Quantum Mechanics: Level 2

Shaun Cole

April 23, 2018

Course Structure

The course text book is "Introduction to Quantum Mechanics", by Griffiths. (Pearson, 2nd edition or new international edition, **not** the 1st edition). In the notes, I will reference sections in this book as [Gri 2.2] etc, meaning section 2.2 of Griffiths. There are copies available in the University and college libraries, but you might find it useful to buy your own copy.

For historical reasons the course syllabus is specified by reference to sections of "Quantum Mechanics" by B.H. Bransden and C.J. Joachain (Prentice Hall, 2nd Edition). This has a much more mathematically formal approach and is not used in this course. What you need to know will be covered in the lecture notes and references to Griffiths.

There are 21 lectures, 9 Workshops and 8 coursework problems.

Quantum mechanics is necessarily quite mathematical and we will come across many special functions (Hermite polynomials, Spherical Harmonics, ...), but don't worry. You won't be asked to memorise them. You will be expected to be able to use them once they are given to you.

Course work is *formative*, i.e. does not contribute marks towards your degree result. They are *required* as part of keeping term and you are responsible for marking and uploading your marks. Workshops are weekly (except 1st week) and are designed to support your learning. These give an opportunity to talk through things you don't understand in small groups with your peers or with a tutor or lecturer. There are also workshop questions to give you more practice with the material. Attendance at workshops is required and again formative.

The formative work gives you responsibility - you get to choose how to spend the time and you choose which things that you think will most benefit your understanding. But its possible to think everything is OK with understanding when its not. So to catch this, we have a midterm test. These again are *formative* but they are marked by the workshop team, and they give you an external check on your progress.

There are great online resources available at The Quantum Mechanics Visualisation Project (QuVis)[1] at St. Andrews. The pdf version of my notes provides active links.

---

[1] http://www.st-andrews.ac.uk/physics/quvis

1 Basics of Quantum Mechanics

1.1 Wavefunctions and Schrödinger's equation

The double slit experiment is one of the conceptually most important experiments of quantum mechanics.

The interference pattern when both slits are open means that the electrons also have wave-like properties. The separation of the fringes gives us the wavelength of the electron waves, and we get the de Broglie relation

$$\boxed{\lambda = h/p,}$$

where p is momentum. If it is behaving like a wave then it must be described by some form of **wavefunction**, $\Psi(\mathbf{r}, t)$ - a function of position and time.

But this is quite counter intuitive. We are used to particles having well defined positions but a simple wave

$$\Psi(x) = A\cos(kx) \quad \text{with} \quad \lambda = 2\pi/k$$

extends over the entire x axis and so how do we define its position?

If we add together two waves of slightly different k (or wavelength) we get beats and more waves with spread in k produces a localised wave packet, but we reduced the spread in position only by introducing a spread in momentum $p = h/\lambda = \hbar k$.

We are not getting any closer to a classical mechanics approach where if we know $x(t)$ then we automatically know $v = dx/dt$ and $p = mv = mdx/dt$ and $E_{\text{kin}} = mv^2/2 = p^2/2m$. With a wave packet we can localise $x(t)$ at some time t, but only at the expense of not knowing p very well.

Figure 1: Electron 2-slit interference.

Figure 2: **Top:** single cosine/sine wave. **Mid:** Beats of two waves of similar k. **Bottom:** Many waves with a spread in k producing a localised wavepacket.

This is the **Heisenberg Uncertainty Principle** (better to call it indeterminacy!)

$$\Delta x \Delta p \geq \hbar/2$$

and forms an absolute limit to our knowledge. This is quite unlike classical mechanics!

Classical waves, on a string obey a wave equation. If the string is in the x direction then a transverse wave in the y direction is the solution of the 1D wave equation:

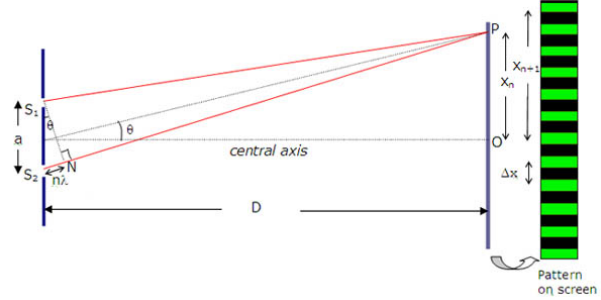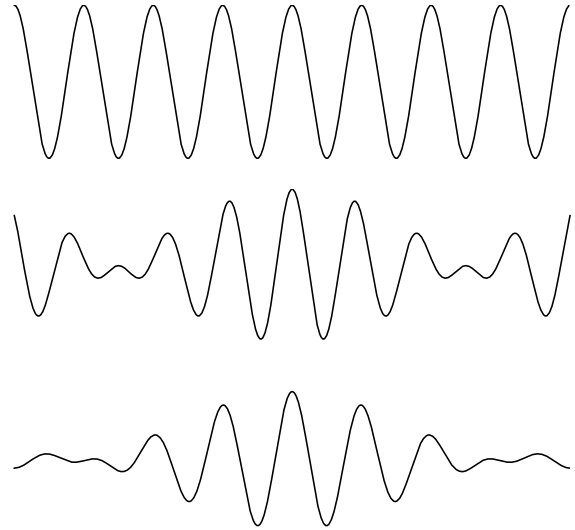$$\frac{\partial^2 y(x,t)}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 y(x,t)}{\partial t^2}.$$

A solution is a simple travelling wave $y(x,t) = A\cos(kx - \omega t)$ where $\omega = 2\pi f$ and $\lambda = 2\pi/k$ and the wave velocity $v = \lambda f = \omega/k$. This wave travels in the +ve x direction (left to right) along a string with constant velocity v irrespective of wavelength, whereas we know that in the quantum world we have $p = mv = h/\lambda$ so the speed of an electron wave depends on its wavelength.

The quantum behaviour is neither classical mechanics nor classical waves.

Quantum mechanics: Suppose our electron is described by a wavefunction Ψ, travelling along the x axis. Then for a free particle (potential $V = 0$) we could try a similar solution

$$\Psi(x,t) = A\cos(kx - \omega t) + B\sin(kx - \omega t).$$

We know that its kinetic energy $E_{\text{kin}} = \frac{1}{2}mv_x^2 = p_x^2/2m = \hbar^2 k^2/2m$. Let us assume the total energy is $E_{\text{tot}} = hf = \hbar\omega$, just as for a photon. Hence by simple energy conservation we have $\hbar\omega = \hbar^2 k^2/2m$.

We can turn this algebraic equation into a wave equation by the following procedure:

Differentiating ψ twice wrt x pulls out a factor of $-k^2$, i.e.

$$\begin{aligned}
\frac{\partial^2 \Psi(x,t)}{\partial x^2} &= \frac{\partial}{\partial x}[-Ak\sin(kx - \omega t) + Bk\cos(kx - \omega t)] \\
&= -Ak^2\cos(kx - \omega t) - Bk^2\sin(kx - \omega t) = -k^2\Psi(x,t)
\end{aligned}$$

and so we could write $E_{\text{tot}}\Psi = (\hbar^2 k^2/2m)\Psi$ as

$$E\Psi(x,t) = -\frac{\hbar^2}{2m}\frac{\partial^2 \Psi(x,t)}{\partial x^2}$$

with the spatial part looking quite like the classical wave equation.

For the time part we can pull out an ω via a single derivative wrt t

$$\frac{\partial \Psi(x,t)}{\partial t} = \omega[A\sin(kx - \omega t) - B\cos(kx - \omega t)]$$

What we would really like the term in square brackets to equal $C\Psi$

$$[A\sin(kx - \omega t) - B\cos(kx - \omega t)] = C[A\cos(kx - \omega t) + B\sin(kx - \omega t)]$$

equating coefficients of sine we get $A = CB$, while for cosine its $-B = CA$. Divide and get $-A/B = B/A$ i.e. $B^2 = -A^2$ and $B = \pm iA$. Take the +ve root and first our electron wave is **complex**

$$\Psi = A[\cos(kx - \omega t) + i\sin(kx - \omega t)] = Ae^{i(kx - \omega t)}.$$

Also, we get $C = -B/A = -i$ so $\partial \Psi/\partial t = -iE/\hbar\Psi$. Multiply both sides by $i\hbar$ and get $i\hbar\partial \Psi/\partial t = E\Psi$. Putting everything together **for a free particle in 1D** our wave equation is

$$\frac{-\hbar^2}{2m}\frac{\partial^2 \Psi}{\partial x^2} = i\hbar\frac{\partial \Psi}{\partial t} = E_{\text{tot}}\Psi$$

Since this is just simply saying $p_x^2/2m = \hbar\omega = E_{\text{tot}}$ then we can easily see how to extend it to put an external potential, $V(x,t)$, as total energy is KE+PE so $p_x^2/2m + V(x,t) = \hbar\omega = E_{\text{tot}}$

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

This is the 1D **time dependent Schrödinger equation**.

• It is linear and homogeneous so the superposition principle holds i.e. if Ψ_1 and Ψ_2 are two different solutions of the Schrödinger equation, then linear combinations $\Psi = c_1\Psi_1 + c_2\Psi_2$ are also a solutions.

• It is also first order in the time derivative, with only $\partial/\partial t$ rather than the $\partial^2/\partial t^2$ of the classical wave equation. For classical waves we need both the wave function Ψ and its first derivative $\partial\Psi/\partial t$ in order to determine its behaviour as a function of time. Here, we only need the wavefunction at some time t_0 in order to completely specify the subsequent behaviour of the system.

1.2 Interpretation of the wavefunction [Gri 1.2]

The wavefunction has complex values and so is not a wave in some substance (e.g. the aether). One view is that wave functions are computational devices that only have meaning in Schrödinger theory (this is unsatisfactory and we will return to its meaning at the end of the course).

The wave function contains all in information which the uncertainty principle allows us to know. What is this information?

Go back to the double slit experiment, but turn down the intensity of the electron source so that only one electron comes through at a time. We can see the single hits on the screen where each electron lands (showing that individual electrons really are going through the slits). But after a while, when many 'single' particles have built up a pattern, we get the characteristic pattern of interference. Interference does not occur between different electrons, but is a property of a single electron.

One single measurement does not have a predictable result - the electron hit can be anywhere in the pattern. Only when we consider many identical systems do we get the full pattern. This suggests that for an individual particle the process is of a statistical nature, so it is telling us something about the probability that the particle will hit the screen at a certain point.

Since our wavefunction is complex, and probabilities must be real, this suggests that we associate a probability $\propto |\Psi(x,t)|^2$. This would be similar to classical waves where the intensity of a wave $I \propto$ Amplitude^2. Then the probability of finding a particle between x and $x + dx$ is

$$\boxed{P(x,t)\,dx \propto |\Psi(x,t)|^2\,dx = \Psi^*(x,t)\Psi(x,t)\,dx}$$

where the position probability density $P(x,t)$ has units 1/length and the wavefunction $\Psi(x,t)$ has units $1/(\text{length})^{1/2}$. This is always real even when our wavefunction is imaginary. Suppose $\Psi = a + ib$ where a, b are real. Then $\Psi^*\Psi = (a - ib)(a + ib) = a^2 + b^2$.

Thus the interpretation of the wave function is a statistical one. we talk about the wavefunction of an individual particle but it is more useful to think about it as describing the behaviour expected from an ensemble of identical systems.

1.3 Normalisation of a wavefunction [Gri 1.4]

For a single particle, the probability of finding it somewhere in space should be unity so this gives us a normalisation

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

4

Hence if we have some unnormalised function $f(x,t)$, which is a solution of the Schrödinger equation, we can get its normalised version $\Psi(x,t)$ as $\Psi(x,t) = Nf(x,t)$, where N is some normalisation constant, using

$$\int |\Psi(x,t)|^2 \, dx = 1 \qquad \Rightarrow \qquad \int (Nf)^*(Nf) \, dx = |N|^2 \int f^* f \, dx = 1$$

$$\Rightarrow \qquad |N|^2 = \frac{1}{\int f^* f \, dx} \quad \text{or } N = \frac{1}{\sqrt{\int f^* f \, dx}}.$$

In full we have

$$\Psi(x,t) = \frac{f(x,t)}{\sqrt{\int_{-\infty}^{+\infty} |f(x,t)|^2 \, dx}}$$

1.3.1   Normalisation for free particle

Consider our 1D Schrödinger equation for a free particle $\Psi(x,t) = Ae^{i(kx-\omega t)}$. Then

$$\int_{-\infty}^{+\infty} \Psi^*(x,t)\Psi(x,t)dx = |A|^2 \int_{-\infty}^{+\infty} e^{-i(kx-\omega t)} e^{i(kx-\omega t)} dx = |A|^2 \int_{-\infty}^{+\infty} dx = \infty$$

This is problematic because a plane wave along the x axis is completely delocalised - there is constant probability to find it anywhere along the x-axis. In practice we'd confine it by the experiment to a box of length $L \gg \lambda$ (e.g. the size of the room!) so

$$|A|^2 \int_0^L e^{-i(kx-\omega t)} e^{i(kx-\omega t)} \, dx = |A|^2 L = 1$$

so then $A^*A = 1/L$ so we are free to choose this to be any (complex) number where $|A|^2 = 1/L$ e.g. $A = 1/\sqrt{L}$ or $A = -1/\sqrt{L}$ or $A = i/\sqrt{L}$ or $A = -i/\sqrt{L}$ or.... so pick the one which makes life easy, $A = 1/\sqrt{L}$. We know that we don't have a physical meaning for the wavefunction itself, only its square, so this phase doesn't matter.

Our normalised wavefunction is

$$\Psi(x,t) = \frac{1}{\sqrt{L}} \, e^{i(kx-\omega t)}.$$

The probability of finding it in any section from a to $a + \Delta a$ is then

$$\int_a^{a+\Delta a} \Psi^*(x,t)\Psi(x,t) \, dx = \frac{1}{L} \int_a^{a+\Delta a} dx = \frac{(a + \Delta a - a)}{L} = \frac{\Delta a}{L}$$

and there is equal probability of finding a particle with this wavefunction anywhere.

1.3.2   A Gaussian: $\Psi(x, t = 0) = N \exp(-ax^2/2)$

Here we are using the Gaussian just a convenient example of simple function representing a particle with a localized and so normalizable wave function.

The normalisation condition is

$$\int_{-\infty}^{+\infty} N^* e^{-ax^2/2} N e^{-ax^2/2} \, dx = N^2 \int_{-\infty}^{+\infty} e^{-ax^2} \, dx = 1$$

Wolfram alpha site lets you calculate definite integrals `http://www.wolframalpha.com/`
functions have capital letters, so here type

`integrate Exp[-a x^ 2] between -infty and infty`

in the box and hit return to get $= \sqrt{\pi/a}$ hence

$$N^2 \sqrt{\frac{\pi}{a}} = 1$$

$$N = \left(\frac{a}{\pi}\right)^{1/4}$$

so then the probability of finding the particle between $0 < x < a$ is

$$\int_0^a \Psi^* \Psi \, dx = \sqrt{\frac{a}{\pi}} \sqrt{\frac{\pi}{a}} \frac{\text{erf}(a^{3/2})}{2} = \frac{\text{erf}(a^{3/2})}{2}$$

The time evolution of such Gaussian wavepackets is illustrated at QuVis:GaussianWavePacket[2].

