actual measurement, $\Delta \hat{F}$ represents the distance of those measurements from the mean. We'll do an example in a moment which should clarify things.

One more point that sometimes confuses people is that we define $\Delta \hat{F}$ as an operator minus a number ($\langle \hat{F} \rangle$). How can you subtract a number from an operator? The easiest way is just to remember that numbers are operators too! It's just they don't change the wave function except for a constant. Let's look at how this actually works in an example.

Example 5.4.1.

In one particular state, the physical quantity \hat{F} can be either 2, 4 or 9 with equal probability. Hence the expectation value (which is the weighted mean of the measurements) is $\langle \hat{F} \rangle = 5$.

Therefore $\Delta\hat{F}$ in this case will take the form

$$\begin{aligned}\Delta\hat{F} &= \hat{F} - \langle F \rangle \\ &= \hat{F} - 5\end{aligned}$$

and because \hat{F} can be 2,4, or 9, $\Delta\hat{F}$ can take values -3,-1 and 4 – which are simply the distances of all possible measurements from the mean.

The variance is the average of the square of these. Since the values of \hat{F} occur with equal probability, so will the values of $\Delta\hat{F}$, and thus of $\Delta\hat{F}^2$. Note that if say $\Delta\hat{F}$ could be -3,2 or 3, then $\Delta\hat{F}^2$ would be 9 twice as often as 4, and this would need to be taken into account. In this case, however, we just add the squares and divide by 3:

$$\langle \Delta\hat{F}^2 \rangle = \frac{9 + 1 + 16}{3} = \frac{26}{3}$$

To double check, let's use the other definition (which is the one more commonly used, in fact) where

$$\langle \Delta\hat{F}^2 \rangle = \langle \hat{F}^2 \rangle - \langle \hat{F} \rangle^2$$

\hat{F}^2 could be either 4,16 or 81, and we know \hat{F} , so

$$\langle \Delta\hat{F} \rangle = \frac{4 + 16 + 81}{3} - 5^2 = \frac{26}{3}$$

just as we hoped!

We can also calculate the variance in terms on an integration, given a wave function $\Psi(\mathbf{r}, r)$. We expand the expectation value like

$$\begin{aligned}\langle \Delta\hat{F}^2 \rangle &= \int \Psi^*(\mathbf{r}, t) \Delta\hat{F}^2 \Psi(\mathbf{r}, t) d\mathbf{r} \\ &= \int \Psi^*(\mathbf{r}, t) \Delta\hat{F} \Delta\hat{F} \Psi(\mathbf{r}, t) d\mathbf{r} \\ &= \int \left(\Delta\hat{F}^* \Psi^*(\mathbf{r}, t) \right) \left(\Delta\hat{F} \Psi(\mathbf{r}, t) \right) d\mathbf{r} \quad (\text{By Eq ??}) \\ &= \int \left| \Delta\hat{F} \Psi(\mathbf{r}, t) \right|^2 d\mathbf{r}\end{aligned}\tag{5.19}$$

Notice we used the property of an operator acting from the right being the complex conjugate acting from the left? This form is sometimes a useful way of calculating the variance.

Example 5.4.2.

What is the result of acting \hat{p}_x on a 1-D Gaussian wave function $\Psi(x) = e^{-x^2/2}$? Hence calculate the expectation value of p_x (using the position representation).

Answer.

$$\begin{aligned}\hat{p}_x \Psi(x) &= -i\hbar \frac{\partial}{\partial x} e^{-x^2/2} \\ &= i\hbar x e^{-x^2}\end{aligned}$$

Then,

$$\begin{aligned}\langle p_x \rangle \equiv \langle \hat{p}_x \rangle &= \int_{-\infty}^{\infty} \Psi^*(x) \hat{p}_x \Psi(x) dx \\ &= -i\hbar \int_{-\infty}^{\infty} e^{-x^2/2} \frac{\partial}{\partial x} e^{-x^2/2} dx \\ &= i\hbar \int_{-\infty}^{\infty} x e^{-x^2} dx \\ &= -\frac{i\hbar}{2} \left(e^{-x^2} \Big|_{-\infty}^{\infty} \right) \\ &= 0\end{aligned}$$

Hence the expectation value of momentum is zero. Let's look at what the variance in momentum is. First, $\Delta \hat{p}_x = \hat{p}_x - \langle \hat{p}_x \rangle = \hat{p}_x$. Now, we can find the variance defined by the expectation value of the square of $\Delta \hat{p}_x$

$$\langle \Delta \hat{p}_x^2 \rangle = \int_{-\infty}^{\infty} \Psi^*(x) \Delta \hat{p}_x^2 \Psi(x) dx \quad (5.20)$$

$$= -\hbar^2 \int_{-\infty}^{\infty} e^{-x^2/2} \frac{\partial^2}{\partial x^2} e^{-x^2/2} dx \quad (5.21)$$

$$= -\hbar^2 \int_{-\infty}^{\infty} e^{-x^2/2} \left(-e^{-x^2/2} + x^2 e^{-x^2/2} \right) dx \quad (5.22)$$

$$= \hbar^2 \left(\int_{-\infty}^{\infty} e^{-x^2} dx - \int_{-\infty}^{\infty} x^2 e^{-x^2} dx \right) \quad (5.23)$$

$$= \hbar^2 \left(\sqrt{\pi} - \frac{\sqrt{\pi}}{2} \right) \quad (5.24)$$

$$= \frac{\sqrt{\pi} \hbar^2}{2} \quad (5.25)$$

Hence the expectation value of momentum is 0, and its variance is $\frac{\sqrt{\pi} \hbar^2}{2}$

Example 5.4.3.

(As yet unfinished) Change to mom rep, and recal. (Should be $e^{-k^2/2}$) $(\int_{-\infty}^{\infty} e^{-x^2} e^{-ikx} dx = e^{-k^2/2} \sqrt{2\pi}$, then change base)

5.5 Hermitian Operators

We keep talking about operators that “represent physical quantities”, things like position and momentum. However, there are other operators which do not correspond to physical quantities. So is there anything special that a physical operator must have? There answer is yes – we need the expectation value of a physical quantity to be real. After all, it doesn’t make sense to get an imaginary position, does it? But look at the momentum operator in the position representation:

$$\hat{p}_x = -i\hbar \frac{\partial}{\partial x}$$

It’s imaginary! But this doesn’t cause a problem: we can prove the expectation value of this operator is real by showing that the complex conjugate of the expectation value is the same as the expectation value (equivalent to saying that the negative of the imaginary part is the same – which only works if the imaginary part is zero.)

$$\langle \hat{p}_x \rangle = \int \Psi^*(\mathbf{r}, t) \hat{p}_x \Psi(\mathbf{r}, t) d\mathbf{r}$$

***** (Will add soon...)

In general, the expectation value of an operator is real if the operator is **Hermitian**, which says that given an operator \hat{F} and two wave functions $\Psi_1(\mathbf{r}, t)$ and $\Psi_2(\mathbf{r}, t)$,

$$\int \Psi_1^*(\mathbf{r}, t) \hat{F} \Psi_2(\mathbf{r}, t) d\mathbf{r} = \int \Psi_2(\mathbf{r}, t) \hat{F}^* \Psi_1^*(\mathbf{r}, t) d\mathbf{r} \quad (5.26)$$

$$= \int \left(\hat{F} \Psi_1(\mathbf{r}, t) \right)^* \Psi_2(\mathbf{r}, t) d\mathbf{r} \quad (5.27)$$