---

[2]`https://www.st-andrews.ac.uk/physics/quvis/simulations_html5/sims/gaussian/gaussian.html`

2 Operators and Expectation Values [Gri 1.5]

As we have a probability distribution we can not say what value we will get when we make one measurement, but we can work out the average value we will get from many repeat measurements of identical systems. In QM the average value is called the **expectation value**. To get an average, we multiply each possible value by the probability of getting that value and add them up.

2.1 Normalisation is independent of time [Gri 1.4]

To prove the normalization condition is independent of time consider its rate of change

$$\frac{d}{dt} \int \psi^* \psi \, dx = \int \frac{\partial \psi^*}{\partial t} \psi + \psi^* \frac{\partial \psi}{\partial t} \, dx.$$

But from the Schrödinger equation we know

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

and by taking the complex conjugate

$$\frac{\partial \psi^*}{\partial t} = -\frac{i\hbar}{2m} \frac{\partial^2 \psi}{\partial x^2} + \frac{i}{\hbar} V \psi^*.$$

Hence if we substitute in the terms involving V cancel and we are left with

$$
\begin{aligned}
\frac{d}{dt} \int \psi^* \psi \, dx &= \frac{i\hbar}{2m} \int \left(\psi^* \frac{\partial^2 \psi}{\partial x^2} - \psi \frac{\partial^2 \psi^*}{\partial x^2} \right) dx \\
&= \frac{i\hbar}{2m} \int \frac{\partial}{\partial x} \left(\psi^* \frac{\partial \psi}{\partial x} - \psi \frac{\partial \psi^*}{\partial x} \right) dx \\
&= \frac{i\hbar}{2m} \left[\psi^* \frac{\partial \psi}{\partial x} - \psi \frac{\partial \psi^*}{\partial x} \right]_{-\infty}^{\infty} = 0,
\end{aligned}
$$

where the last equality comes from assuming that ψ and ψ^* go to zero at $x \to \pm\infty$.

The fact that $d/dt \left(\int \Psi^* \Psi dx \right) = 0$ means we only need to normalise at $t = 0$ and then the wavefunction has the same normalization for all subsequent times.

2.2 The position operator

The expectation, or average, value of position will be denoted by angle brackets and is simply

$$\langle x(t) \rangle = \int x \, P(x,t) \, dx = \int x \, \Psi^*(x,t) \Psi(x,t) \, dx = \int \Psi^*(x,t) \, x \, \Psi(x,t) \, dx,$$

where the reason for the ordering will become apparent soon.

So for the free electron we have

$$\langle x(t) \rangle = \frac{1}{L} \int_0^L e^{-i(kx-\omega t)} \, x \, e^{i(kx-\omega t)} \, dx = \frac{1}{L} \int_0^L x \, dx = \frac{1}{L} \left[\frac{x^2}{2} \right]_0^L \, dx = \frac{L}{2}$$

and for $\Psi(x, t=0) = (a/\pi)^{1/4} \exp(-ax^2/2)$ we have

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

this is another standard integral we could look up on Wolfram alpha – or by symmetry we see that its zero.

2.3   The Momentum operator

Consider letting the **operator** $\partial/\partial x$ act on the free space wavefunction $\Psi = Ae^{i(kx-\omega t)}$.

$$\frac{\partial \Psi}{\partial x} = \frac{\partial Ae^{i(kx-\omega t)}}{\partial x} = ik\Psi = \frac{ip}{\hbar}\Psi,$$

where we know $p = h/\lambda = \hbar k$. Hence

$$\frac{\hbar}{i}\frac{\partial \Psi}{\partial x} = p\psi \quad \text{or} \quad -i\hbar\frac{\partial \Psi}{\partial x} = p\Psi$$

So there is an association between the dynamical quantity p and the differential operator, $-i\hbar\partial/\partial x$. The effect of multiplying Ψ by p is the same as the effect of operating on it with the differential operator $p\Psi = -i\hbar\partial\Psi/\partial x$. So we can define a momentum operator $\hat{p} = -i\hbar\partial/\partial x$.

This provides us with a recipe for determining the average momentum for any specified wavefunction as

$$\langle p \rangle = \int_{-\infty}^{\infty} \Psi^* \hat{p}\Psi\, dx = \int_{-\infty}^{\infty} \Psi^* \left(-i\hbar\frac{\partial \Psi}{\partial x}\right) dx$$

Note that since $\hat{p}$ is an operator its position in this expression matters, $\int \Psi^*(\partial\Psi/\partial x)dx \neq \int \partial/\partial x(\Psi^*\Psi)dx$.

Example: Consider the Gaussian wavefunction $\Psi(x, t = 0) = (a/\pi)^{1/4}\exp(-ax^2/2)$ we get

$$\begin{aligned}
\langle p \rangle &= \left(\frac{a}{\pi}\right)^{1/2}\int \exp\left(\frac{-ax^2}{2}\right)(-i\hbar)\frac{\partial}{\partial x}\exp\left(\frac{-ax^2}{2}\right)dx \\
&= i\hbar\left(\frac{a}{\pi}\right)^{1/2}\int \exp\left(\frac{-ax^2}{2}\right)(ax)\exp\left(\frac{-ax^2}{2}\right)dx = i\hbar\left(\frac{a^3}{\pi}\right)^{1/2}\int x\exp(-ax^2)dx
\end{aligned}$$

Again, go to Wolfram Alpha or by symmetry – its zero! So in this example the mean momentum is actually zero, and this is telling us physically that it's a standing wave!

2.4   General operators [Gri 3.2]

Any dynamical variable $Q(x, p, t)$ can be written classically as some function of position, momentum and time. We can associate this with an operator $\hat{Q}(x, p, t)$ by substituting $-i\hbar\partial/\partial x$ for p. Then the expectation value of this variable is

$$\boxed{\langle Q(x, t)\rangle = \int \Psi^*(x, t)\hat{Q}(x, -i\hbar\partial/\partial x, t)\Psi(x, t)\, dx}$$

2.5   Equations from operators

The Schrödinger equation is $E = T + V$. In §1.1, where we justified the Schrödinger equation, we saw how to express T in terms of operators and that we could associate energy with the operator $i\hbar\,\partial/\partial t$. Hence we can write this energy equation as

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

In classical mechanics the sum of kinetic and potential energy is called the Hamiltonian. Hence we have

$$\boxed{\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V}$$

and $E\Psi = \hat{H}\Psi$ is the Schrödinger equation.

As another example of turning operators into equations consider the fully relativistic energy equation for a free particle $E^2 = p^2c^2 + m^2c^4$. If we write this in operator form (operating on a wavefunction) we have

$$
\begin{aligned}
i\hbar\frac{\partial}{\partial t}i\hbar\frac{\partial\Psi}{\partial t} &= c^2 \times -i\hbar\frac{\partial}{\partial x} \times -i\hbar\frac{\partial\Psi}{\partial x} + m^2c^4\Psi \\
-\hbar^2\frac{\partial^2\Psi}{\partial t^2} &= -\hbar^2c^2\frac{\partial^2\Psi}{\partial x^2} + m^2c^4\psi
\end{aligned}
$$

this is the **Klein-Gordon equation**, the relativistic form of the Schrödinger equation for a free particle. We can see it makes sense for photons, massless particles, as it reduces to

$$
-\hbar^2\frac{\partial^2\Psi}{\partial t^2} = -\hbar^2c^2\frac{\partial^2\Psi}{\partial x^2}
$$
$$
\frac{\partial^2\Psi}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2\Psi}{\partial t^2}
$$

which is of course the wave equation with velocity c.

3 The origin of uncertainty and other key results

The next few sections are quite formal maths, but prove some worthwhile results.

3.1 Real observables imply Hermitian operators [Gri 3.2]

The expectation value of an observable, Q, can be written as

$$\langle Q \rangle = \int \psi^* \hat{Q} \psi \, dx.$$

Taking the complex conjugate
$$\langle Q \rangle^* = \int \psi \hat{Q}^* \psi^* \, dx = \int \hat{Q}^* \psi^* \psi \, dx.$$

Hence if $\langle Q \rangle$ is real, $\langle Q \rangle = \langle Q \rangle^*$ and $\hat{Q}$ is **Hermitian**, i.e.
$$\boxed{\int \psi^* \hat{Q} \psi \, dx = \int \hat{Q}^* \psi^* \psi \, dx.}$$

The **Hermitian** property is alternatively stated as
$$\boxed{\int f^* \hat{Q} g \, dx = \int \hat{Q}^* f^* g \, dx}$$

for any pair of functions f and g. Proven using $\psi = f + g$ and $\psi = f + ig$ (see Appendix: C).

Any dynamical quantity $Q(x, p, t)$ can be associated with an operator $\hat{Q}(x, -i\hbar \partial/\partial x, t)$. The expectation value $\langle Q \rangle = \int \Psi^* (Q\Psi) dx$ is guaranteed to be real if (and only if) it is Hermitian.

3.2 Non-Hermitian operators and commutators

The wavefunction contains **all** the information there is to know about the associated particle. It encodes the probability density and the expectation values of physical quantities, but must also respect the limitations of the **uncertainty principle**. For example, we cannot know position and momentum to arbitrary accuracy. Hence when we look at operators which constrain *both* position and momentum then they are not necessarily Hermitian.

Consider

$$\langle xp \rangle = \int \Psi^* \hat{x}\hat{p} \, \Psi \, dx = \int \Psi^* x \left(-i\hbar \frac{\partial}{\partial x} \right) \Psi \, dx$$

integrate by parts

$$= -i\hbar \left([\Psi^* x . \Psi]_{-\infty}^{\infty} - \int \Psi \frac{\partial}{\partial x} [x\Psi^*] \, dx \right)$$

the first term is zero as the wavefunction must be zero at $\pm\infty$ as there is no possibility of finding the particle at infinity. Hence

$$= i\hbar \int \Psi . \frac{\partial}{\partial x} [x\Psi^*] \, dx = \int \Psi (\hat{p}\hat{x}\Psi)^* \, dx \neq \langle xp \rangle^* \quad !$$

The reason for the (surprising) final inequality is that $\hat{x}\hat{p} \neq \hat{p}\hat{x}$, i.e. the two operators **do not commute**.

Consider

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

while

$$\hat{p}\hat{x}\Psi = -i\hbar \frac{\partial (x\Psi)}{\partial x} = -i\hbar \left(\Psi + x \frac{\partial \Psi}{\partial x} \right) = -i\hbar \Psi - i\hbar x \frac{\partial \Psi}{\partial x} = -i\hbar \Psi + \hat{x}\hat{p}\psi$$

and so $(\hat{x}\hat{p} - \hat{p}\hat{x})\Psi = i\hbar\Psi$ or equivalently $[\hat{x}, \hat{p}]\Psi = i\hbar\Psi$, where we define the square bracket as the commutator of two operators so $[\hat{a}, \hat{b}] = \hat{a}\hat{b} - \hat{b}\hat{a}$.

When calculating a commutator, then its always *much* safer to explicitly put the wavefunction in the calculation, e.g. $[\hat{x}, \hat{p}]\Psi = \hat{x}\hat{p}\Psi - \hat{p}\hat{x}\Psi = i\hbar\Psi$

As $\hat{x}\hat{p} \neq (\hat{p}\hat{x})^*$ the (composite) operator $\hat{x}\hat{p}$ is **not Hermitian** and so the corresponding quantity is not real and cannot measured! The quantities x and p are examples of incompatible variables. They cannot be measured at the same time.

3.3 Commutators and the Uncertainty principle [Gri 3.5]

The derivation is rather involved but on can show (see [Gri 3.5] or Appendix: A) that for any two operators $\hat{A}$ and $\hat{B}$ with commutator $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$, the uncertainties in A and B satisfy

$$\sigma_A^2 \sigma_B^2 \geq \left(\frac{1}{2i}\left\langle [\hat{A}, \hat{B}] \right\rangle\right)^2 = \left(\frac{1}{2i}(\langle AB \rangle - \langle BA \rangle)\right)^2$$

For example, for $\hat{x}$ and $\hat{p}$ we know $\langle \hat{x}\hat{p} \rangle - \langle \hat{p}\hat{x} \rangle = i\hbar$ and so

$$\sigma_x^2 \sigma_p^2 \geq \left(\frac{i\hbar}{2i}\right)^2$$

or $\boxed{\sigma_x \sigma_p = \Delta x \Delta p \geq \hbar/2}$

3.3.1 Gaussian example:

For $\Psi(x, t = 0) = (a/\pi)^{1/4}\exp(-ax^2/2)$, work out the expectation values $\langle x \rangle$, $\langle p \rangle$, $\langle x^2 \rangle$, etc.

$$
\begin{aligned}
\langle x \rangle &= \sqrt{\frac{a}{\pi}}\int_{-\infty}^{\infty} e^{-ax^2/2}\, x\, e^{-ax^2/2}\, dx = 0 \quad \text{by symmetry}\\
\langle p \rangle &= \sqrt{\frac{a}{\pi}}\int_{-\infty}^{\infty} e^{-ax^2/2}\left(-i\hbar\frac{\partial}{\partial x}\right) e^{-ax^2/2}\, dx = ia\hbar\sqrt{\frac{a}{\pi}}\int_{-\infty}^{\infty} e^{-ax^2/2}\, x\, e^{-ax^2/2}\, dx = 0\\
\langle x^2 \rangle &= \sqrt{\frac{a}{\pi}}\int_{-\infty}^{\infty} e^{-ax^2/2}\, x^2\, e^{-ax^2/2}\, dx = \sqrt{\frac{a}{\pi}}\frac{1}{2}\sqrt{\frac{\pi}{a^3}} = \frac{1}{2a}\\
\langle p^2 \rangle &= \sqrt{\frac{a}{\pi}}\int_{-\infty}^{\infty} e^{-ax^2/2}\left(-\hbar^2\frac{\partial^2}{\partial x^2}\right) e^{-ax^2/2}\, dx\\
&= \hbar^2\sqrt{\frac{a}{\pi}}\int_{-\infty}^{\infty} e^{-ax^2/2}\left(\frac{\partial}{\partial x}\right) ax\, e^{-ax^2/2}\, dx\\
&= \hbar^2\sqrt{\frac{a}{\pi}}\int_{-\infty}^{\infty} e^{-ax^2/2}\,(a - a^2 x^2)\, e^{-ax^2/2}\, dx = \hbar^2\frac{a}{2}
\end{aligned}
$$

As $\langle x \rangle = 0$, $\Delta x = \sigma_x \equiv \langle (x - \langle x \rangle)^2 \rangle^{1/2} = \sqrt{\langle x^2 \rangle}$ and similarly $\Delta p = \sqrt{\langle p^2 \rangle}$ and so

$$\Delta x \Delta p = \sqrt{\frac{1}{2a}}\,\hbar\sqrt{\frac{a}{2}} = \hbar/2,$$

the minimum allowed by Heisenberg!

For completeness let's calculate

$$
\begin{aligned}
\langle xp \rangle &= \sqrt{\frac{a}{\pi}} \int_{-\infty}^{\infty} e^{-ax^2/2}\, x \left(-i\hbar \frac{\partial}{\partial x} \right) e^{-ax^2/2}\, dx \\
&= -i\hbar \sqrt{\frac{a}{\pi}} \int_{-\infty}^{\infty} e^{-ax^2/2}\, x \left(\frac{-a}{2} \right) 2x\, e^{-ax^2/2}\, dx \\
&= ia\hbar \sqrt{\frac{a}{\pi}} \int_{-\infty}^{\infty} x^2 e^{-ax^2}\, dx \\
&= ia\hbar \sqrt{\frac{a}{\pi}} \frac{1}{2} \sqrt{\frac{\pi}{a^3}} \\
&= \frac{i\hbar}{2}
\end{aligned}
$$

Similarly you can show $\langle px \rangle = -i\hbar/2$, and hence show that $\langle xp \rangle - \langle px \rangle = i\hbar$. It is true for this example wavefunction and all wavefunctions.

It is nevertheless odd – we know for this system that $\langle x \rangle = 0$ and $\langle p \rangle = 0$ so the only 'physical' answer to this should be zero! The way to actually form an operator for incompatible variables is to average them i.e. $\langle xp + px \rangle /2 = 0$ as it should be for our example.

3.4   Ehrenfest's theorem [Gri 1.5]

Ehrenfest's theorem states that Newton's laws are exactly satisfied by the expectation values of the corresponding operators in quantum mechanics. I.E

$$
m \frac{d\langle x \rangle}{dt} = \langle p \rangle,
$$

or as $\langle v \rangle = d\langle x \rangle /dt$, $\langle p \rangle = m\langle v \rangle$ and

$$
\frac{d\langle p \rangle}{dt} = -\left\langle \frac{\partial V}{\partial x} \right\rangle = F,
$$

which is Newton's law for a particle moving in a potential $V(x)$.

Proof:

$$
\frac{d}{dt} \langle x \rangle = \frac{d}{dt} \int \psi^* x \psi\, dx = \int \psi^* x \frac{\partial \psi}{\partial t} + \frac{\partial \psi^*}{\partial t} x\psi\, dx
$$

If we use Schrödinger's equation

$$
\frac{\partial \psi}{\partial t} = (i\hbar)^{-1} \left(-\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + V\psi \right)
$$

and its complex conjugate

$$
\frac{\partial \psi^*}{\partial t} = -(i\hbar)^{-1} \left(-\frac{\hbar^2}{2m} \frac{\partial^2 \psi^*}{\partial x^2} + V\psi^* \right)
$$

we can rewrite this as

$$
\frac{d}{dt} \langle x \rangle = (i\hbar)^{-1} \left[\int \psi^* x \left(-\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + V\psi \right) dx - \left(-\frac{\hbar^2}{2m} \frac{\partial^2 \psi^*}{\partial x^2} + V\psi^* \right) x\psi\, dx \right].
$$

The terms involving the potential cancel out and we are left with

$$\frac{d}{dt}\langle x \rangle = \frac{i\hbar}{2m} \int \left[\psi^* x \frac{\partial^2 \psi}{\partial x^2} - \frac{\partial^2 \psi^*}{\partial x^2} x\psi \right] dx. \tag{rate}$$

Now consider the second term in the integral and integrate by parts

$$\int \frac{\partial^2 \psi^*}{\partial x^2} x\psi \, dx = \left[\frac{\partial \psi^*}{\partial x} x\psi \right]_{-\infty}^{\infty} - \int \frac{\partial \psi^*}{\partial x} \frac{\partial(x\psi)}{\partial x} \, dx.$$

The first term vanishes as $\psi = 0$ at $x = \pm\infty$. We can then integrate by parts again to get

$$\int \frac{\partial^2 \psi^*}{\partial x^2} x\psi \, dx = -\left[\psi^* \frac{\partial(x\psi)}{\partial x} \right]_{-\infty}^{\infty} + \int \psi^* \frac{\partial^2(x\psi)}{\partial x^2} \, dx$$

where again the boundary term vanishes. We can now use this to replace the second term in Eq. (rate) which results in

$$
\begin{aligned}
\frac{d}{dt}\langle x \rangle &= \frac{i\hbar}{2m} \int \psi^* \left[x\frac{\partial^2 \psi}{\partial x^2} - \frac{\partial^2(x\psi)}{\partial x^2} \right] dx \\
&= \frac{i\hbar}{2m} \int \psi^* \left[x\frac{\partial^2 \psi}{\partial x^2} - \frac{\partial}{\partial x}\left(\psi + x\frac{\partial \psi}{\partial x} \right) \right] dx \\
&= \frac{i\hbar}{2m} \int \psi^* \left[x\frac{\partial^2 \psi}{\partial x^2} - \frac{\partial \psi}{\partial x} - \frac{\partial \psi}{\partial x} - x\frac{\partial^2 \psi}{\partial x^2} \right] dx \\
&= \frac{-i\hbar}{m} \int \psi^* \frac{\partial \psi}{\partial x} \, dx \\
&= \frac{1}{m} \int \psi^* \hat{p} \psi \, dx = \frac{\langle p \rangle}{m} \qquad \text{QED.}
\end{aligned}
$$

A similar proof starting with

$$\frac{d}{dt}\langle p \rangle = -i\hbar \frac{d}{dt} \int \psi^* \frac{\partial \psi}{\partial x} dx$$

yields [3]

$$\frac{d\langle p \rangle}{dt} = -\left\langle \frac{\partial V}{\partial x} \right\rangle.$$

These are the expected classical results and are examples of the **correspondence principle**. I.E. on macroscopic scales quantum mechanics must give the same results as classical mechanics.

Expectation values obey classical laws, even though the wavefunction itself is probabilistic and thoroughly quantum in nature.

---

[3]One can prove a more general form of Ehrenfest's theorem yielding an expression for the rate of change of the expectation value of any Hermitian operator.

$$\frac{d}{dt}\langle A \rangle = \int \psi^* \frac{\partial \hat{A}}{\partial t} \psi \, dx + \frac{1}{i\hbar} \int \psi^* [\hat{A}, \hat{H}] \psi \, dx$$

For those doing the classical mechanics course in the theory module you will see this is analogous to the classical result $dA/dt = \partial A/\partial t + \{A, \mathcal{H}\}$ where $\{A, \mathcal{H}\}$ is the Poisson bracket and $\mathcal{H}$ the classical Hamiltonian.

4 The Schrödinger equation and eigenfunctions

4.1 Time independent Schrödinger equation [Gri 2.1]

The **time dependent Schrödinger equation**

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

simplifies if the potential does not depend on time, $V(x,t) = V(x)$, as the wave function is separable, $\Psi(x,t) = \psi(x)T(t)$.

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

$$\frac{-\hbar^2}{2m}T(t)\frac{d^2\psi}{dx^2} + V(x)\psi(x)T(t) = i\hbar\psi(x)\frac{dT}{dt}$$

Divide by $\psi(x)T(t)$

$$\frac{-\hbar^2}{2m}\frac{1}{\psi(x)}\frac{d^2\psi(x)}{dx^2} + V(x) = i\hbar\frac{1}{T(t)}\frac{dT}{dt}$$

The LHS is a function only of x, while the RHS is a function only of t. Only possible if both equal a constant, E. For the RHS

$$E = i\hbar\frac{1}{T(t)}\frac{dT(t)}{dt}$$

$$\frac{E}{i\hbar}dt = \frac{dT(t)}{T(t)}$$

which integrates to give $T(t) = e^{-iEt/\hbar}$ on the LHS

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

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

which is the **time independent Schrödinger equation**.

By the separation of variables, We have turned a partial differential equation into two ordinary differential equations. The time equation was easy to solve. The spatial one is **eigenvalue equation**, i.e. of the form

$$\hat{A}(x)f(x) = af(x),$$

where $\hat{A}$ is an operator and a is a number. In other words, the operator acting on a function returns the same function multiplied by a number. In our case $\hat{A}$ is the Hamiltonian operator $\hat{H}$ and we have

$$\boxed{\hat{H}(x)\psi_E(x) = E\psi_E(x)}$$

where the subscript E labels the energy **eigenfunction** as there are typically multiple solutions, each corresponding to a different energy E. The full wavefunction for each eigenfunction, $\Psi_E(x,t)$ is simply given by multiplying ψ_E by $e^{-iEt/\hbar}$.

The general solution to the time dependent Schrödinger equation is the weighted sum of all of these separate wavefunctions,

$$\Psi(x,t) = \sum_E c_E \Psi_E(x,t) = \sum_E c_E \psi_E(x) e^{-iEt/\hbar}.$$

This total wavefunction is *not* a solution of the time independent Schrödinger equation, only each individual $\Psi_E(x,t)$ is.

4.2   Orthogonality of Eigenfunctions

Consider two distinct eigenfunctions ψ_n and ψ_m with energy eigenvalues E_n and E_m respectively. We have

$$\hat{H}\psi_n(x) = E_n\,\psi_n(x) \tag{Eq. N}$$

and

$$\hat{H}\psi_m(x) = E_m\,\psi_m(x). \tag{Eq. M}$$

Multiplying Eq. N by ψ_m^* and integrating we have

$$\int \psi_m^* \hat{H}\psi_n(x)dx = E_n \int \psi_m^* \psi_n(x)dx,$$

but $\hat{H}$ is Hermitian (invariant to taking complex conjugate of the operator and swapping the order), i.e.

$$\int \psi_m^* \hat{H}\psi_n(x)dx = \int \psi_n(x)\hat{H}^* \psi_m^* dx$$

but then using Eq. M

$$\int \psi_n \hat{H}^* \psi_m^* dx = E_m^* \int \psi_n \psi_m^* dx = E_m \int \psi_n \psi_m^* dx,$$

where we have also used that E_m is real.

Hence putting this together we have

$$E_m \int \psi_n \psi_m^* \, dx = E_n \int \psi_n \psi_m^* \, dx$$

or

$$(E_m - E_n) \int \psi_n \psi_m^* \, dx = 0.$$

This requires $\int \psi_n \psi_m^* dx = 0$ for $E_n \neq E_m$ i.e. we have shown eigenstates with non-degenerate eigenvalues are orthogonal. Normalised energy eigenstates are orthonormal i.e. $\int \psi_n \psi_m^* dx = \delta_{nm}$.

Hermitian operators yield real eigenvalues and eigenfunctions which are orthogonal.

They also form a complete set i.e. any function can be written as a linear combination of these eigenfunctions $f(x) = \sum_n c_n \psi_n$.

4.3 Dirac Notation [Gri 3.6]

The maths of operators, eigenfunctions and eigenvalues is completely analogous with the linear algebra of matrices, eigenvectors and eigenvalues.

Consider the eigenvector equation

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}.$$

One can show that the eigenvalues, λ, are real if the matrix $\mathbf{A}$ is Hermitian, i.e. $A_{ij} = A_{ji}^*$, or $\mathbf{A} = \mathbf{A}^{\mathrm{T}*} = \mathbf{A}^\dagger$.

This equivalence can be made clear by using Dirac's bra-ket notation. We write a wavefunction ψ as $|\psi\rangle$ ($|\ \rangle$ is called a ket) and its complex conjugate ψ^* as $\langle\psi|$ ($\langle\ |$ is called a bra). When bra and ket are put together it indicates that one integrates over space. Hence

$$\langle\psi_1|\psi_2\rangle = \int \psi_1^*\psi_2 \, dx$$

and this is analogous to the **dot product** between two vectors. Using this notation we can rewrite some of the equations we had earlier.

The time independent Schrödinger equation becomes

$$\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle.$$

The orthonormality of eigenfunctions becomes

$$\langle\psi_n|\psi_m\rangle = \delta_{nm}.$$

The expansion of an arbitrary wavefunction in terms of stationary states (eigenfunctions of the Hamiltonian) is

$$|\Psi\rangle = \sum_n c_n|\psi_n\rangle$$

This is equivalent to expanding a vector in terms of a set of orthonormal basis vectors, i.e. $\mathbf{v} = \sum_n c_n\mathbf{e}_n$. Here we would find the coefficients c_n by forming the dot product $c_m = \mathbf{e_m}.\mathbf{v}$. Likewise for our wavefunction

$$c_m = \langle\psi_m|\Psi\rangle$$

Expectation value of a quantity given by the operator $\hat{A}$ is

$$\langle A\rangle = \langle\psi|\hat{A}|\psi\rangle \equiv \int \psi^*\hat{A}\psi \, dx.$$

For this set of basis functions an operator is characterised by the matrix elements

$$A_{mn} = \langle\psi_m|A\psi_n\rangle$$

and if $\hat{A}$ is Hermitian then the matrix A_{mn} is Hermitian, i.e.

$$\langle\psi_m|\hat{A}|\psi_n\rangle = \langle\hat{A}\psi_m|\psi_n\rangle \quad \Rightarrow \quad A_{mn} = A_{nm}^*.$$

The proof that non-degenerate eigenfunctions of a Hermitian operator are orthogonal is just

$$\langle\psi_m|\hat{H}|\psi_n\rangle = \langle\hat{H}\psi_m|\psi_n\rangle$$

for any Hermitian operator. Hence

$$E_n\langle\psi_m|\psi_n\rangle = E_m^*\langle\psi_m|\psi_n\rangle = E_m\langle\psi_m|\psi_n\rangle$$

and so

$$(E_n - E_m)\langle\psi_m|\psi_n\rangle = 0$$

requiring $\langle\psi_m|\psi_n\rangle = 0$ if the energies are distinct.

5 Eigenfunctions versus Superpositions

5.1 Quantisation — bound particles

For a free particle then the wavelength and therefore wavenumber k could take any value. But if we trap a particle in some potential V, i.e., it has energy $E = \hbar^2 k^2 / 2m < V$, then effectively multiple reflections set up interference so the only waves which fit are those with some integer wavelength condition and hence some quantisation of energy.

We have multiple eigenstate solutions of the Schrödinger equation ψ_n, each with energy with energy E_n. The observable energy only has a deterministic value if the system is in one of these eigenstates ψ_n. If it is not, and has instead some arbitrarily shaped wavefunction ψ, then this can be expanded out as a sum of all the different eigenstates so that $\psi = \sum_n a_n \psi_n$. The outcome is then *not* deterministic and we will see we can only predict the probability of measuring each of the possible values of E.