Note the way we’ve used brackets in Equation 5.27 – the operator acts *only* on $\Psi_1(\mathbf{r}, t)$ in this expression. Later, this definition of Hermitian operators will seem a little less contrived, but roughly speaking it says that operating on the second wave function and integrating together is the same as operating on the first wave function instead. But the most important thing to understand is that if this property holds then the expectation value must be real. To see why, let’s take the expectation value of an arbitrary wave function $\Psi(\mathbf{r}, t)$. We’ll make use of Equation , only now $\Psi_1(\mathbf{r}, t) = \Psi_2(\mathbf{r}, t) = \Psi(\mathbf{r}, t)$.

$$\begin{aligned} \langle \hat{F} \rangle &= \int \Psi^*(\mathbf{r}, t) \hat{F} \Psi(\mathbf{r}, t) d\mathbf{r} \\ &= \int \Psi(\mathbf{r}, t) \hat{F}^* \Psi^*(\mathbf{r}, t) d\mathbf{r} \quad (\text{from Equation } 5.27) \\ &= \left(\int \Psi^*(\mathbf{r}, t) \hat{F} \Psi(\mathbf{r}, t) d\mathbf{r} \right)^* \\ &= \langle \hat{F} \rangle^* \end{aligned}$$

and hence the imaginary part of the expectation value is zero, no matter what wave function we choose.

Example 5.5.1.

Let's prove that \hat{p}_x is Hermitian:

$$\begin{aligned} \int \Psi_1^*(\mathbf{r}, t) \hat{p}_x \Psi_2(\mathbf{r}, t) d\mathbf{r} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Psi_1^*(\mathbf{r}, t) \left(-i\hbar \frac{\partial}{\partial x} \right) \Psi_2(\mathbf{r}, t) dx dy dz \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[-i\hbar \Psi_1^*(\mathbf{r}, t) \Psi_2(\mathbf{r}, t) \Big|_{-\infty}^{\infty} + i\hbar \int_{-\infty}^{\infty} \Psi_2(\mathbf{r}, t) \frac{\partial \Psi_1^*(\mathbf{r}, t)}{\partial x} dx \right] dy dz \\ &= \int \Psi_2(\mathbf{r}, t) \left(i\hbar \frac{\partial}{\partial x} \right) \Psi_1^*(\mathbf{r}, t) d\mathbf{r} \quad (\text{First term vanishes}) \\ &= \int \Psi_2(\mathbf{r}, t) \hat{p}_x^* \Psi_1^*(\mathbf{r}, t) \end{aligned}$$

which is the condition for a Hermitian operator. Note that the first term vanishes because the wave functions must go to zero at infinity. Thus \hat{p}_x is Hermitian, and its expectation value is real as we calculated before.

5.6 The Commutator Bracket

We mentioned before that $\hat{A}\hat{B}$ does not necessarily do the same as $\hat{B}\hat{A}$. A measure of how much $\hat{A}\hat{B} \neq \hat{B}\hat{A}$ is given by the **commutator bracket**, which is commonly referred to as the **commutator**, and is defined by

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \quad (5.28)$$

Notice that if $\hat{A}\hat{B} = \hat{B}\hat{A}$ then $[\hat{A}, \hat{B}]$ is zero, and we say that the two operators commute. Two operators commute if their commutator is zero, and visa versa. If they don't commute, then the commutator will be non zero, and is in fact itself an operator, because it's just a combination of two other operators. It would act on a wave function like

$$[\hat{A}, \hat{B}] \Psi(\mathbf{r}, t) = \hat{A}(\hat{B}\Psi(\mathbf{r}, t)) - \hat{B}(\hat{A}\Psi(\mathbf{r}, t))$$

where we've put in brackets to show the order of operation: in the first term, you operate \hat{B} first, and then operate \hat{A} on the resulting function.

What sort of operators commute, and what does the commutator look like if they don't? First of all, the commutator of the same operator is always zero:

$$[\hat{A}, \hat{A}] = \hat{A}\hat{A} - \hat{A}\hat{A} = 0$$

Now, let's derive some important relations between the momentum and position operators, which we'll need for the rest of the course, but are good examples of the way we use commutators.

Firstly, \hat{x}, \hat{y} and \hat{z} all commute. The proof is almost “trivial”, but by acting the commutator on a dummy function, we can see that the commutator is indeed zero.

$$\begin{aligned}
 [\hat{x}, \hat{y}] \Psi(\mathbf{r}) &= \hat{x}\hat{y}\Psi(\mathbf{r}) - \hat{y}\hat{x}\Psi(\mathbf{r}) \\
 &= xx\Psi(\mathbf{r}) - yx\Psi(\mathbf{r}) \quad (\hat{x} = x) \\
 &= yx\Psi(\mathbf{r}) - yx\Psi(\mathbf{r}) \quad (xy = yx \text{ in normal algebra}) \\
 &= 0
 \end{aligned}$$

And because this is zero no matter what $\Psi(\mathbf{r})$ we choose, this means that the commutator must be the zero operator. Because the \hat{x} and \hat{y} operators in the position representation are just variables which don't affect each other, order doesn't matter when operating them and hence the commutator must be zero. Similarly, we could show that the other position co-ordinates commute.

We can also show that the momentum operators all commute, by repeating the above procedure:

$$\begin{aligned}
 [\hat{p}_x, \hat{p}_y] \Psi(\mathbf{r}) &= \hat{p}_x\hat{p}_y\Psi(\mathbf{r}) - \hat{p}_y\hat{p}_x\Psi(\mathbf{r}) \\
 &= -i\hbar \left(\frac{\partial}{\partial x} \frac{\partial}{\partial y} \Psi(\mathbf{r}) - \frac{\partial}{\partial y} \frac{\partial}{\partial x} \Psi(\mathbf{r}) \right) \quad (\text{Taking operators in } \mathbf{r}\text{-rep}) \\
 &= -i\hbar \left(\frac{\partial}{\partial x} \frac{\partial}{\partial y} \Psi(\mathbf{r}) - \frac{\partial}{\partial y} \frac{\partial}{\partial x} \Psi(\mathbf{r}) \right) \quad \left(\frac{\partial^2}{\partial x \partial y} = \frac{\partial^2}{\partial y \partial x} \right) \\
 &= 0 \\
 \Rightarrow [\hat{p}_x, \hat{p}_y] &= 0
 \end{aligned}$$

Of course, if we used a function of momentum as our dummy function, then $\hat{p}_x \equiv p_x$, which would make this calculation look just like the previous example with the position operators, only with p_x instead of x and so on. Thus we have an important point – the commutator is representation independent.³ This fits with our idea of operators as representing physical concepts without actually describing how to use them in a particular representation. It means that we can just say $[\hat{p}_x, \hat{p}_y] = 0$ and this implies that $[p_x, p_y] = 0$ in the momentum representation and $[-i\hbar \frac{\partial}{\partial x}, i\hbar \frac{\partial}{\partial y}] = 0$ in the position rep, and so on for any representation we choose to use.

What about position and momentum for *different* co-ordinates? Let's check whether \hat{x} and \hat{p}_y commute:

$$\begin{aligned}
 [\hat{x}, \hat{p}_y] \Psi(\mathbf{r}) &= \hat{x}\hat{p}_y\Psi(\mathbf{r}) - \hat{p}_y\hat{x}\Psi(\mathbf{r}) \\
 &= -i\hbar x \frac{\partial}{\partial y} \Psi(\mathbf{r}) + i\hbar \frac{\partial}{\partial y} (x\Psi(\mathbf{r})) \\
 &= -i\hbar x \frac{d\Psi(\mathbf{r})}{dy} + i\hbar \frac{\partial \Psi(\mathbf{r})}{\partial y} \\
 &= 0
 \end{aligned}$$

³Note we haven't actually proved this here, just demonstrated that it works in one case.

and similarly for all other combinations. Therefore, operators corresponding to *different* co-ordinates do commute.

Now let's see an example of operators that don't commute: \hat{p}_x and \hat{x} **do not** commute – we can see this by again taking a trial function:

$$\begin{aligned}
 [x, p]_x \Psi(\mathbf{r}) &= \hat{x}\hat{p}_x\Psi(\mathbf{r}) - \hat{p}_x x\Psi(\mathbf{r}) \\
 &= -i\hbar x \frac{\partial}{\partial x} \Psi(\mathbf{r}) + i\hbar \frac{\partial}{\partial x} (x\Psi(\mathbf{r})) \\
 &= -i\hbar x \frac{\partial \Psi(\mathbf{r})}{\partial x} + i\hbar \Psi(\mathbf{r}) + i\hbar x \frac{\partial \Psi(\mathbf{r})}{\partial x} \quad (\text{Product rule}) \\
 &= i\hbar \Psi(\mathbf{r})
 \end{aligned}$$

Therefore,

$$[x, p_x] \Psi(\mathbf{r}) = i\hbar \Psi(\mathbf{r})$$

and since this holds on matter what $\Psi(\mathbf{r})$ we choose, it must mean that

$$[x, p_x] = i\hbar \quad (5.29)$$

We could repeat this for the y and z components, and get the same result. Therefore, we have the **very** important result that the position and momentum operators of the *same* co-ordinate do not commute.

This perhaps seems rather counterintuitive, but such is the way of quantum. These last results have deep consequences for quantum mechanics, in relation to the ***Heisenberg Uncertainty Principle*** which you may have heard about. We won't look at this just yet, but you can see that in some way, the position and momentum (of the same coordinate) of a particle are intertwined somehow, in contrast to classical mechanics where they are separate properties.

We can combine the relations between position and momentum into one equation, writing

$$[\hat{x}_i, \hat{p}_j] = i\hbar \delta_{ij} \quad (5.30)$$

where we have used the ***Kronecker delta*** function, which is 1 if $i = j$ and zero otherwise. Written mathematically, this is

$$\begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

You'll hopefully notice the similarities between this function and the Dirac δ -function. The difference is that the Kronecker delta applies to discrete systems, while the Dirac function is for continuous systems.

The Kronecker δ -function has much the same properties as the Dirac version, only they apply to discrete functions, and that means sums not integrations. So whereas

$$\int \delta(x - x') f(x) dx = f(x')$$

we have that

$$\sum_i \delta(i - j) t_i = t_j$$

Take a moment to understand this new function, as we will be making great use of both δ -functions very soon.

5.7 Eigenfunctions and Eigenvalues

This section is so important it should almost have a chapter of its own. It provides the basis for much of the rest of the course, so take the time to work through it slowly. This whole section begins by asking a simple question: given any physical quantity (an operator), do there exist states which are *pure states* of that operator, that is, states where the operator will only ever take one value? We already know them for the position and momentum operators: in the \mathbf{r} -representation, they are $\delta(x - x')$ and $e^{i(\mathbf{p}\cdot\mathbf{r}-\omega t)/\hbar}$ respectively. What about angular momentum – can we find a state where the particle has exactly one value for angular momentum? And do states exist for which there is only one value for both x and p_x ? The answer to these questions can be found by *eigenfunctions*.

Let's say we're measuring some physical quantity \hat{F} , which could be position, momentum or something else entirely, and let's assume that there does indeed exist a state $\Psi_F(\mathbf{r})$ in which \hat{F} only takes one value, say F (note the lack of a hat – F is just a number⁴.) What can we say about this state? The average of measurements (of \hat{F}) made would have to be F (because we can't measure anything else!) which means mathematically that $\langle \hat{F} \rangle = F$. The variance of the operator in this state would be zero, because every measurement lies exactly on the mean, so

$$\langle \Delta \hat{F}^2 \rangle = 0$$

Then, from equation Equation 5.19, we can write the expectation value in terms of $\Psi_F(\mathbf{r}, t)$, like

$$\langle \Delta \hat{F}^2 \rangle = \int |\Delta \hat{F} \Psi_F(\mathbf{r}, t)|^2 d\mathbf{r}$$

Note that the integrand is always positive, but the integral equals zero. This could only happen if the integrand itself is zero everywhere! (If the integrand could be negative, then a zero area might just mean that the area of the graph below the x-axis cancels with the area above. But because $|\Delta \hat{F} \Psi_F(\mathbf{r}, t)|^2$ is positive, this can't happen, so the area must physically be zero.) Thus we can say

$$\begin{aligned} |\Delta \hat{F} \Psi_F(\mathbf{r}, t)|^2 &= 0 \\ \Delta \hat{F} \Psi_F(\mathbf{r}, t) &= 0 \\ (\hat{F} - \langle \hat{F} \rangle) \Psi_F(\mathbf{r}, t) &= 0 \end{aligned}$$

⁴Unfortunately, this sort of notation is convention, and until you get used to it it can be somewhat confusing.

Therefore, writing $\langle \hat{F} \rangle = F$ and, just for convenience, dropping the (\mathbf{r}, t) from Ψ_F we have

$$\hat{F}\Psi_F = F\Psi_F \quad (5.31)$$

The first \hat{F} is the operator representing some physical quantity. The subscripts denote wave functions that represent states where \hat{F} takes only one value, namely F , which is just a number. What this equation says is that operating \hat{F} on Ψ_F doesn't give a completely different wave function. Instead, it just multiplies Ψ_F by F , the value we said was the only possible measurement for \hat{F} .

Equation 5.31 is called the *eigenvalue equation* or the *eigenfunction equation*. You should have heard these terms before, when you tried to find the eigenvalues and eigenvectors of matrices. In those problems, you would have been given a matrix A , and told to find a vector \mathbf{x} and a number λ (the eigenvector and eigenvalue respectively) such that

$$A\mathbf{x} = \lambda\mathbf{x},$$

meaning that “operating” A on this vector returns a vector that, although is a different length, points in the same directions as \mathbf{x} . The idea in quantum mechanics is very similar – Ψ_F is called the *eigenfunction of F* , or an *eigenstate*. When we operate \hat{F} on it, we don't get a completely different function. Instead, the resulting function looks qualitatively exactly the same, except it is stretched or shrunk in height. If you think about it in terms of probability, the relative probability of all the measurements is the same, with peaks in all the right places, it's just that the function isn't normalised anymore.

Although the terms are used interchangeably, we generally say eigenfunction when talking about the wave function but eigenstate when describing a particle, like “the particle was in an eigenstate of momentum” as opposed to “we sum up all the eigenfunctions of momentum”.

Our motivation for this discussion was to find if there were states such that \hat{F} only ever gave one value (e.g. one specific position). We now know these states are the eigenfunctions of \hat{F} , and to find them we have to somehow solve the eigenvalue equation. Doing this would give us not only Ψ_F but its corresponding eigenvalue as well. Let's see an example of this in action.

Example 5.7.1.