5.2 Eigenfunctions of the Infinite Square Well [Gri 2.2]

An example that you know well is the **infinite square well**, $V = 0$ for $0 < x < L$ and ∞ elsewhere. Outside the well ψ must be zero to satisfy Schrödinger's equation. In the well

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

reduces to

$$\frac{d^2 \psi(x)}{dx^2} = -\frac{2mE}{\hbar^2}\psi(x)$$

which is simply the SHM equation $d^2\psi/dx^2 = -k^2\psi$ with $k = \sqrt{2mE}/\hbar$ and solution

$$\psi(x) = A\sin kx + \cancel{B\cos kx}.$$

The 2nd term vanishes due to the **boundary condition** $\psi(0)$, while $\psi(L) = 0 = A\sin kx$ requires $k = n\pi/L$ where n is an integer $n = 1, 2 \ldots$. Thus $\psi(x) = A\sin(n\pi x/L)$ and A is set by the normalisation condition $\int \psi^* \psi \, dx = 1$ giving $A = \sqrt{2/L}$.

The quantum number n specifies the shape of the wavefunction (number of anti-nodes) and the energy through $k = n\pi/L = \sqrt{2mE}/\hbar$ giving

$$\psi_n(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi x}{L}\right) \quad \text{with} \quad E = \frac{n^2\pi^2\hbar^2}{2mL^2}.$$

We can make these into the full energy eigenfunctions by inserting the time dependence

$$\boxed{\Psi_n(x,t) = \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi x}{L}\right)e^{-iE_n t/\hbar}.}$$

Formally, for any eigenfunction, the probability of finding an electron within dx of x is

$$P(x,t)dx = \Psi^*(x,t)\Psi(x,t)dx = \psi^*(x)\,e^{iEt/\hbar}\,\psi(x)\,e^{-iEt/\hbar} = \psi^*(x)\psi(x)dx$$

i.e. is **not time dependent**. We say these are **stationary states** and all their expectation values are time independent

The fully general solution to the time dependent Schrödinger equation is a **superposition** of the eigenstates

$$\Psi(x,t) = \sum_n c_n \Psi_n(x,t) = \sum_n c_n \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right) e^{-iE_n t/\hbar}$$

where the c_n must be chosen so as to normalise the total wavefunction, $\int \Psi^*(x,t)\Psi(x,t)\,dx = 1$ as required.

Explore the infinite square well eigenfunctions on QuVis:InfSqWell[4].

5.2.1 Orthogonality

If we form the overlap integral of two states n and m we have

$$\langle \psi_n | \psi_m \rangle = \int \psi_n^*(x)\psi_m(x)\,dx = \frac{2}{L}\int_0^L \sin(n\pi x/L)\sin(m\pi x/L)\,dx$$

e.g. Wolfram alpha `integrate (2/L) sin(n Pi x/L) sin(m Pi x/L)` gives the indefinite integral

$$= \left[\frac{\sin(\pi(n-m)x/L)}{\pi(n-m)} - \frac{\sin(\pi(n+m)x/L)}{\pi(n+m)}\right]_0^L$$

$$= \frac{\sin(\pi(n-m))}{\pi(n-m)} - \frac{\sin(\pi(n+m))}{\pi(n+m)}$$

when $n \neq m$ then both sin() terms are for an integer times π and so both are zero. When $n = m$ the denominator on the first term goes to zero and we have $\sin(\pi\epsilon)/\pi\epsilon$ which equals 1 in the limit $\epsilon \to 0$. We have explicitly shown that these are orthonormal.

They also span the space (form a complete set) as we know from Fourier analysis that any function can be expressed as a sum of sine waves.

5.3 Wavefunctions in a Superposition of Eigenstates [Gri 2.2]

The general solution is an arbitrary linear superposition of different stationary states given by the eigenfunctions $\psi_n(x)$ associated with energy E_n.

$$\Psi(x) = \sum_n c_n \psi_n(x) \quad \text{or} \quad |\Psi\rangle = \sum_n c_n |\psi_n\rangle$$

and these eigenfunctions are mutually orthogonal. But if we *measure* the energy of the system, it can only be in *one* of these eigenfunctions. The multiple terms are telling us about the probability of finding the electron in a given state.

Consider the superposition of just two states

$$\psi(x) = c_1\psi_1(x) + c_2\psi_2(x).$$

The normalization condition requires

$$\int \psi(x,0)^*\psi(x,0)dx = \int (c_1^*\psi_1^* + c_2^*\psi_2^*)(c_1\psi_1 + c_2\psi_2)dx$$

---

[4] `https://www.st-andrews.ac.uk/physics/quvis/simulations_html5/sims/infwell1d/infwell1d.html`

$$= |c_1|^2 \int \psi_1^* \psi_1 dx + c_1^* c_2 \int \psi_1^* \psi_2 dx + c_1 c_2^* \int \psi_2^* \psi_1 dx + |c_2|^2 \int \psi_2^* \psi_2 dx = |c_1|^2 + |c_2|^2 = 1$$

Now in the general case the wavefunction normalization gives

$$\int \psi^*(x)\psi(x)\,dx = 1 \qquad\qquad \langle \psi(x)|\psi(x)\rangle = 1$$

$$\int \sum_m c_m^* \psi_m^*(x) \sum_n c_n \psi_n(x)\,dx = 1 \qquad\qquad \left\langle \sum_m c_m \psi_m \Big| \sum_n c_n \psi_n \right\rangle = 1$$

$$\sum_m \sum_n \int c_m^* c_n \psi_m^*(x)\psi_n(x)\,dx = 1 \qquad\qquad \sum_m \sum_n c_m^* c_n \langle \psi_m(x)|\psi_n(x)\rangle = 1$$

$$\sum_m \sum_n c_m^* c_n \delta_{mn} = 1 \qquad\qquad \sum_m \sum_n c_m^* c_n \delta_{mn} = 1$$

$$\sum_m |c_m|^2 = 1 \qquad\qquad \sum_m |c_m|^2 = 1$$

Furthermore, the average energy we know is given by

$$
\begin{aligned}
\langle E \rangle &= \int \Psi^* H \Psi\,dx = \int \sum_n c_n^* \psi_n^*(x) H \sum_m c_m \psi_m(x)\,dx \\
&= \sum_n \sum_m c_n^* c_m E_m \int \psi_n^* \psi_m dx = \sum_n \sum_m c_n^* c_m \delta_{mn} E_m = \sum_n |c_n|^2 E_n
\end{aligned}
$$

The only way this can work is if $|c_n|^2$ is the probability of being in state ψ_n and measuring energy E_n

Quantum Mechanics predicts only the probabilities of measuring particular energies. It is inherently non-deterministic, a statistical theory.

5.4 Expectation for Energy of a State in a Superposition of Eigenfunction [Gri 2.2]

The average (expectation) value of the energy is

$$
\begin{aligned}
\langle E(t=0)\rangle = \int \psi^*(x)\hat{H}\psi(x)\,dx &\qquad = \langle \psi(x)|\hat{H}|\psi(x)\rangle \\
= \int \sum_m c_m^* \psi_m^*(x)\hat{H} \sum_n c_n \psi_n(x)\,dx &\qquad = \left\langle \sum_m c_m \psi_m \Big| \hat{H} \sum_n c_n \psi_n \right\rangle \\
= \sum_m \sum_n \int c_m^* c_n \psi_m^*(x)\hat{H}\psi_n(x)\,dx &\qquad = \sum_m \sum_n c_m^* c_n \langle \psi_m(x)|\hat{H}\psi_n(x)\rangle \\
= \sum_m \sum_n c_m^* c_n E_n \delta_{mn} &\qquad = \sum_m \sum_n c_m^* c_n E_n \delta_{mn} \\
\langle E(t=0)\rangle = \sum_m |c_m|^2 E_m &\qquad = \sum_m |c_m|^2 E_m
\end{aligned}
$$

i.e. the expectation value of the energy is the probability weighted sum of the energies associated with each state.

Example: Normalize the superposed state, $\Psi(x,0) = N(2\psi_1 + \psi_2)$, given the individual ψ_1, ψ_2 are already normalized.

We have $\sum_n |c_n|^2 = 1$ where $c_1 = 2N$ and $c_2 = N \Rightarrow 4N^2 + N^2 = 1$ or $N = 1/\sqrt{5}$.

Hence $\Psi(x,0) = \frac{1}{\sqrt{5}}(2\psi_1 + \psi_2)$

Implying the probability of measuring E_1 is $(2/\sqrt{5})^2 = 4/5$ and the probability of measuring E_2 is $1/5$.

5.5 Time dependence for the superposition of eigenstates

For a superposed state the probability

$$\Psi(x,t) = \sum_n c_n \psi_n(x) e^{-iE_n t/\hbar}$$

of being in any particular state oscillates with time.

We can see this explicitly by looking at the probability for a two state system

$$\Psi(x,t) = c_1 \psi_1(x) e^{-iE_1 t/\hbar} + c_2 \psi_2(x) e^{-iE_2 t/\hbar}.$$

$$P(x,t) = \Psi^*\Psi = (c_1^* \psi_1^* e^{iE_1 t/\hbar} + c_2^* \psi_2^* e^{iE_2 t/\hbar})(c_1 \psi_1 e^{-iE_1 t/\hbar} + c_2 \psi_2 e^{-iE_2 t/\hbar})$$
$$= |c_1|^2 |\psi_1|^2 + |c_2|^2 |\psi_2|^2 + c_1^* c_2 \psi_1^* \psi_2 e^{iE_1 t/\hbar} e^{-iE_2 t/\hbar} + c_1 c_2^* \psi_1 \psi_2^* e^{-iE_1 t/\hbar} e^{iE_2 t/\hbar}$$

The last two terms contain $e^{\pm i(E_2 - E_1)t/\hbar}$ which we can write as $e^{\pm i\omega t}$ where $\hbar\omega = E_2 - E_1$. This gives

$$P(x,t) = |c_1|^2 |\psi_1|^2 + |c_2|^2 |\psi_2|^2 + c_1^* c_2 \psi_1^* \psi_2 e^{-i\omega t} + c_1 c_2^* \psi_1 \psi_2^* e^{i\omega t}$$

This is real (as a probability must be) as the last two terms are the complex conjugate of each other (remember $e^{i\omega t} = \cos\omega t + i\sin\omega t$) and so when added together their imaginary parts cancel, but note that form of this probability varies with time. **Superposed states are not stationary states.**

It is still normalised, $\int P(x,t) = 1$, as the orthogonality of eigenfunctions makes each of the last two terms zero when integrated over x.

For superposed states there is a nice interactive illustration and on line challenge at
QuVis:SuperpositionStates[5]

---

[5]http://www.st-andrews.ac.uk/physics/quvis/simulations_html5/sims/SuperpositionStates/SuperpositionStates.html

6 Superposition of eigenstates and transitions

6.1 Eigenstates and transitions

We want to solve the full, time dependent Schrödinger equation to obtain the wavefunction $\Psi(x,t)$. If the potential is constant in time then the wavefunction is separable in space and time. We find the spatial part by solving the time independent Schrödinger equation. In general there are multiple solutions, $H\psi_n = E_n\psi_n$. Then a particular solution is $\Psi_n(x,t) = \psi_n(x)e^{-iE_n t/\hbar}$ and the general solution is $\Psi(x,t) = \sum_n c_n\psi_n(x)e^{-iE_n t/\hbar}$.

A single particular solution (energy eigenfunction) is a stationary state – all expectation values are constant in time. A general solution which is made of multiple particular solutions is *not*. We saw that for a superposed state the probability distribution $|\Psi(x,t)|^2$ was time dependent, which implies all its expectation values are also not constant.

e.g. for 2 terms in the infinite square well the probability density oscillates with a frequency $\omega = (E_2 - E_1)/\hbar$

Consider an electron in the ground state of an atom. This is a single eigenfunction – $\langle x \rangle$ is not dependent on time as its a standing wave. As an electron is charged, then this is also saying that the charge distribution is constant, so there is no radiation. QM provides a way to resolve the paradox you get by thinking of electrons as particles which orbit around atoms. In that picture the electron is accelerating (circular motion) so should radiate and hence the orbit should decay. But in QM the probability density is stationary so the charge is not 'moving' so it doesn't radiate.

But atoms in their excited states do eventually make a transition down to the ground state. So if we have an atom perturbed slightly from the $n = 2$ state then it has some probability of being in the $n = 1$ state as well – of decaying down. If its a superposition of two states then its probability density and hence charge distribution oscillate in time with frequency $\omega = (E_2 - E_1)/\hbar$ which is precisely the frequency of the photon which is emitted to carry away the energy difference. This can't happen for the ground state as there is no lower state for this to mix with.

Wavefunctions which are superpositions of energy eigenstates give expectation values which depend on time, pure energy eigenstates do not.

6.2 How to find the c_m [Gri 3.4]

Going from the particular solution to the general solution we are making the assumption that the energy eigenfunctions span the entire space – that they are a complete set of basis functions. Thus any arbitrary function can be expanded as a sum of these basis functions.

We want the c_n such that $f(x) = \sum_n c_n\psi_n$ so multiply by ψ_m^* and integrate

$$\int \psi_m^* f(x)\,dx = \int \psi_m^* \sum_n c_n\psi_n\,dx = \sum_n c_n \int \psi_m^* \psi_n\,dx = \sum_n c_n\delta_{mn} = c_m$$

or

$$|f\rangle = \sum_n c_n|\psi_n\rangle \quad \text{and so} \quad \langle\psi_m|f\rangle = \sum_n c_n\langle\psi_m|\psi_n\rangle = \sum_n c_n\delta_{mn} = c_m$$

so for any functional form $f(x)$ we can decompose it into a weighted sum of energy eigenfunctions $c_n\psi_n$, by calculating each $c_n = \int \psi_n^* f(x)\,dx$

Example: Consider the wavefunction

$$f(x) = \begin{cases} A & \text{for } 0 < x < L/2 \\ 0 & \text{for } L/2 < x < L \end{cases}$$

in an infinite square well potential spanning $0 < x < L$.

Normalisation requires $A = \sqrt{2/L}$. Then we can decompose it into the sum of the energy eigenfunctions for this potential, $\psi_n(x) = \sqrt{2/L} \sin(n\pi x/L)$

$$c_n = \frac{2}{L} \int_0^{L/2} \sin n\pi x/L \, dx = \frac{2}{L} \left[-\frac{\cos n\pi x/L}{(n\pi/L)} \right]_0^{L/2}$$

$$= -\frac{2L}{Ln\pi} \left[\cos n\pi/2 - \cos 0 \right] = -\frac{2}{n\pi} \left(\cos(n\pi/2) - 1 \right)$$

this is different for even and odd n. For odd n, $\cos(n\pi/2) = 0$ while for even $\cos(n\pi/2) = (-1)^{n/2}$. Hence $c_1 = 2/\pi$, $c_2 = 2/\pi$, $c_3 = 2/3\pi$, $c_4 = 0$, $c_5 = 2/5\pi$ etc.
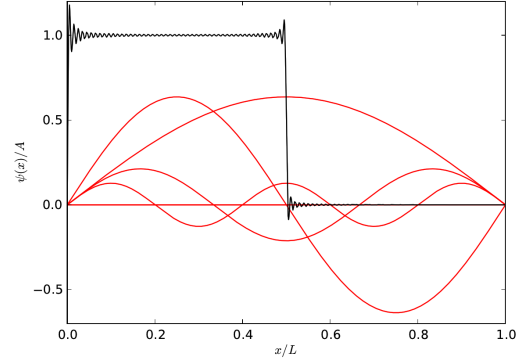
QuVis:EigenstateExpansion[6]

Figure 3: Synthesis of a superposed state, showing the first four non-zero eigenfunctions and the sum up to 200 terms.