*Let's find the eigenfunctions and eigenvalues of the x -component of momentum, that is, states in which measuring the momentum of the particle in the x -direction would only result in one value. For simplicity, let's just consider a 1D particle, ignore the time dependence, and let's work in the position representation, so that $\Psi(\mathbf{r}, t) \equiv \Psi(x, t)$. (You'll notice we do this quite a bit, particularly regarding the time dependence. That's because, as we'll see later, time dependence is derived from what's known as the **Shrodinger Equation**)*

Assume that such a state exists, say Ψ_{p_x} and let's say that the particular momentum value is p_x . The eigenvalue equation (Equation 5.31) then becomes

$$\hat{p}_x\Psi_{p_x} = p_x\Psi_{p_x}$$

In the position representation, the momentum operator is given by $-i\hbar\frac{\partial}{\partial x}$. Therefore, the equation above becomes

$$-i\hbar\frac{\partial\Psi_{p_x}}{\partial x} = p_x\Psi_{p_x}$$

This is a differential equation, which we can solve for Ψ_{p_x} by rearranging and integrating:

$$\begin{aligned}\frac{\partial\Psi_{p_x}}{\Psi_{p_x}} &= \frac{p_x}{-i\hbar}\partial x \\ \ln\Psi_{p_x} &= \frac{ip_x x}{\hbar} + c \\ \Psi_{p_x} &= Ae^{ip_x x/\hbar}, \quad \text{where } A = e^c\end{aligned}$$

where, technically, A could depend on time.

Does this look familiar? It should! It's just like the pure momentum state we discussed in Postulate 2 (Section 3.3) except that we've taken the time dependent part into the co-efficient, similar to Equation 4.26. And, of course, this result is what we'd expect, almost by definition: the states for which momentum can have only one value are the pure momentum states. Note also that there are no restrictions on the eigenvalue p_x – choose any real number, and we can find an eigenstate with that value of momentum.

This means that just by considering eigenvalues we've been able to rederive at least part of Postulate 2. In fact, at the end of this chapter we will be able to do away with most of the first three Postulates that we've introduced so far, and replace them with a more powerful single Postulate.

The only thing missing is the normalisation constant $A(t)$, but we'll leave that until Example 5.11.2.

5.8 Interpreting the Eigenvalues

This is just a short section, to try and justify why we devote almost the whole rest of this chapter to eigenvalues. The reason is that the eigenvalues are not just mathematical curiosities – they in fact represent the *actual physical measurements* than one can obtain when doing experiments to measure that operator. Just think on this a moment – if we calculate the eigenvalues of an operator, like position say, then these values, and only these values, are the ones that we could get if we measured the position of a particle.

As an example, when we calculated the momentum eigenvalues above, we found that any (real) number was an eigenvalue of momentum. So this means that we could potentially make a particle with any value of momentum – exactly what we'd expect. A free particle (one that isn't constrained by walls and so on) has the ability move in any direction and any speed. It doesn't, however, mean that given any particle we could find it to have any value of momentum. We

still need to consider the *probability* of obtaining each of those measurements, and some of these probabilities could be zero, depending on the wave function.

Perhaps all this doesn't seem particularly Earth-shattering as yet. But if we found a situation where we could only make eigenfunctions for certain values of an operator, then these would be the only values for that physical quantity that a particle could have – which is **quantisation**. We've already seen an example of this – Bohr suggested that the angular momentum for an electron in a Hydrogen atom was quantised. So let's investigate what the eigenvalues of angular momentum are.

Example 5.8.1.

Let's try and find the eigenvalues of angular momentum. Let's say that we have a particle moving in a circle, like an electron orbiting around an atom. In Cartesian co-ordinates, the angular momentum is $\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$. If the particle is orbiting in the $x - y$ plane, then the angular momentum will point up in the z direction, so let's just try and find the eigenvalues of the z -component of $\hat{\mathbf{L}}$. It turns out that it is far easier to calculate this in spherical polar co-ordinates (r, θ, ϕ) such that

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

In these co-ordinates, ϕ is like the angle around the circle, so that when the particle moves from ϕ to $\phi + 2\pi$ it's back where it started.

It is possible (but messy!) to show that the z -component of angular momentum can be written

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial \phi}$$

We won't do the derivation here, because it really is just maths, and it really isn't worth it.

Using this definition, our eigenvalue equation is

$$-i\hbar \frac{\partial}{\partial \phi} \Psi_{L_z}(\phi) = L_z \Psi_{L_z}(\phi) \quad (5.32)$$

Notice that here we are writing our wave function as a function of ϕ . We have changed to a different representation, one still based on position, but where we use spherical polar co-ordinates. We know that the angular momentum is zero in the other directions, so this is the only equation we need to consider.

Solving that equation give us

$$\Psi_{L_z}(\phi) = A e^{i \frac{L_z}{\hbar} \phi} \quad (5.33)$$

*However, unlike when we did the momentum eigenvalues, we have a **constraint** on this solution: we must have that $\Psi_{L_z}(\phi) = \Psi_{L_z}(\phi + 2\pi)$, because*

both correspond to the same position. This means that

$$\begin{aligned} Ae^{i\frac{L_z}{\hbar}\phi} &= Ae^{i\frac{L_z}{\hbar}(\phi+2\pi)} \\ e^{i\frac{L_z}{\hbar}\phi} &= e^{i\frac{L_z}{\hbar}\phi} e^{i\frac{L_z}{\hbar}2\pi} \\ 1 &= e^{i\frac{L_z}{\hbar}2\pi} \\ \frac{L_z}{\hbar}2\pi &= 2m\pi, \quad m = \dots, -2, -1, 0, 1, 2, \dots \end{aligned}$$

The last follows because $e^{i\beta}$ equals 1 when β is an integer multiple of 2π , i.e. $\beta = 2\pi m$. (This comes from **Euler's formula** which says $e^{i\theta} = \cos\theta + i\sin\theta$

Thus, we have the restriction that

$$L_z = m\hbar \tag{5.34}$$

where m is an integer. So we say the eigenvalues are $L_{z_m} = m\hbar$, and the eigenfunctions are $\Psi_m(\phi) = Ae^{im\phi}$. (We dropped the L_z from Ψ because it's just too many subscripts!)

And there, lo and behold, is the quantisation that Bohr proposed for an electron in a Hydrogen atom (see Section 2.6). So, finally, we are getting back the results that were suggested rather ad hoc to fit observations, now fully justified by quantum theory, if we assume that the eigenvalues are really the only values that we could get as measurements, but we'll get to that soon enough. We also need to fix the normalisation constant, but we'll do that in Example ??.

Note again that when we measure the angular momentum of a specific particle, we won't necessarily be able to obtain all of these values. The actual wave function of the particle determines which of these eigenvalues we might actually get for measurements, and the relative probabilities. If the wave function was a sum of two of those eigenstates, then we would expect to obtain those two eigenvalues with equal probability when making our measurements. This is just like when we added together pure momentum states, way back in Section 3.4. This suggests then, that an arbitrary state of the particle could be made by adding together *all* of these wave functions, which is exactly the case, and we'll explore this in Section 5.11.3.

We're now going to explore some of the properties of eigenvalues, and eventually we will discuss *why* they actually represent the physical measurements. Hopefully, these brief examples will be enough to justify doing all the maths we're about to delve into.