7 Eigenfunctions of various potentials

We will look at various potentials leading up to that of the hydrogen. These are typically symmetric about the origin (a symmetry we can exploit), and so let us first revisit the infinite square well but making it symmetric about the origin, i.e. spanning $-L/2 < x < L/2$.

7.1 The infinite square well

We know the solutions (see §5.2) when spanning $0 < x < L$, $\Psi_n(x, t) = \psi_n(x)e^{-iE_n t/\hbar}$ where

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right) \qquad \text{and} \qquad E_n = \frac{n^2\pi^2\hbar^2}{2mL^2}.$$

We can make it symmetric about the origin just by a change of variable $x' = x - L/2$. The energy levels are unchanged but the solutions for odd n will become cosines (most easily seen by drawing a diagram). Formally we have

$$\psi_n(x') = \sqrt{\frac{2}{L}} \sin(n\pi x'/L + n\pi/2) = \sqrt{\frac{2}{L}} \left(\sin(n\pi x'/L) \cos(n\pi/2) + \cos(n\pi x'/L) \sin(n\pi/2) \right)$$

which depends on n. Eg for $n = 1$, $\sin \pi/2 = 1$ and $\cos \pi/2 = 0$

$$\Rightarrow \psi_1(x') = \sqrt{\frac{2}{L}} \cos(\pi x'/L)$$

---

similarly for $n = 2$ then $\sin \pi = 0$ and $\cos(\pi) = -1$ so

$$\psi_2(x') = -\sqrt{\frac{2}{L}} \sin\left(\frac{2\pi x'}{L}\right)$$

$$\psi_3(x') = -\sqrt{\frac{2}{L}} \cos\left(\frac{3\pi x'}{L}\right)$$

$$\psi_4(x') = \sqrt{\frac{2}{L}} \sin\left(\frac{4\pi x'}{L}\right).$$

Hence we see that n odd wavefunctions are even functions, $\psi(x') = \psi(-x')$ and even n wavefunctions are odd functions $\psi(x') = -\psi(-x')$. Note that in either case the probability distribution $P(x) = |\psi(x)|^2$ is symmteric about $x' = 0$ as you might expect from the symmetry.

7.2   The finite square well [Gri 2.6]

Now let the potential outside the well be a finite constant, V_0, rather than infinity.

Physically, we should expect the (bound) wavefunctions to correspond to those of infinite square well, but to be somewhat broader and leak out slightly beyond the edges of the well.

Mathematically, we can solve the Schrödinger equation separately in each of the three regions $x < -L/2,\quad -L/2 < x < L/2$ and $x > L/2$.

For the central region $V = 0$ and we have

$$\frac{-\hbar^2}{2m} \frac{d^2\psi}{dx^2} = E\psi$$

$$\frac{d^2\psi}{dx^2} = \frac{-2mE}{\hbar^2}\psi = -k^2\psi \qquad \text{where} \qquad k^2 = 2mE/\hbar^2$$

$$\Rightarrow \psi = A\cos(kx) + B\sin(kx) \qquad \text{for} \qquad -L/2 < x < L/2$$

For the $x > L/2$ Schrödinger's equation is $\quad -\dfrac{\hbar^2}{2m}\dfrac{d^2\psi}{dx^2} + V_0\psi = E\psi$ or

$$\frac{d^2\psi}{dx^2} = \frac{2m}{\hbar^2}(V_0 - E)\psi = \alpha^2\psi \qquad \text{where} \qquad \alpha^2 = 2m(V_0 - E)/\hbar^2.$$

Here we have used that for bound states $V_0 > E$ and so the constant, $-\alpha^2$, is negative. The general solution is $\psi(x) = Ce^{\alpha x} + De^{-\alpha x}$. However for the solution to be normalizable, $\int |\psi|^2\,dx = 1$, C must be zero, i.e

$$\psi(x) = De^{-\alpha x} \qquad \text{for} \qquad x > L/2.$$

Similarly for $x < -L/2$,
$$\psi(x) = Fe^{\alpha x} \qquad \text{for} \qquad x < -L/2.$$

As there is not an infinite discontinuity in V then $\partial\psi/\partial x$ as well as ψ must be continuous at $x \pm L/2$. In principle this is all we need to find all the solutions, but it is rather laborious. Let us instead use the fact that the solutions will either be even or odd, $\psi(x) = \pm\psi(-x)$.

Even solutions: We only need consider the join at $x = L/2$, as the behaviour at $-L/2$ is fixed by the symmetry. Applying continuity

$$\text{in } \psi \quad : \quad A\cos kL/2 = De^{-\alpha L/2}$$
$$\text{in } \partial\psi/\partial x \quad : \quad -kA\sin kL/2 = -\alpha De^{-\alpha L/2}$$

Dividing the second by the first and multiplying through by $L/2$ we find

$$(kL/2)\tan kL/2 = \alpha L/2$$

Meanwhile for the definitions of k and α in terms of the energy we also have

$$\left(\frac{\alpha L}{2}\right)^2 + \left(\frac{kL}{2}\right)^2 = \frac{2mV_0}{\hbar^2}\left(\frac{L}{2}\right)^2.$$

Solving these two simultaneous equations to find the allowed values of k and hence the energy levels and wavefunctions isn't trivial (could use Wolfram alpha), but we can find an informative graphical solution by setting $X = kL/2$ and $Y = \alpha L/2$. In terms of these variables the two equations we need to solved are

$$Y = X\tan X \quad \text{and} \quad X^2 + Y^2 = \frac{2mV_0}{\hbar^2}\left(\frac{L}{2}\right)^2.$$

We plot Y(X) for both equations and where they intersect are the allowed solutions. These are shown by the red curves of $X\tan X$ and the black circles, $X^2 + Y^2 = \text{const}$. We have plotted three circles corresponding to three values of the well depth V_0
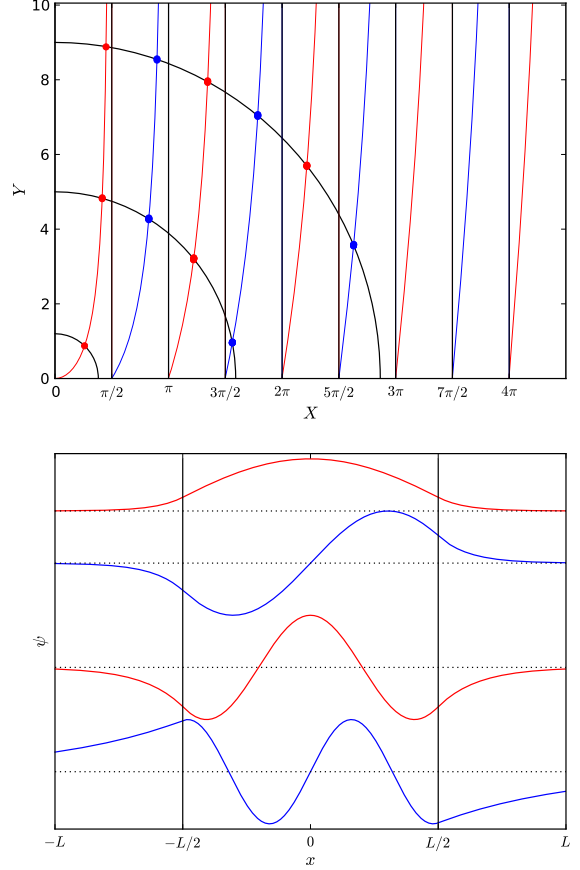

Figure 4: **Top:** Graphical solution for the allowed values of $k[= XL/2]$ for three different values of V_0 (see text). **Bottom:** The ground state and three excited states for the intermediate case.

Odd Solutions: Following the same method for the odd solutions (sin rather cos in the middle) then sin swaps with cos and we find

$$Y = -X\cot X \qquad \text{and} \qquad X^2 + Y^2 = \frac{2mV_0}{\hbar^2}\left(\frac{L}{2}\right)^2.$$

These are the blue curves in the top panel.

Looking at the plot:

1. There is always at least one bound state, but if the circle has radius $R < \pi/2$, i.e. $(V_0 < \hbar^2\pi^2/(2mL^2))$ then this is the only state.

2. For the intermediate circle there are 4 allowed states corresponding to the four points of intersection. Reading of their X and Y values determines their k and α values. Plugging these back into either of the continuity equations gives D (and F by symmetry) in terms of A. To get the value of A we use the normalization condition $\int \Psi^*\Psi\,dx = 1$. We have to do this in piecewise fashion as our wavefunction is piecewise continuous. i.e.

$$\int_{-\infty}^{-L/2} \psi^*\psi\,dx + \int_{-L/2}^{L/2} \psi^*\psi\,dx + \int_{L/2}^{\infty} \psi^*\psi\,dx = 1.$$

These wavefunctions are shown in the lower panel.

3. In the limit as $V_0 \to \infty$ the radius of the circle become very large and the intersections occur at $X(= kL/2) = n\pi/2$ and so the energy values are as in the infinite square well $E_n = n^2\hbar^2k^2/2m = \pi^2\hbar^2/(2mL^2)$.
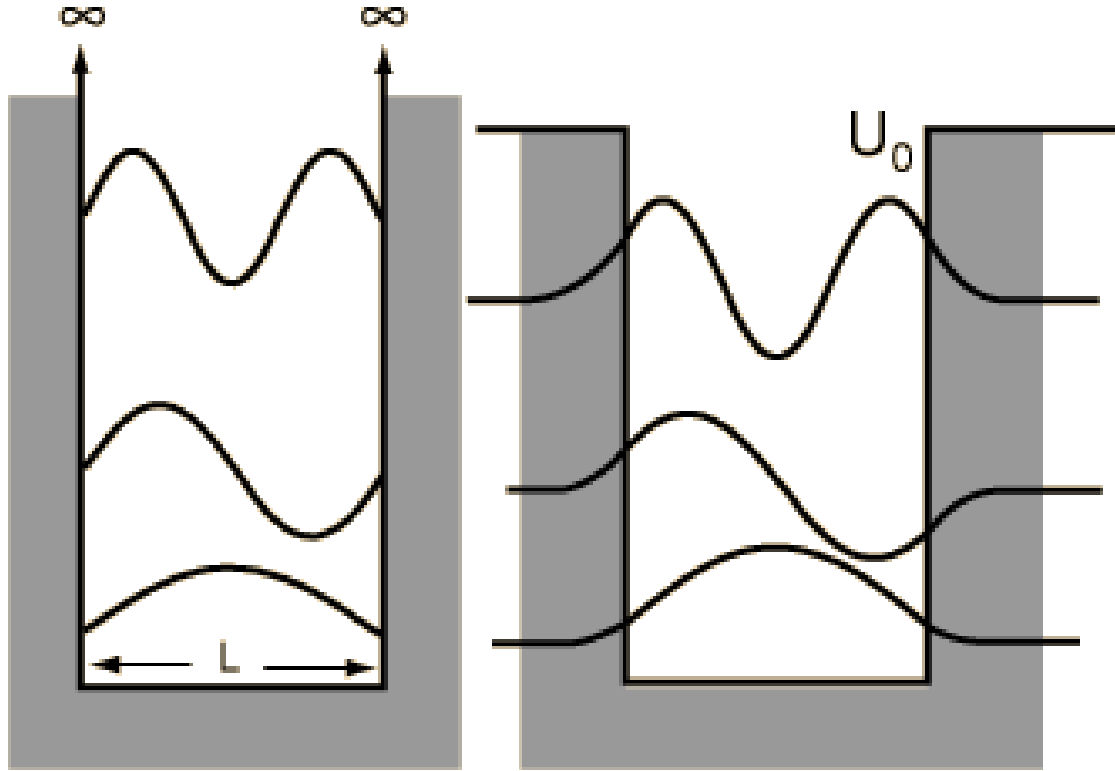

Figure 5: Comparison of infinite and finite square well wavefunctions.

Explore varying the parameters of the finite square well on QuVis:FiniteSqWell[7].

7.3 Commutator algebra

We saw $[A, B] = AB - BA = -[B, A]$ and that this is *not* necessarily equal to zero i.e **operators do not have to commute**. E.g. position and momentum operators do not commute, $[x, p] = i\hbar$.

Some readily provable commutator identities:

$$[A, A] = AA - AA = 0$$

$$[A + B, C] = (A + B)C - C(A + B) = AC + BC - CA - CB = [A, C] + [B, C]$$

$$[AB, C] = ABC - CAB = ABC - ACB + ACB - CAB = A[B, C] + [A, C]B$$

$$[A, BC] = ABC - BCA = ABC - BAC + BAC - BCA = [A, B]C + B[A, C]$$

Example: Consider $[H, x]$ (always good to give this a wavefunction to work on)

$$[H, x]\psi = [p^2/2m + V(x), x]\psi = [p^2/2m, x]\psi + [V(x), x]\psi = 1/2m[p^2, x]\psi$$

$$= 1/2m(p[p, x] + [p, x]p)\psi = 1/2m(p[-i\hbar] + [-i\hbar]p)\psi = -(2i\hbar/2m)p\psi = -(i\hbar/m)\,p\psi$$

Then take away the test wavefunction and we have shown $[H, x] = -(i\hbar/m)p$.

---

[7]https://www.st-andrews.ac.uk/physics/quvis/simulations_html5/sims/finite-infinite-well/finite-infinite-well.html

7.4 Consequences of non-commutation

All operators A, B which commute share a common set of eigenfunctions.

To see this let $Af_n = a_n f_n$, then $[A, B] = 0$ implies $ABf_n - BAf_n = 0$. so $A(Bf_n) = a_n(Bf_n)$ but this means that $Bf_n \propto f_n$ i.e. $Bf_n = b_n f_n$.

This tells us a bit more about the uncertainty principle – if operators commute then they share a common set of eigenfunctions, so measuring one does not disturb the wavefunction for the measurement of the next. If they don't commute then in measuring A we change the wavefunction to be one of the eigenfunctions of A, f_n $(Af_n = a_n f_n)$. If these are not eigenfunctions of B we can still expand f_n in terms of the eigenfunctions of B $(Bg_n = b_n g_n)$ I.E

$$f_n = \sum_m c_m g_m, \quad \text{where} \quad c_m = \int g_m^* f_n \, dx.$$

Now when we measure B we do not get a deterministic value of B. Instead we measure the value b_m with probability $|c_m|^2$.

Non-commutation means the two associated variables cannot simultaneously have deterministic values. $\Rightarrow$ An uncertainty principle applies.

7.5 The linear harmonic oscillator [Gri 2.3]

This has a potential $V(x) = \frac{1}{2}kx^2$, and classically it oscillates with frequency $\omega = \sqrt{k/m}$ so $V(x) = \frac{1}{2}m\omega^2x^2$. It has wide applications because about an equilibrium position x_0 any arbitrary continuous potential $V(x)$ is harmonic to leading order

$$V(x) \approx V(x_0) + \frac{dV}{dx}\Big|_{x_0}(x - x_0) + \frac{1}{2}\frac{d^2V}{dx^2}\Big|_{x_0}(x - x_0)^2 + ...$$

The time independent Schrödinger equation is

$$\left(\frac{\hat{p}^2}{2m} + V\right)\psi = E\psi$$

$$\frac{1}{2m}\left(\hat{p}^2 + (m\omega x)^2\right)\psi = E\psi$$

7.5.1 Ladder operators [Gri 2.3.1]

If this were an algebraic equation we would factorize, i.e. $(a^2 + b^2) = (-ib + a)(ib + a)$, *but these are operators*, and worse, operators that don't commute, $[x, \hat{p}] = i\hbar$

However let us still consider

$$a_\pm = \frac{1}{\sqrt{2\hbar m\omega}}(\mp ip + m\omega x),$$

where the prefactor is chosen to make things neater later on.

$$
\begin{aligned}
a_+a_- &= \frac{1}{2\hbar m\omega}(-ip + m\omega x)(ip + m\omega x) \\
&= \frac{1}{2\hbar m\omega}(p^2 - im\omega px + im\omega xp + m^2\omega^2 x^2) \\
&= \frac{1}{2\hbar m\omega}(p^2 + im\omega(xp - px) + m^2\omega^2 x^2) \\
&= \frac{1}{2\hbar m\omega}(p^2 + im\omega i\hbar + m^2\omega^2 x^2) \\
&= \frac{1}{2\hbar m\omega}(p^2 - m\omega\hbar + m^2\omega^2 x^2) \\
&= \left(\frac{1}{\hbar\omega}H - \frac{1}{2}\right)
\end{aligned}
$$

and so $H = \hbar\omega(a_+a_- + 1/2)$.

Similarly $a_-a_+ = (\frac{1}{\hbar\omega}H + \frac{1}{2})$ and so $H = \hbar\omega(a_-a_+ - 1/2)$

Hence the Hamiltonian does not quite factor perfectly and we have

$$\hbar\omega\left(a_+a_- + \frac{1}{2}\right)\psi = E\psi$$

or

$$\hbar\omega\left(a_-a_+ - \frac{1}{2}\right)\psi = E\psi.$$

Suppose we could find a solution, ψ_n, with associated energy E_n. Then if we operate on it with a_+ we get $a_+\psi_n$, which is also a solution!

$$H(a_+\psi_n) = \hbar\omega\left(a_+a_- + \frac{1}{2}\right)a_+\psi_n$$

$$= \hbar\omega\left(a_+a_-a_+ + \frac{1}{2}a_+\right)\psi_n$$

$$= \hbar\omega a_+\left(a_-a_+ + \frac{1}{2}\right)\psi_n$$

$$= a_+\left(\hbar\omega\left(\frac{1}{\hbar\omega}H + \frac{1}{2}\right) + \frac{1}{2}\hbar\omega\right)\psi_n = a_+(H + \hbar\omega)\psi_n$$

$$= a_+(E_n\psi_n + \hbar\omega\psi_n) = a_+(E_n + \hbar\omega)\psi_n = (E_n + \hbar\omega)a_+\psi_n$$

so if ψ_n satisfies the Schrödinger equation with energy E then $a_+\psi_n$ satisfies it with energy $E_n + \hbar\omega$

Similarly, $H(a_-\psi_n) = (E_n - \hbar\omega)a_-\psi_n$

so $a_\pm$ are *ladder* operators. We have one solution, and we can find all the rest by moving up and down the ladder in energies!

Hence all we need is *one* solution to get started. There must be a bottom rung, with wavefunction ψ_0 at which if we tried to go lower we couldn't. I.E. $a_-\psi_0 = 0$.

$$a_-\psi_0 = \frac{1}{\sqrt{2\hbar m\omega}}\left(\hbar\frac{d}{dx} + m\omega x\right)\psi_0 = 0$$

$$\frac{d\psi_0}{dx} = -\frac{m\omega}{\hbar}x\psi_0$$

$$\ln\psi_0 = -\frac{m\omega}{2\hbar}x^2 + c$$

where c is a constant. Hence $\psi_0 = Ne^{-m\omega x^2/2\hbar}$ normalise this and we get

$$\psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4}e^{-m\omega x^2/2\hbar}$$

which is the Gaussian wavefunction we used in the first couple of lectures with $a = m\omega/\hbar$.

The energy associated with this is also easy to find from $H\psi = \hbar\omega(a_+a_- + 1/2)\psi = E\psi$. We know $a_-\psi_0 = 0$ so $1/2\hbar\omega\psi_0 = E_0\psi_0$ so $E_0 = \hbar\omega/2$. I'm going to call this $n = 0$ state as all the rest are $\psi_n = A_n a_+^n \psi_0$ and have energies $E_n = (n + 1/2)\hbar\omega$.

7.5.2 Brute force solution [Gri 2.3.2]

If we hadn't been smart, we could have solved it by more standard maths techniques. Starting with

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

we can change variables to $\zeta = \sqrt{\frac{m\omega}{\hbar}}x$ and then we have

$$\frac{d^2\psi}{d\zeta^2} = (\zeta^2 - K)\psi$$

where $K = 2E/\hbar\omega$.

This can be solved using the Frobenius/Power Series technique (see Maths Methods lectures). One finds polynomial solutions known as **Hermite polynomials** $H_n(\zeta)$ where n denotes the highest power of ζ present so H_0 is constant, $H_1 = a\zeta + b$ etc. Also n is also the quantization condition with $K = 2n + 1$ so $(2n + 1)\hbar\omega/2 = E_n$ or $E_n = (n + 1/2)\hbar\omega$ (as before). The

$$\psi_n(\zeta) = \left(\frac{m\omega}{\pi\hbar}\right)\frac{1}{\sqrt{2^n n!}} H_n(\zeta) e^{-\zeta^2/2}$$

where

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

so

$$H_0(\zeta) = 1$$

$$H_1(\zeta) = 2\zeta \quad \text{etc}$$

Higher orders become more tedious but one can also use a recurrence relation

$$H_{n+2}(\zeta) - 2\zeta H_{n+1}(\zeta) + 2(n+1)H_n(\zeta) = 0$$

which Hermite polynomials satisfy. This yields

Figure 6: Harmonic Oscillator wavefunctions and probability distributions.

$$H_2(\zeta) = ... = 4\zeta^2 - 2$$

$$H_3(\zeta) = ... = 8\zeta^3 - 12\zeta$$

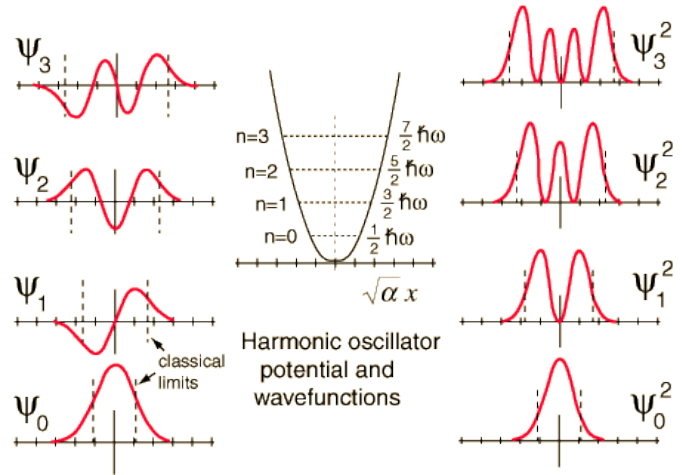$$H_4(\zeta) = ... = 16\zeta^4 - 48\zeta^3 + 12$$

Finally turn it back from ζ to x!

7.6 Properties of harmonic potential

The wavefunctions of the infinite square well, finite square well and harmonic oscillator have similar properties. Quantized energies and alternately wavefunctions with even and odd symmetry.

$\boxed{E_n = (n + 1/2)\hbar\omega}$ and n runs from 0 rather than 1.

The lowest energy state is $E_0 = \hbar\omega/2$.

The system has **zero point energy** which is non-zero due to the Heisenburg uncertainty principle $\Delta x \Delta p \geq \hbar/2$. The system cannot sit motionless at the bottom of its potential well, for then its position and momentum would both be completely determined to arbitrarily high precision. Therefore, the lowest-energy state (the ground state) of the system must have a distribution in position and momentum that satisfies the uncertainty principle, which implies its energy must be greater than the minimum of the potential well.

The harmonic oscillator wavefunctions are quite similar in shape to the finite square well potential wavefunctions, but the nice thing here is that the potential doesn't have unphysical discontinuities. The oscillations come from polynomials rather than sin/cos functions, but the general form is still quite similar. [Gri Fig 2.7]

QuVis:HarmonicOscillator[8]

---

[8]http://www.st-andrews.ac.uk/physics/quvis/simulations_html5/sims/QuantumOscillator/oscillator2.html

8 The 3-D Schrödinger Equation [Gri 4.1]

So far we have only looked at 1D problems, but we will need 3D to treat many physical situations such as electrons in atoms.

8.1 Cartesian coordinates

In 3D, we still have $E = T + V$ but the kinetic energy now has 3 components, $T = p_x^2/2m + p_y^2/2m + p_z^2/2m$

Hence

$$H = \frac{p^2}{2m} + V(x, y, z, t) = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{p_z^2}{2m} + V(x, y, z, t)$$

$$i\hbar \frac{\partial}{\partial t} = \frac{1}{2m}\left(-i\hbar\frac{\partial}{\partial x}\right)^2 + \frac{1}{2m}\left(-i\hbar\frac{\partial}{\partial y}\right)^2 + \frac{1}{2m}\left(-i\hbar\frac{\partial}{\partial z}\right)^2 + V(x, y, z, t)$$

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

so our 3D Schrödinger equation is

$$i\hbar \frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2\Psi}{\partial x^2} + \frac{\partial^2\Psi}{\partial y^2} + \frac{\partial^2\Psi}{\partial z^2}\right) + V(x, y, z, t)\Psi$$

where $\Psi = \Psi(x, y, z, t)$. We can write this more compactly in vector form by recognizing that $\underline{r} = x\mathbf{i} = y\mathbf{j} + z\mathbf{k}$ so $\Psi(x, y, z, t) = \Psi(\underline{r}, t)$ and $V(x, y, z, t) = V(\underline{r}, t)$. Then the 3D **time dependent Schrödinger equation** is

$$\boxed{i\hbar\frac{\partial\Psi(\underline{r}, t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi(\underline{r}, t) + V(\underline{r}, t)\Psi(\underline{r}, t)}$$

and we can use the same separation of variables techniques to show that as long as the potential is *not* dependent on time then the equation is separable in time and space as $\Psi_n(\underline{r}, t) = \psi_n(\underline{r})e^{-iE_n t/\hbar}$ where the **time independent Schrödinger equation** is

$$\boxed{\frac{-\hbar^2}{2m}\nabla^2\psi_n(\underline{r}) + V\psi_n(\underline{r}) = E_n\psi_n(\underline{r})}$$

or going back explicitly to Cartesian coordinates

$$\frac{-\hbar^2}{2m}\left(\frac{\partial^2\psi_n}{\partial x^2} + \frac{\partial^2\psi_n}{\partial y^2} + \frac{\partial^2\psi_n}{\partial z^2}\right) + V(x, y, z)\psi_n = E\psi_n(x, y, z).$$

This is separable into $\psi_n(x, y, z) = X(x)Y(y)Z(z)$ if $V(x, y, z) = V_x(x) + V_y(y) + V_z(z)$.

$$\frac{-\hbar^2}{2m}\left(YZ\frac{\partial^2 X}{\partial x^2} + XZ\frac{\partial^2 Y}{\partial y^2} + XY\frac{\partial^2 Z}{\partial z^2}\right) + (V_x + V_y + V_z)XYZ = E\,XYZ$$

divide by XYZ to get

$$\left(\frac{-\hbar^2}{2mX}\frac{\partial^2 X}{\partial x^2} + V_x\right) + \left(-\frac{\hbar^2}{2mY}\frac{\partial^2 Y}{\partial y^2} + V_y\right) + \left(-\frac{\hbar^2}{2mZ}\frac{\partial^2 Z}{\partial z^2} + V_z\right) = E$$

The three terms are dependent on only x, y and z respectively so *none* of them can depend on x,y or z - they must each be constants which we'll call E_x, E_y and E_z. This gives the three equations

$$\frac{-\hbar^2}{2m}\frac{\partial^2 X(x)}{\partial x^2} + V_x X(x) = E_x\, X(x)$$

$$\frac{-\hbar^2}{2m}\frac{\partial^2 Y(x)}{\partial y^2} + V_y Y(y) = E_y\, Y(y)$$

$$\frac{-\hbar^2}{2m}\frac{\partial^2 Z(z)}{\partial z^2} + V_z Z(z) = E_z\, Z(z)$$

where $E_x + E_y + E_z = E$.

There are not many useful potentials that can be written as $V(x, y, z) = V_x(x) + V_y(y) + V_z(z)$ but one important exception is the Harmonic Oscillator $V \propto r^2 \propto x^2 + y^2 + z^2$.