5.9 The Eigenvalue Spectrum

We've so far seen two possible cases for the eigenvalues – one where any value is allowed, and one where only discrete values were possible. There are in fact four types of eigenvalue solutions possible:

1. The Continuous Spectrum

This arises when there are no bounds on the operator, like a particle being able to be at any position, or have any momentum. We can then find a Ψ_F that will satisfy the eigenvalue equation

$$\hat{F}\Psi_F = F\Psi_F$$

for every possible eigenvalue.

This is the situation with the eigenvalues of momentum and position.

2. The Discrete Spectrum

When an operator has a discrete spectrum, it means the eigenvalue equation has solutions only for particular values of F . This means that those values are the only possible measurements that an operator \hat{F} can give. We say then that the physical quantity that corresponds to \hat{F} is *quantised*.

When we have a discrete spectrum, say of \hat{F} , we usually label the eigenfunctions by Ψ_{F_i} and the eigenvalues by F_i , where i is an integer, possibly negative as well as positive, depending on the situation. We found the angular momentum was quantised, and we labeled its eigenvalues by $L_{z_m} = m\hbar$. You can think about these eigenvalues as functions of m – given an m , the eigenvalue is $m\hbar$.

3. The Mixed Spectrum

The mixed spectrum is a combination of the two above. Below a certain value of F , the spectrum is discrete, but above F , we can suddenly take any value. Perhaps this sounds a little strange, but there is an example which should be familiar to everyone – electron levels in an atom. Initially, the electron has well defined energy levels – they're quantised. But once you reach a certain energy, the atom ionises, and the electron becomes a free particle, and can have any energy it wants.

Of course, you could also have a mixed spectrum the other way round, with a continuous spectrum below a certain level, and then discrete beyond that, but this doesn't occur frequently in physical situations.

4. The Band Spectrum

The final possibility is that the solutions only exist in certain continuous bands, as per Figure ***.

This spectrum occurs when we look at the motion of electrons in conductors and semi-conductors. It represents a case where particles can only exist in pockets of space, but while there, they're free to move as they wish.

5.10 Degeneracy

Sometimes, we find that a given eigenvalue might have more than one eigenfunction associated with it, meaning that the eigenvalue equation is satisfied with

two distinct functions, but the same eigenvalue ⁵. This is *degeneracy*, and we say the eigenvalues where this occurs are *degenerate*. If an eigenvalue has two eigenfunctions, we say it is 2-fold degenerate. If it had g eigenfunctions, it would be g -fold degenerate. We label the eigenfunctions as Ψ_{Fl} , or perhaps Ψ_{il} if we have a discrete spectrum, where l is an integer between 1 and g . Formally, we can write this as

$$\hat{F}\Psi_{Fl} = F\Psi_{Fl}, \quad 1 \leq l \leq g \quad (5.35)$$

We won't deal with degenerate eigenfunctions too much in this course, but they are common in many applications of quantum mechanics. For instance, the energy levels of an atom are degenerate, which means that given a certain energy, there are lots of states an electron can be in. Many of the things we derive in the following sections are only for non-degenerate eigenvalues.

Example 5.10.1.

For a particle restricted to move in one dimension, the eigenvalues of the operator $\frac{\hat{p}_x^2}{2m}$, which corresponds kinetic energy, are 2-fold degenerate.

First of all, any eigenfunction of the momentum operator \hat{p}_x will also be an eigenfunction of \hat{p}_x^2 . To show this, let Ψ_{p_x} be an eigenfunction of \hat{p}_x . Then,

$$\begin{aligned} \hat{p}_x^2 \Psi_{p_x} &= \hat{p}_x (\hat{p}_x \Psi_{p_x}) \\ &= \hat{p}_x (p_x \Psi_{p_x}) \\ &= p_x (\hat{p}_x \Psi_{p_x}) \\ &= p_x^2 \Psi_{p_x} \end{aligned}$$

This means that the eigenfunctions of \hat{p}_x are also eigenfunctions of \hat{p}_x^2 , but with eigenvalue p_x^2 . The degeneracy comes from the fact that if we considered the eigenfunction corresponding to the negative momentum, it would also be an energy eigenfunction with eigenvalue p_x^2 . Hence we have two distinct eigenfunctions with the same eigenvalue, and hence the eigenfunctions of energy are 2-fold degenerate.

Exercise 5.10.1.

Verify that if Ψ_{-p_x} is an eigenfunction of momentum such that $\hat{p}_x \Psi_{-p_x} = -p_x \Psi_{-p_x}$, then it is also an eigenfunction of \hat{p}_x^2 with eigenvalue p_x^2 .

Degeneracy is important when are trying to determine what state a particle is in. Say we know that a particle is in an eigenstate of a particular operator. For operators without degeneracy, like position, once we've measured the position of a particle, we know exactly what state it's in (the corresponding eigenstate). However, if we were measuring something like energy which is 2-fold degenerate (for a one dimensional particle), then the particle could be traveling either to the left or to the right (corresponding the negative and positive momentum

⁵By distinct functions we mean they differ by more than a scalar multiple, because any scalar multiple of an eigenfunction is also an eigenfunction (check it!). However, we always choose a scalar that normalises the eigenfunction to unity.

values). Hence, we don't know exactly the state of the particle without doing more measurements.

5.11 Properties of Eigenfunctions

Now that you've been thoroughly inundated with concepts, let's throw a few more in. We now look at the properties that these eigenfunctions and eigenvalues have.

5.11.1 Eigenvalues of Hermitian Operators

Firstly, recall that in Section 5.4 we proved that the expectation value of a Hermitian operator were real. Now, we want to show that the eigenvalues of a Hermitian operator are also real. We would *hope* they were, because if (as we've claimed) the eigenvalues are the possible measurements, we don't want to get an imaginary momentum!

To prove this, we start with the eigenvalue equation, and take its complex conjugate:

$$\begin{aligned}\hat{F}\Psi_i &= F_i\Psi_i \\ \hat{F}^*\Psi_i^* &= F_i^*\Psi_i^*\end{aligned}$$

Next, we multiply the first equation by Ψ_i^* and the second by Ψ_i so that they start to look like the terms we used to define a Hermitian operator. Then we integrate over all space, and subtract the two, giving

$$\int \Psi_i^* \hat{F} \Psi_i d\mathbf{r} - \int \Psi_i \hat{F}^* \Psi_i^* d\mathbf{r} = (F_i - F_i^*) \int \Psi_i^* \Psi_i d\mathbf{r}$$

By definition of a Hermitian operator, the left hand side must be zero. Also, $\int \Psi_i^* \Psi_i d\mathbf{r}$ won't be zero, so we have

$$(F_i - F_i^*) \int \Psi_i^* \Psi_i d\mathbf{r} = 0$$

which implies $F_i - F_i^* = 0$. Thus, the imaginary part must be zero, and hence F_i is real.

5.11.2 Orthogonality of Eigenfunctions

We now will show that eigenfunctions corresponding to different eigenvalues of a Hermitian operator are *orthogonal*. And because we can (almost always) normalise our wave functions, we can make them *orthonormal* (i.e. orthogonal and normalised)

We can do just like we did above, only with a different index in the complex conjugate equation:

$$\hat{F}\Psi_i = F_i\Psi_i \quad \hat{F}^*\Psi_j^* = F_j^*\Psi_j^*$$

Following the same working as above, we find that

$$(F_i - F_j) \int \Psi_j^* \Psi_i d\mathbf{r} = 0$$

So if we are talking about different eigenvalues, i.e. $i \neq j$, this must mean that

$$\int \Psi_j^* \Psi_i d\mathbf{r} = 0 \quad (5.36)$$

We say that this means that the eigenfunctions are **orthogonal**. In Chapter ?? we will discuss what this really means, but for now, just treat it as a name – if the functions are orthogonal, then Equation 5.36 holds.