8.1.1 Infinite square well potential

$V = 0$ for $0 < x < L_x$, $0 < y < L_y$ and $0 < z < L_z$ and ∞ elsewhere. Inside the well we have

$$\frac{-\hbar^2}{2m}\frac{\partial^2 X(x)}{\partial x^2} = E_x\, X(x),$$

$$\frac{-\hbar^2}{2m}\frac{\partial^2 Y(x)}{\partial y^2} = E_y\, Y(y),$$

$$\frac{-\hbar^2}{2m}\frac{\partial^2 Z(z)}{\partial z^2} = E_z\, Z(z).$$

Each one of these is just the same as the 1D case. So we can write down the solution as

$$X(x) = \sqrt{\frac{2}{L_x}}\sin n_x\pi x/L_x \quad E_x = \frac{n_x^2\pi^2\hbar^2}{2mL_x^2}$$

$$Y(y) = \sqrt{\frac{2}{L_y}}\sin n_y\pi y/L_y \quad E_y = \frac{n_y^2\pi^2\hbar^2}{2mL_y^2}$$

$$Z(z) = \sqrt{\frac{2}{L_z}}\sin n_z\pi z/L_z \quad E_z = \frac{n_z^2\pi^2\hbar^2}{2mL_z^2}$$

Hence the full wavefunction is

$$\psi(x, y, z) = X(x)Y(y)Z(z) = \sqrt{\frac{8}{L_xL_yL_z}}\,\sin(n_x\pi x/L_x)\,\sin(n_y\pi y/L_y)\,\sin(n_z\pi z/L_z)$$

where allowed energy levels are

$$E = E_x + E_y + E_z = \frac{\pi^2\hbar^2}{2m}\left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2}\right).$$

In the special case of a cube, $L_x = L_y = L_z = L$, $(V = L^3)$, we have

$$\psi_{\underline{n}}(x, y, z) = \sqrt{\frac{8}{V}}\,\sin(n_x\pi x/L)\,\sin(n_y\pi y/L)\,\sin(n_z\pi z/L)$$

with energy

$$E_{\underline{n}} = \frac{\hbar^2 \pi^2}{2mL^2}(n_x^2 + n_y^2 + n_z^2)$$

The ground state has $n_x = n_y = n_z = 1$ and there is only one wavefunction with energy $E = 3\hbar^2\pi^2/(2mL^2)$ (degeneracy one or non-degenerate).

The next energy level has one of the dimensions in the $n = 2$ state. but this could be either of n_x, n_y or n_z. So there are 3 possible different wavefunctions with this energy, where $(n_x, n_y, n_z) = (2, 1, 1)$, or (1,2,1) or (1,1,2). These all have energy $E = 6\hbar^2\pi^2/(2mL^2)$ so the level is three fold degenerate.

We get degeneracies because of the *symmetry* of the potential. Each dimension has its own quantization condition and quantum number. If the dimensions are the same then rotating the wave around gives the same energy as before.

QuVis:2DSqWell[9]

8.2 The Schrödinger equation in spherical polars [Gri 4.1.1]

We now want to transform the Schrödinger equation into spherical polar coordinates, because we often have spherically symmetric potentials potentials such as $V(r) \propto 1/r$, that cannot be separated into $V(r) = V_x(x) + V_y(y) + V_z(z)$.

Instead the most natural coordinates to use are spherical polar coordinates, where the variables are (r, θ, ϕ). Then a potential such as $V(r) = -Ze^2/4\pi\epsilon_0 r$ is a function only of r and not of θ or ϕ and so is separable in these coordinates.

We need to take our 3D Cartesian Schrödinger equation and transform to spherical polar coordinates where

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

and so $r^2 = x^2 + y^2 + z^2$.

Volume integrals become

$$\int_{x=-\infty}^{+\infty} \int_{y=-\infty}^{+\infty} \int_{z=-\infty}^{+\infty} dxdydz = \int_{r=0}^{+\infty} \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} r^2 \sin\theta \, d\theta d\phi dr.$$

The probability of being in some small volume dV about position $\underline{r}$ goes from

$$
\begin{aligned}
P(\underline{r})dV &= \psi^*(x,y,z)\psi(x,y,z)\,dxdydz \quad \text{to} \\
&= \psi^*(r,\theta,\phi)\psi(r,\theta,\phi)\,r^2 \sin\theta \, d\theta d\phi dr.
\end{aligned}
$$

A probability density in a single coordinate is obtained by integrating over all the other coordinates e.g. the radial probability density is $P(r) = \int_\theta \int_\phi \psi^*(r,\theta,\phi)\psi(r,\theta,\phi)\,r^2 \sin\theta d\theta d\phi$.

We need to express the Laplacian

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

---

[9]`http://www.st-andrews.ac.uk/physics/quvis/simulations_chem/ch11-Comparison-particles.html`

in spherical polar coordinates **(Derived by chain rule. See Appendix: B)**

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

so then we write a 3D Hamiltonian in spherical polar coordinates

$$H = \frac{-\hbar^2}{2m}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right] + V(\underline{r}).$$

Now separating variables let $\psi(r,\theta,\phi) = R(r)\Theta(\theta)\Phi(\phi) = R(r)Y(\theta,\phi)$ and substitute into the time independent Schrödinger equation.

$$\frac{-\hbar^2}{2m}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2}\right] + V(\underline{r})\psi = E\psi$$

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2} - \frac{2m}{\hbar^2}(V(r) - E)\psi = 0$$

$$\frac{Y}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) + \frac{R}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{R}{r^2\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2} - \frac{2m}{\hbar^2}(V(r) - E)RY = 0$$

$$\frac{1}{R}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) - \frac{2mr^2}{\hbar^2}(V(r) - E) = -\left(\frac{1}{Y\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{1}{Y\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2}\right)$$

The RHS is a function only of r, the LHS is a function only of θ and ϕ. The only way these two can be equal to each other is if *neither* has any r, θ or ϕ dependence i.e. they equal a constant.

For *any* spherically symmetric potential, the angular dependencies are the same! Hence the solutions, called **Spherical Harmonics** $Y(\theta,\phi)$ are very general and crop up in much of physics.

To understand what they quantify consider the special case of a rigid rotator. A particle of mass m is attached to a light rod of length a pivoting about the origin. With zero potential, $V = 0$, and no radial motion (zero radial derivatives) all the radial terms vanish so the Schrödinger equation becomes

$$\frac{2ma^2}{\hbar^2}E = -\left(\frac{1}{Y\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{1}{Y\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2}\right).$$

As the radius is fixed the kinetic energy related to these functions is only in angular rotation and so is related to angular momentum.

9 Angular Momentum

9.1 Angular momentum operators (Cartesian coordinates) [Gri 4.3]

In 3D, for central potentials, angular momentum is an important concept (in 1D we can only have linear momentum). We know from classical mechanics that the orbital angular momentum vector is $\underline{L} = \underline{r} \times \underline{p}$ where $\underline{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ is the position vector and $\underline{p} = p_x\mathbf{i} + p_y\mathbf{j} + p_z\mathbf{k}$ is linear momentum.

$$\underline{L} = \underline{r} \times \underline{p} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x & y & z \\ p_x & p_y & p_z \end{vmatrix} = (yp_z - zp_y)\mathbf{i} + (zp_x - xp_z)\mathbf{j} + (xp_y - yp_x)\mathbf{k} = L_x\mathbf{i} + L_y\mathbf{j} + L_z\mathbf{k},$$

where

$$L_x = (yp_z - zp_y), \quad L_y = (zp_x - xp_z) \text{ and } L_z = (xp_y - yp_x).$$

Recall in terms of operators, we replace x with $\hat{x} = x$ and p_x with $\hat{p}_x = -i\hbar\dfrac{\partial}{\partial x}$ so

$$\hat{L}_x = -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)$$

$$\hat{L}_y = -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)$$

$$\hat{L}_z = -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)$$

so $\underline{\hat{L}} = -i\hbar(\underline{r} \times \underline{\nabla}) = \hat{L}_x\mathbf{i} + \hat{L}_y\mathbf{j} + \hat{L}_z\mathbf{k}$.

All the operators $\hat{L}_x$, $\hat{L}_y$ and $\hat{L}_z$ are Hermitian and so correspond to measurable observables.

Measuring one of them does not actually constrain both position and momentum along the same direction – they constrain position along one axis together with the momentum along another axis and we can measure both of these together. Lets prove this by evaluating the commutator of e.g. for x and p_y

$$[x, p_y]\psi = (xp_y - p_yx)\psi = x\left(-i\hbar\frac{\partial\psi}{\partial y}\right) - -i\hbar\frac{\partial}{\partial y}(x\psi) = -i\hbar\left(x\frac{\partial\psi}{\partial y} - x\frac{\partial\psi}{\partial y}\right) = 0$$

The result being zero means we can measure both at the same time without any uncertainty.

The Heisenberg uncertainty principle applies to position and momentum along the *same* direction as their operators do not commute, $[x, \hat{p}_x] = i\hbar$.

As with all Hermitian operators the expectation values of $\hat{L}_x$, $\hat{L}_y$ and $\hat{L}_z$ are real and their normalised eigenfunctions are orthonormal.

9.1.1 Combinations of L_x, L_y and L_z

The operators L_x, L_y and L_z do not commute. Consider

$$[L_x, L_y] = [(yp_z - zp_y), (zp_x - xp_z)]$$

$$= [yp_z, zp_x] - \cancel{[yp_z, xp_z]} - \cancel{[zp_y, zp_x]} + [zp_y, xp_z]$$

only position and momentum along the same coordinate have commutator $\neq 0$ e.g. $[x, p_x] = i\hbar$.

$$= [yp_z, zp_x] + [zp_y, xp_z] = yp_x[p_z, z] + xp_y[z, p_z] = -i\hbar yp_x + i\hbar xp_y = i\hbar L_z$$

likewise $[L_y, L_z] = i\hbar L_x$ and $[L_z, L_x] = i\hbar L_y$.

(The non-commutation of the angular momentum in QM related to the non-commutation of rotations in classical mechanics. For example, rotating 90° about the x-axis then 90° about the y-axis gives a different overall rotation than rotating 90° about the y-axis then 90° about the x-axis.)

Non-commutation means that we cannot simultaneously measure all the components of L. We can't simultaneously know any pair of L_x, L_y, L_z. Consequently there is an uncertainty principle which in this case reads

$$\sigma_{L_x}\sigma_{L_y} \geq \frac{1}{2}|\langle [L_x, L_y]\rangle| = \frac{\hbar}{2}|\langle L_z\rangle|.$$

Is there any more information we can know about angular momentum together with one of its components?

9.1.2   Total angular momentum L^2

It turns out that the thing we can measure alongside one of the components is the magnitude of the total angular momentum $L^2 = L_x^2 + L_y^2 + L_z^2$. This commutes with each of the components so $[L^2, L_x] = [L^2, L_y] = [L^2, L_z] = 0$. E.g.

$$
\begin{aligned}
[L_z, L_x^2 + L_y^2 + L_z^2] &= [L_z, L_x^2] + [L_z, L_y^2] + \cancel{[L_z, L_z^2]} \\
&= [L_z, L_x]L_x + L_x[L_z, L_x] + [L_z, L_y]L_y + L_y[L_z, L_y] \\
&= i\hbar\left(L_yL_x + L_xL_y - L_xL_y - L_yL_x\right) = 0
\end{aligned}
$$

So we can measure any single component of angular momentum simultaneously with the total orbital angular momentum. We will choose L_z as the one to measure along with L^2 because in spherical polars L_z has a simpler form than L_x and L_y. By choosing L_z and L^2 we know these commute, so they have a common set of eigenfunctions – now we need to find them!

9.2   Spatial symmetry and angular momentum

Consider a wavefunction that is spherically symmetric and hence only a function of the radial coordinate $\psi(r)$, where $r^2 = x^2 + y^2 + z^2$.

Note that

$$2r\frac{\partial r}{\partial x} = 2x \quad \Rightarrow \quad \frac{\partial r}{\partial x} = \frac{x}{r}$$

and similarly for $\partial r/\partial y$ and $\partial r/\partial z$.

Hence

$$L_z\psi(r) = -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\psi(r) = -i\hbar\left(x\frac{\partial r}{\partial y} - y\frac{\partial r}{\partial x}\right)\frac{d\psi(r)}{dr} = -i\hbar\left(x\frac{y}{r} - y\frac{x}{r}\right)\frac{d\psi(r)}{dr} = 0.$$

Similarly for $L_x\psi(r)$ and $L_y\psi(r)$. Hence the angular momentum is zero for any spherically symmetric wavefunction.

Consider $\psi(\underline{r}) = x + iy$

$$L_z(x + iy) = -i\hbar \left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} \right)(x + iy) = -i\hbar \left(ix\frac{\partial y}{\partial y} - y\frac{\partial x}{\partial x} \right) = \hbar\,(x + iy)\,.$$

Hence $(x + iy)$ is an eigenfunction of L_z with eigenvalue $\hbar$. So we see that angular momentum about the z-axis is related to angular dependence (oscillations between real and imaginary) in the x-y plane. Note $x + iy = re^{i\theta}$ in polar coordinates.

10 Angular momentum and spherical harmonics

10.1 Angular momentum operators (spherical polars)

For potential with spherical symmetry we want to work in spherical polars. Hence we need the angular momentum operators expressed in spherical polars:

$$L_z = -i\hbar \frac{\partial}{\partial \phi}$$

$$L^2 = -\hbar^2 \left(\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial}{\partial\theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2} \right).$$

These can be obtained by repeated application of chain rule
(e.g. see `http://quantummechanics.ucsd.edu/ph130a/130_notes/node216.html`)

We see that L^2 is essentially the angular part of the Laplacian and so we can rewrite our spherically symmetric Schrödinger equation as

$$\frac{1}{R} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right) - \frac{2mr^2}{\hbar^2} (V(r) - E) = - \left(\frac{1}{Y\sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial Y}{\partial\theta} \right) + \frac{1}{Y\sin^2\theta} \frac{\partial^2 Y}{\partial\phi^2} \right) = \frac{L^2 Y}{\hbar^2 Y}.$$

or as

$$\frac{-\hbar^2}{2m} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \psi}{\partial r} \right) + V(r)\psi + \frac{L^2 \psi}{2mr^2} = E\psi, \quad \text{where } \psi = R(r)Y(\theta, \phi).$$

Hence the angular dependence of the wavefunction is related to angular momentum.

10.2 Eigenfunctions of L_z

We want to solve

$$L_z \Phi_m = m\hbar \Phi_m$$

for the eigenvalues m ($\hbar$ inserted for convenience) and eigenfunctions Φ_m. In spherical polars

$$-i\hbar \frac{\partial \Phi_m}{\partial \phi} = m\hbar \Phi_m$$

$$\frac{\partial \Phi_m}{\partial \phi} = im\Phi_m$$

$\Rightarrow \Phi_m \propto e^{im\phi}$, or when normalised $\Phi_m = (2\pi)^{-1/2} e^{im\phi}$

The equation is satisfied for any value of m, but in order for the solution to be single valued we require $\Phi_m(\phi) = \Phi_m(\phi + 2\pi)$ i.e.

$$(2\pi)^{-1/2} e^{im\phi} = (2\pi)^{-1/2} e^{im2\pi + im\phi}$$

$$1 = e^{im2\pi} = \cos m2\pi + i \sin m2\pi$$

require m be an integer, $m = 0, \pm 1, \pm 2, \ldots$

Hence the eigenvalues of L_z are $0, \pm\hbar, \pm 2\hbar \ldots$. Because the z-axis could be chosen to be along any arbitrary direction then the component of orbital angular momentum about *any* axis is quantised. In atoms, m is called the magnetic quantum number, due to the role it plays in the response to a magnetic field.

These are eigenfunctions are orthonormal (as for all Hermitian operators)

$$\int_0^{2\pi} \Phi_n^* \Phi_m \, d\phi = \delta_{nm}$$

and so we can expand any angular function as $f(\phi) = \sum c_m \Phi_m$.