Also, because we always take our wave functions to be normalised (i.e. $\int \Psi^*(\mathbf{r}, t) \Psi(\mathbf{r}, t) d\mathbf{r} = 1$), we say that the eigenfunctions are **orthonormal** – orthogonal and normalised. These functions then have the nice property that

$$\int \Psi_j^* \Psi_i d\mathbf{r} = \delta_{ij} \quad (5.37)$$

where we have again used the Kronecker delta function δ_{ij} , which you hopefully recall is 1 if $i = j$ and 0 otherwise.

Example 5.11.1.

?? In Example ??, we found the eigenvalues of \hat{L}_z to be

$$\Psi_m(\phi) = A e^{im\phi}$$

where A was a constant that might depend on time. Let's find the what A must be, but let's not worry about time dependence.

We need the wave function to be normalised, so we must have

$$\int \Psi_m^*(\phi) \Psi_m(\phi) d\phi = 1$$

Here, the integration over all space means integrating ϕ from 0 to 2π . So,

$$\begin{aligned} \int \Psi_m^*(\phi) \Psi_m(\phi) d\phi &= \int_0^{2\pi} A^* A e^{-im\phi} e^{im\phi} d\phi \\ &= |A|^2 \int_0^{2\pi} d\phi \\ &= 2\pi |A|^2 \end{aligned}$$

Thus, $A = (2\pi)^{-\frac{1}{2}}$, and so our full normalised eigenfunctions are

$$\Psi_m(\phi) = \frac{1}{\sqrt{2\pi}} e^{im\phi} \quad (5.38)$$

Let's check then that these wave functions are orthonormal. The normal part is easy - we just showed it. Let's check the orthogonality, by considering

$$\int_0^{2\pi} \Psi_{m'}^*(\phi) \Psi_m(\phi) d\phi = (2\pi)^{-1} \int_0^{2\pi} e^{i(m-m')\phi} d\phi$$

If $m = m'$, then

$$\begin{aligned} (2\pi)^{-1} \int_0^{2\pi} e^{i(m-m')\phi} d\phi &\equiv \int_0^{2\pi} d\phi \\ &= (2\pi)^{-1} (2\pi) \\ &= 1 \end{aligned}$$

which is our normalisation condition. Otherwise, the integral will vanish, for the same reason as the integral did when we derived the delta function. Therefore, we have

$$\int_0^{2\pi} \Psi_{m'}^*(\phi) \Psi_m(\phi) d\phi = \delta_{m' m} \quad (5.39)$$

which is the orthonormal condition we required.

Example 5.11.2.

Let's try and find the normalisation factor for the eigenstates of the x -component of momentum, which have the form

$$\Psi_{p_x}(x) = Ae^{ip_x x/\hbar}$$

If we just went ahead and integrated, we would hit a problem:

$$\begin{aligned} \int \Psi_{p_x}^*(x) \Psi_{p_x}(x) dx &= \int A^* e^{-ip_x x/\hbar} \cdot Ae^{ip_x x/\hbar} \\ &= A^* A \int e^{i(p_x - p_x)x/\hbar} dx \\ &= |A|^2 \int dx \end{aligned}$$

which is infinity!

Instead, we use a clever trick. We consider integrating two different eigenstates together, and we'll assume that the normalisation constant is just that - a constant, that's the same for all eigenfunctions.

$$\begin{aligned} \int \Psi_{p_x}^*(x) \cdot \Psi_{p_x}(x) dx &= \int A^* e^{-} \\ &= \text{*****} (Willdothislater) \end{aligned}$$

5.11.3 Completeness

When we say that a set of wave functions is complete, we mean that any other wave function can be written as a sum of these. This is very much like the idea of a basis of vectors – we say a set of vectors are complete, that is, they form a basis, if any other vector can be written as a linear combination of them. Thus $(1, 0)$ and $(0, 1)$ form a basis for \mathbf{R}^2 , as do $(1, 2)$ and $(3, 4)$. Both pairs of those vectors are complete, in that any other vector (of the same type, in this case two dimensional) can be written as a combination of those.

The set of wave functions we are most interested in is the eigenfunctions of some operator. Let's look at the discrete case first, where the eigenfunctions of an operator \hat{F} are labeled by $\Psi_i(\mathbf{r})$, and the associated eigenvalues are F_i . For them to be complete, we must be able to write

$$\Psi(\mathbf{r}) = \sum_i c_i \Psi_i(\mathbf{r}) \quad (5.40)$$

where the sum is over all permissible values of i , and the c_i 's are (potentially) complex numbers. This is just a formal way of writing what we discussed with vectors above – that every wave function can be written as a linear combination of the set of states we're proposing form a basis.

If the spectrum is continuous, we suggest the wave function could be written as

$$\Psi(\mathbf{r}) = \int c_F \Psi_F(\mathbf{r}) dF, \quad (5.41)$$

that is, we integrate over every possible eigenvalue, just like the sum in the discrete case. Although we're using our shorthand integration symbol (that is, we're not putting any limits on it) it doesn't necessarily have to be from negative infinity to positive infinity. The eigenvalues might only lie in a particular range, like 0 to 2π in the case of the z -component of angular momentum. Just keep this in mind when evaluating these integrals. Note also that the co-efficients c_F are now continuous – there is one c for every value of F , the eigenvalue, as opposed to the c_i 's where the i didn't directly represent the eigenvalue.

We're going to assume the more general case that the operator \hat{F} has a mixed spectrum, with both a discrete part, with eigenfunctions labeled by $\Psi_i(\mathbf{r})$, and a continuous part with eigenfunctions $\Psi_F(\mathbf{r})$, and prove that any wave function

$$\Psi(\mathbf{r}) = \sum_i c_i \Psi_i(\mathbf{r}) + \int c_F \Psi_F(\mathbf{r}) dF, \quad (5.42)$$

To prove this, we're going to use something that might at first appear to be circular logic! We're going to pretend that it's true, and then show that we can find co-efficients c_i that satisfy it. To do this, we multiply our expression above

by $\Psi_j^*(\mathbf{r})$, and integrate as follows:

$$\begin{aligned}
 \int \Psi_j^*(\mathbf{r})\Psi(\mathbf{r})d\mathbf{r} &= \int \Psi_j^*(\mathbf{r}) \left(\sum_i c_i \Psi_i(\mathbf{r}) \right) d\mathbf{r} + \int \Psi_j^*(\mathbf{r}) \left(\int c(F)\Psi_F(\mathbf{r}) dF \right) d\mathbf{r} \\
 &= \sum_i c_i \int \Psi_j^*(\mathbf{r})\Psi_i(\mathbf{r})d\mathbf{r} + \int c_F \left(\int \Psi_j^*(\mathbf{r})\Psi_F(\mathbf{r})d\mathbf{r} \right) dF \\
 &= \sum_i c_i \delta_{ij} + 0 \\
 &= c_j
 \end{aligned}$$

where we have used the proof from Section 5.11.2 that wave functions are orthogonal. Note that the second term vanishes completely, because $\Psi_F(\mathbf{r})$ and $\Psi_i(\mathbf{r})$ are always orthogonal – they belong to completely different parts of the spectrum!

So we find that

$$c_j = \int \Psi_j^*(\mathbf{r})\Psi(\mathbf{r})d\mathbf{r} \quad (5.43)$$

Similarly, to find the continuous co-efficients, we do just like we did before, except multiplying this time by $\Psi_{F'}^*$, the complex conjugate of one of the continuous eigenfunctions:

$$\begin{aligned}
 \int \Psi_{F'}^*(\mathbf{r})\Psi(\mathbf{r})d\mathbf{r} &= \int \Psi_{F'}^*(\mathbf{r}) \left(\sum_i c_i \Psi_i(\mathbf{r}) \right) d\mathbf{r} + \int \Psi_{F'}^*(\mathbf{r}) \left(\int c(F)\Psi_F(\mathbf{r}) dF \right) d\mathbf{r} \\
 &= \sum_i c_i \int \Psi_{F'}^*(\mathbf{r})\Psi_i(\mathbf{r})d\mathbf{r} + \int c_F \left(\int \Psi_{F'}^*(\mathbf{r})\Psi_F(\mathbf{r})d\mathbf{r} \right) dF \\
 &= 0 + \int c_F \delta_{F'} - F dF \\
 &= c_{F'}
 \end{aligned}$$

Note we've used the Filtering property of the Dirac δ -function at the end there. c_F is really more like $c(F)$, an actual function, but we write the F as a subscript for convenience.

Thus,

$$c_{F'} = \int \Psi_{F'}^*(\mathbf{r})\Psi(\mathbf{r})d\mathbf{r} \quad (5.44)$$

If the logic is a little confusing, sit and think about it for a little while – we've shown that there exist co-efficients such that we can expand any wave function in that way, which is enough to show that we *can* expand it.

To sum up, the important result here is that the eigenfunctions of any operator are complete, and therefore, in some sense, form a basis for all the possible wave functions that we could come up with, that are of the same type. That is to say, the momentum eigenstates form a basis for all wave functions that are *square integrable*, which means that $\int \Psi(\mathbf{r}, t)\Psi^*(\mathbf{r}, t)d\mathbf{r}$ is finite. The eigenfunctions of angular momentum form a basis for periodic functions, with period

2π . This won't really concern us too much here, but it's worth taking note of for future reference.

Have a think about what it means for the particle if it can be written as a combination of eigenstates, and we'll come back to it soon.

5.12 Expanding Expectation Values

Using the expansion of a wave function we made above, we can expand out the expectation value of an operator, in terms of its eigenvalues. For the discrete case, we have

$$\begin{aligned}
 \langle \hat{F} \rangle &= \int \Psi^*(\mathbf{r}) \hat{F} \Psi(\mathbf{r}) d\mathbf{r} \\
 &= \int \left(\sum_i c_i^* \Psi_i^*(\mathbf{r}) \right) \hat{F} \left(\sum_j a_j \Psi_j(\mathbf{r}) \right) d\mathbf{r} \\
 &= \sum_i \sum_j c_i^* c_j \int \Psi_i^* \hat{F} \Psi_j d\mathbf{r} \\
 &= \sum_i \sum_j c_i^* c_j \int \Psi_i^* F_j \Psi_j d\mathbf{r}, \quad (\text{by defn of eigenvalues}) \\
 &= \sum_i \sum_j c_i^* c_j F_j \int \Psi_i^* \Psi_j d\mathbf{r} \\
 &= \sum_i \sum_j c_i^* c_j F_j \delta_{ij}, \quad (\text{by orthogonality}) \\
 &= \sum_i |c_i|^2 F_i
 \end{aligned}$$

Basically, we've expanded the wave functions in terms of the eigenfunctions of \hat{F} , then used the eigenfunction equation to replace the operator by the observed values, and finally used the orthogonality of the eigenfunctions to simplify it.

So we have

$$\langle \hat{F} \rangle = \sum_i |c_i|^2 F_i \quad (5.45)$$

If we have a continuous spectrum, then we find that if the particle is in state $\Psi(\mathbf{r}) = \int c_F \Psi_F(\mathbf{r}) dF$ (from Section 5.11.3) then

$$\langle \hat{F} \rangle = \int |c_F|^2 F dF \quad (5.46)$$

which looks almost like Equation 5.45.

This means that we can write the expectation as a “weighted sum” of the eigenvalues. Does this sound vaguely familiar? Remember way back when we defined the expectation value as the sum of the possible measurements multiplied by their probabilities? Have a think about this for a moment, then read on.

5.13 Interpreting the Eigenfunctions

Now that we are armed with a bit more knowledge about these eigenfunctions, we are going to try and back up our suggestion from Section 5.8. There we suggested that the eigenvalues of an operator were precisely the values that we could measure for the physical quantity represented by that operator.

What justifies this assumption? There is some hand-wavy evidence, in that the eigenvalues of angular momentum are precisely the values that Bohr predicted for the quantisation of angular momentum in the hydrogen atom. However, we can do better than that! We will consider a *thought experiment* – which is where we prove something just by thinking about it! [Perhaps include a heavy/light body falling t.e. example?] This is one of the more difficult bits of quantum, so read the whole section first, then come back and work through it slowly, because we may answer your question a little later!

Imagine a particle in some state – perhaps a wave packet, or a pure momentum state, or some other wave function. We then measure the x-position of the particle, and find it to be x_0 , say. Now here’s the important bit: if we measure its position again *immediately* afterwards, the particle must still be in the same position – even in quantum mechanics, a particle can’t just jump to a new location instantaneously.

But for the particle to be guaranteed to be at x_0 when we make a measurement, it must be in an *eigenstate* of position – that’s how we defined eigenstates, as states where we could only get one value for the operator. If it wasn’t in an eigenstate then there would be some probability of getting a different value of position, but this can’t happen because we just said our particle can’t jump instantaneously. And if it’s in an eigenstate, then its position must be the corresponding eigenvalue of that state (because that’s the one value that position takes in that state).

Let’s say that again: After we measure the position (or any other value) of a particle, it must be in an eigenstate of that operator. Then because it’s in an eigenstate, the value that we measure must be the eigenvalue corresponding to that state. Therefore, the only measurements we can make are eigenvalues. Even if we don’t actually make that second measurement, all this must be true – the universe still has to allow for the *possibility* of us making another measurement, and it can’t know in advance!

No doubt you’ve realised that there is a problem here – how could the particle be in an eigenstate, when we originally said we had no idea what state the particle was in? Somehow the particle has “jumped” into a different state, simply because we measured it! This is a difficult proposition to swallow, but experiment appears to confirm it.

A piece of supporting evidence was briefly discussed at the end of Section 5.12. In that section we showed that the expectation value of an operator could be written as the weighted sum of the eigenvalues. Thinking back to Section 5.1 when we defined the expectation value, we said it was the sum of the possible measurements multiplied by their probability:

$$\langle \hat{F} \rangle = \int P(F) \cdot F \, dF$$

Compare this to Equation 5.46 above, which looks the same except for $|c_F|^2$ instead of $P(F)$. It suggests that $|c_F|^2$ is the probability that we would measure \hat{F} to be F , and then c_F must be the probability amplitude. Does *this* now make sense?

To see that it does, we can make use of the completeness of the eigenfunctions, from Section 5.11.3. There we said that, in the continuous case,

$$\Psi(\mathbf{r}) = \int c_F \Psi_F(\mathbf{r}) \, dF$$

Just like when we were talking about making wave packets out of the pure momentum states, the c_F coefficients determine “how much” each eigenstate of the operator contributes to the wave function. The higher they are, the more that state contributes. And we expect that the more prominent a state is, the higher the probability that we measure its corresponding value. For example, we now know the pure momentum states to be the eigenstates of momentum, and we suggested that the $A(k)$ coefficient in our wave packet was the probability amplitude of the particle having that wave number. $A(k)$ corresponds to c_F , which is just what we said above!

This is all consistent with the eigenvalues being exactly the values of an operator that the particle could have. Note that what we looked at above was for a continuous spectrum (the reason being that the eigenvalues of position and momentum are continuous, and that’s what we used for the first part of the course!)

5.13.1 How to Analyse a Wave Function

The important thing is that the c_F ’s tell us what eigenvalues we could actually measure given a specific wave function, and with what probabilities. There’s sort of a chain of possible values: we always have that measurements of a physical quantity (a Hermitian operator) must be real. Then, when we know what the operator looks like mathematically, we can find its eigenvalues, and the measurements must be one of the eigenvalues. Then, we look at the actual wave function, and rewrite it as a sum (or integral) of the eigenfunctions (which we can do, because they are complete). The absolute square coefficients we get tell us the probabilities of obtaining each eigenvalue. Only eigenvalues with non-zero probability would actually get measured.

Let’s have a look at a couple of examples (just in one dimension to make it easier!)

Example 5.13.1.

Find the allowed values of momentum for a particle in a pure state ($\Psi(x) = e^{ip_x x/\hbar}$), and the expectation value of momentum using eigenvalues. Then, compute the expectation value explicitly by integrating.

Answer.

The first step is to decide what representation we will work in, and what the operator in question looks like in that rep. Since we are given our state as a function of position, we'll choose the position rep where the momentum operator looks like $\hat{p}_x = -i\hbar \frac{\partial}{\partial x}$.

The second step is to find the eigenvalues and eigenfunctions of the operator. Note that the eigenvalues will be independent of what representation we choose, but the eigenfunctions (being functions!) almost certainly would look different. We've already done this, and we know the eigenfunctions are

$$\Psi_{p_x} = e^{ip_x x/\hbar}$$

We are given

$$\Psi(\mathbf{r}) = e^{ip'_x x/\hbar}$$

We know this is already an eigenstate of momentum, with eigenvalue p_x . Hence the expansion into the eigenstates is

$$\Psi(\mathbf{r}) = 1 \cdot \Psi_{p'_x}$$

So the only coefficients that are non-zero are for the eigenvalue p'_x , so this is the only value of momentum you could measure. The probability of obtaining this measurement is $|1|^2 = 1$, as we would hope. Therefore, the expectation value of momentum must be

$$\langle p_x \rangle = p'_x$$

We can check this explicitly by

$$\begin{aligned} \langle p_x \rangle &= \int \Psi^*(x) p_x \Psi(x) dx \\ &= \int_{-\infty}^{\infty} e^{-ip'_x x/\hbar} \left(-i\hbar \frac{\partial}{\partial x} \right) e^{ip'_x x/\hbar} dx \\ &= -i\hbar \int_{-\infty}^{\infty} e^{-ip'_x x/\hbar} \left(p'_x e^{ip'_x x/\hbar} \right) dx \\ &= -i\hbar \int_{-\infty}^{\infty} e^{-ip'_x x/\hbar} \left(p'_x e^{ip'_x x/\hbar} \right) dx \end{aligned}$$

5.13.2 The Copenhagen Collapse Hypothesis

However, there still remains the matter of the “jump” into an eigenstate. The most popular interpretation is the ***Copenhagen Collapse Hypothesis***, after a congress in Copenhagen adopted it. It says that the act of measuring some property of a particle causes it to jump into an eigenstate, regardless of what state it was in before. We could have a particle in a pure momentum state (where its position is completely unspecified), measure its position, and suddenly have a particle whose position is known precisely. Of course, if we were to “reset”

the particle into its initial momentum state, and measured its position again, we would (most likely) get a different eigenstate.

Note however that the particle doesn't necessarily stay in the eigenstate from then on. It might still evolve in response to its environment. What's different is that whatever the particle was doing before doesn't matter after we examine it, because by the very act of measuring the particle, we've changed its properties.

Can you think of a case where the wave function *doesn't* jump? This would happen if the particle was already in an eigenstate of that operator! In that case, the only value you can measure is the corresponding eigenvalue, the particle stays in that state, and all is good. Just don't be confused into thinking that the particle could jump into some other eigenstate of the operator.

The other interpretation, which is of great appeal to science fiction writers, is the ***Many Worlds Hypothesis***, which says that there are an infinite number of universes, which contain every possible world. By that, we mean that every time you make a decision, the universe splits into two. In one you chose the red shirt, in the other the blue. In one universe a neutron decayed, in the other it didn't. The probabilistic nature of quantum comes from the number of worlds where a particle is in each position – the more worlds, the higher the probability. When we measure the particle then, it's only in one position, hence the eigenstate, only we don't know what universe we're in until we measure it.

The idea of these interpretations is that they interpret the physical results – there's no way to prove for sure which one, if either, is the right one. There are many more theories, some of which are rather more weird. And if you think these theories, or even the whole idea of jumping into another state, sound dubious then you're not alone. Einstein himself believed that, based on this, quantum mechanics was an incomplete theory and that there exists ***hidden variables*** that actually remove the probabilistic nature of quantum. This is well summarised in his famous statement: "God does not play dice". Unfortunately for Einstein, the late British physicist John Bell devised an experiment that disproved the existence of hidden variables. So that is one interpretation, at least, that can be put to rest.

Note that the particle probably won't stay in that eigenstate for all time. Just like a classical particle, it will continue to evolve. We only say that in the instant after we measure it, it must be in an eigenstate. Is this rather strange idea supported by experiment? In fact, there are several experiments that show that measuring a particle affects its state, most notably the Stern-Gerleck experiment that we will discuss later in this course.

5.14 Bringing it together

We can now sum up the contents of this chapter into a new Postulate, which supersedes most of Postulates 2 and 3, by explaining them in even more basic terms.

Postulate 4. 1. *Every real dynamical variable is represented in quantum me-*

chanics by a linear Hermitian Operator. If the dynamical variable can be written in classical mechanics in the form $F(\mathbf{r}, \mathbf{p})$ then the rule for obtaining the position representation form of the operator is to replace \mathbf{p} by $-i\hbar\nabla$ so that $F(\mathbf{r}, \mathbf{p}) \rightarrow \hat{F}(\mathbf{r}, -i\hbar\nabla)$. When the dynamical variable has no classical counterpart (e.g. the spin of an electron) it is necessary to discover the corresponding operator by appeal to experiment combined with intelligent guesswork. These new operators then become part of the basic quantum mechanics framework.

2. The eigenvalues of the operator \hat{F} are the possible results of measurement of the dynamical variable represented by this operator.
3. When the system is in an arbitrary state $\Psi(\mathbf{r})$ (ignoring time dependence) it can be expanded in terms of the eigenfunctions of \hat{F} as follows:

$$\Psi(\mathbf{r}, t) = \sum_i c_i \Psi_i(\mathbf{r}) + \int c_F \Psi_F(\mathbf{r}) dF$$

$|c_i|^2$ is the probability of obtaining F_i as a result of a measurement of the dynamical variable and $|c_F|^2$ is the probability of measuring the dynamical variable to lie between F and $F + dF$.

Example 5.14.1.

This example will soon show how Born's interpretation of the wave function is just a special case of the above postulate - try it for yourself!