10.3 Eigenfunctions of L^2

As $[L_z, L^2] = 0$ we know that there has to be a common set of eigenfunctions which are *both* eigenfunctions of L_z **and** of L^2. We will call these $Y_{lm}(\theta, \phi)$, labelling them with the subscripts related to the eigenvalues of

$$\boxed{L_z Y_{lm}(\theta, \phi) = m\hbar Y_{lm}(\theta, \phi)}$$

and

$$\boxed{L^2 Y_{lm}(\theta, \phi) = l(l+1)\hbar^2 Y_{lm}(\theta, \phi).}$$

The choice of setting the eigenvalue to be be $l(l+1)\hbar^2$ will eventually become clear!

We can see that $Y_{lm}(\theta, \phi)$ must be separable, $\Theta_{lm}(\theta)\Phi_m(\phi)$, where Φ_m are as above and Θ is only a function of θ and not ϕ as otherwise it would be changed by $L_z = -i\hbar\, \partial/\partial\phi$ and then this would not be an eigenfunction of both of them.

Expressing L^2 in spherical polars and subbing for Y_{lm} we have

$$-\hbar^2 \left(\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial}{\partial\theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2} \right) \Theta(\theta)\Phi_m(\phi) = l(l+1)\hbar^2 \Theta(\theta)\Phi_m(\phi)$$

$$\frac{\Phi_m}{\sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial\Theta}{\partial\theta} \right) + \frac{\Theta}{\sin^2\theta} \frac{\partial^2\Phi_m}{\partial\phi^2} = -l(l+1)\Theta(\theta)\Phi_m(\phi)$$

This simplifies further as we know for the L_z eigenfunction

$$\frac{\partial\Phi_m}{\partial\phi} = im\Phi_m$$

$$\frac{\partial^2\Phi_m}{\partial\phi^2} = im\frac{\partial\Phi_m}{\partial\phi} = -m^2\Phi_m.$$

Substituting in

$$\frac{\cancel{\Phi_m}}{\sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial\Theta}{\partial\theta} \right) - \frac{\Theta}{\sin^2\theta} m^2 \cancel{\Phi_m} = -l(l+1)\Theta(\theta)\cancel{\Phi_m(\phi)}.$$

After cancelling Φ_m this is just a function of θ and so $\partial/\partial\theta \to d/d\theta$ and rearranging we get

$$\left(\frac{1}{\sin\theta} \frac{d}{d\theta} \left(\sin\theta \frac{d}{d\theta} \right) + l(l+1) - \frac{m^2}{\sin^2\theta} \right) \Theta_{lm} = 0.$$

This ugly equation is well known mathematically (see maths methods lectures). One can find closed polynomial solutions (not an infinite number of terms) only when l is an integer (this is why we choose $l(l+1)$). When $m = 0$ the solutions are called **Legendre polynomials** $P_l(\cos\theta)$, and l is the order of the polynomial, e.g.

$$
\begin{aligned}
P_0(\cos\theta) &= 1 \\
P_1(\cos\theta) &= \cos\theta \\
P_2(\cos\theta) &= (3\cos^2\theta - 1)/2 \\
P_3(\cos\theta) &= (5\cos^3\theta - 3\cos\theta)/2
\end{aligned}
$$

(These are normalized so that $P_l(1) = 1$)

Like the Hermite polynomials, the pattern is that even l only have even powers of $\cos\theta$, odd l only have odd powers of $\cos\theta$.

For $m \neq 0$ the solutions are **associated Legendre polynomials** which are related to the $|m|^{th}$ derivative of the P_l,

$$P_l^m(x) = (-1)^m (1 - x^2)^{m/2} \frac{d^m P_l(x)}{dx^m}$$

and

$$P_l^{-m}(x) = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(x).$$

Since P_l is a polynomial of degree l, then its $l+1$ derivative will vanish. Hence for a fixed value of l we require $|m| \leq l$. The allowed values of m for a given l are

$$m = -l, -l+1, -l+2 \ldots 0, 1, 2 \ldots (l-1), l$$

so there are $2l+1$ values of m for every l.

If we normalise the polynomials to get our solutions $\Theta_{lm}(\theta)$, setting $\int \Theta_{lm}^*(\theta)\Theta_{lm}(\theta)\sin\theta\, d\theta = 1$ we get

$$\Theta_{lm} = (-1)^m \left(\frac{2l+1}{2} \frac{(l-m)!}{(l+m)!} \right)^{1/2} P_l^m(\cos\theta).$$

10.4 Spherical harmonics $Y_{lm}(\theta, \phi)$ [Gri 4.1.2]

The eigenfunctions common to L^2 and L_z are given by

$$Y_{lm}(\theta, \phi) = \Theta_{lm}\Phi_m = (-1)^m \left(\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!} \right)^{1/2} P_l^m(\cos\theta)\, e^{im\phi}$$

with $l \geq 0$ and $-l \leq m \leq l$.

With this convention

$$Y_{lm}^*(\theta, \phi) = (-1)^m Y_{l,-m}(\theta, \phi).$$

The range of values that m can take can be understood more physically. Consider

$$\left\langle L^2 \right\rangle = \left\langle L_x^2 + L_y^2 + L_z^2 \right\rangle = \left\langle L_x^2 \right\rangle + \left\langle L_y^2 \right\rangle + \left\langle L_z^2 \right\rangle$$

since these operators are all Hermitian then these are all real and non-negative, implying

$$
\begin{aligned}
\left\langle L^2 \right\rangle &\geq \left\langle L_z^2 \right\rangle \\
\int\int Y_{lm}^* L^2 Y_{lm} \sin\theta d\theta d\phi &\geq \int\int Y_{lm}^* L_z^2 Y_{lm} \sin\theta d\theta d\phi \\
\int\int Y_{lm}^* l(l+1)\hbar^2 Y_{lm} \sin\theta d\theta d\phi &\geq \int\int Y_{lm}^* L_z (m\hbar Y_{lm}) \sin\theta d\theta d\phi \\
l(l+1)\hbar^2 &\geq \int\int Y_{lm}^* (m^2\hbar^2 Y_{lm}) \sin\theta d\theta d\phi \\
\Rightarrow l(l+1) &\geq m^2
\end{aligned}
$$

This physical limit implies $|m| \leq l$.

The fact that we have specified L_z means that L_x and L_y cannot simultaneously have well defined values - if we measure them we will get a quantised value, $0, \pm\hbar, \pm 2\hbar \ldots \pm l\hbar$, with varying probability. However, we can explicitly evaluate their averages (expectation values), $\langle L_x \rangle$ and $\langle L_y \rangle$. These both turn out to be zero, i.e. the probability of measuring $\hbar$ and $-\hbar$ and so on are equal.

A semi-classical **vector model** helps in the understand this behaviour. Consider a total angular momentum vector of magnitude $\sqrt{l(l+1)}\hbar$ precessing around the z axis, with fixed z component $m\hbar$. The precession means the time averaged $\langle L_x \rangle$ and $\langle L_y \rangle$ vanish. If we assume m is quantised we deduce from the length of the vector that the possible values of m are $0, \pm 1 \ldots \pm l$.

Each choice of l and m is associated with the corresponding **spherical harmonic** and visualised as
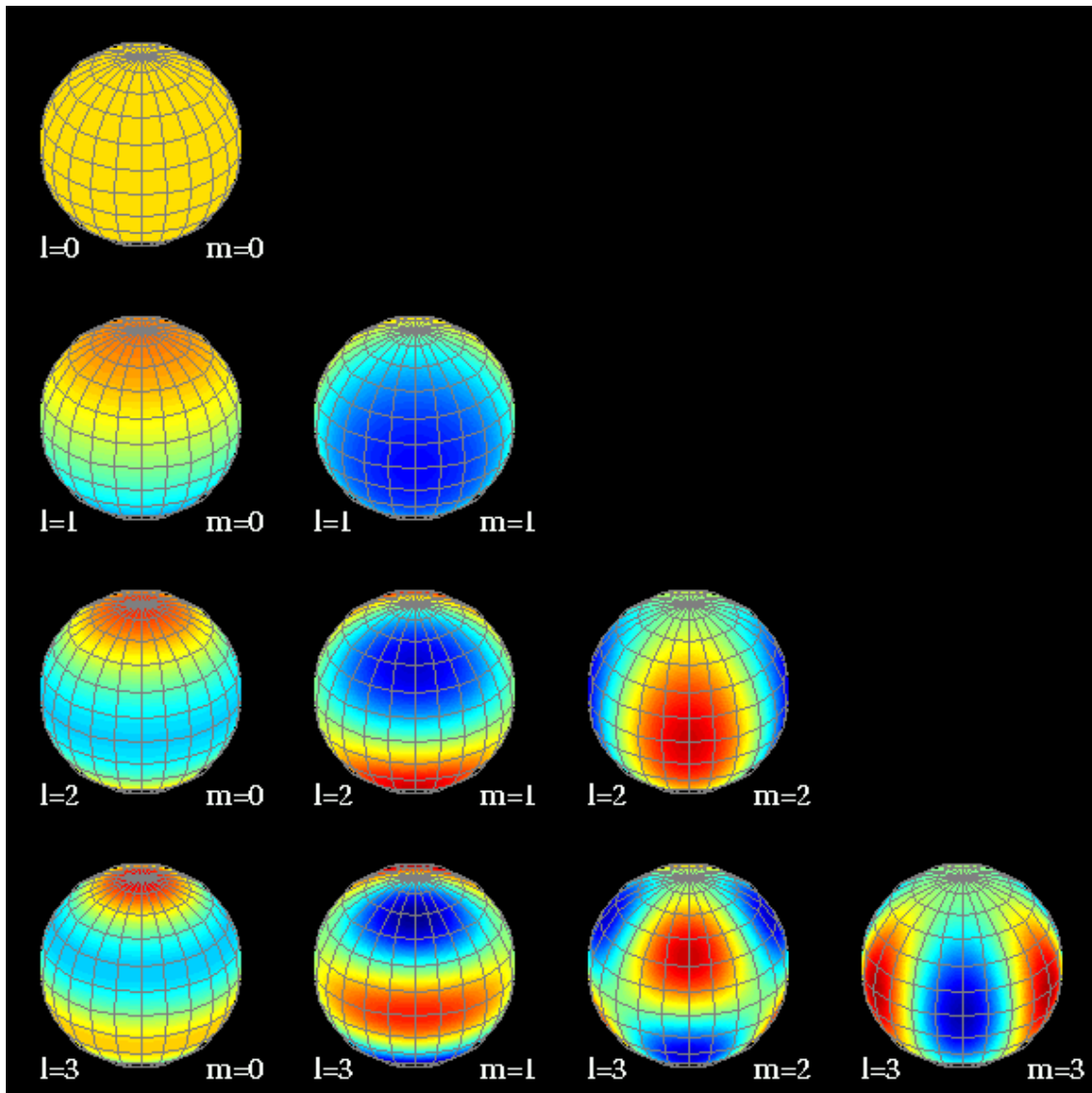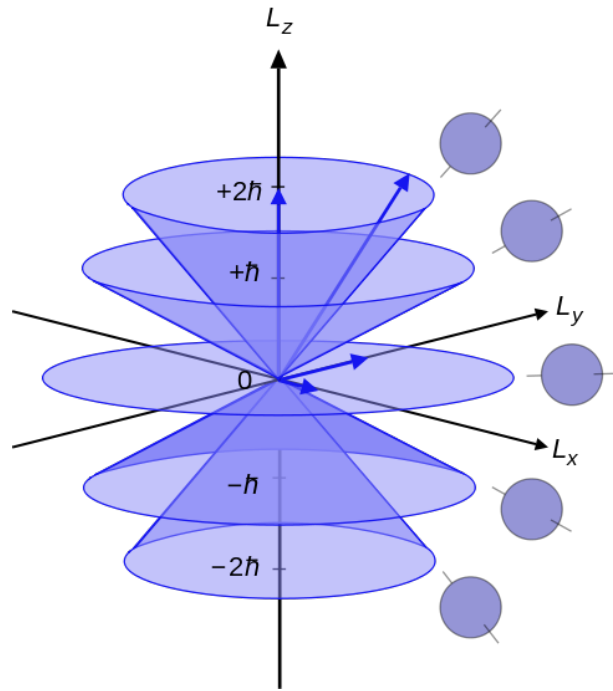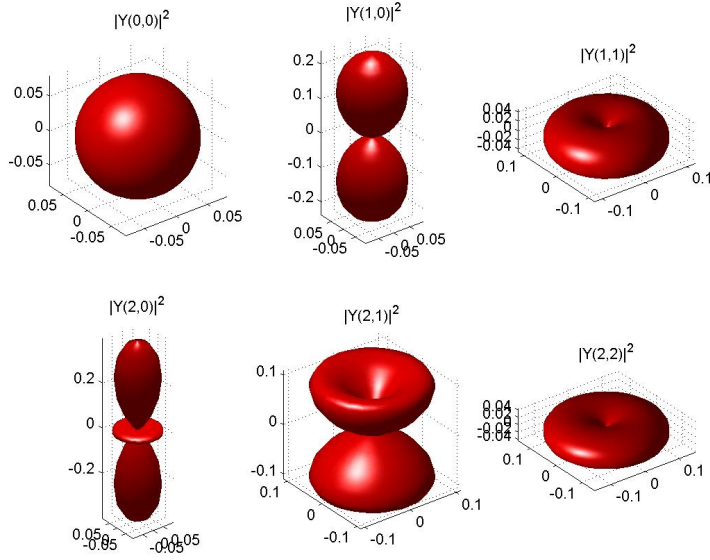
Figure 7: Here the amplitude of $|Y_{lm}(\theta, \phi)|^2$ is shown by the radial distance from the origin.

Perhaps of more interest than the wavefunctions are the associated probability distributions

$$P(\theta, \phi) \, \sin\theta d\theta d\phi =$$

$$Y_{lm}^*(\theta, \phi) Y_{lm}(\theta, \phi) \, \sin\theta d\theta d\phi$$

or

$$P(\theta, \phi) \, d\Omega = Y_{lm}^*(\theta, \phi) Y_{lm}(\theta, \phi) \, d\Omega$$

where $d\Omega$ is the element of solid angle, $\sin\theta d\theta d\phi$. These are easier to visualise as ϕ dependence of the $e^{-im\phi}$ and $e^{im\phi}$ of the Y_{lm}^* and Y_{lm} cancel out.

If we just want the θ dependence of the probability we can integrate over ϕ,

$$P(\theta) \, d\theta = \int_{\phi=0}^{2\pi} |Y_{lm}(\theta, \phi)|^2 d\phi \, \sin\theta \, d\theta$$

11 Finding the Hydrogen Wavefunction

11.1 The radial equation for a spherical potential [Gri 4.1.3]

Our Schrödinger equation now looks like

$$\frac{1}{R}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) - \frac{2mr^2}{\hbar^2}\left(V(r) - E)\right) = \frac{L^2 Y}{\hbar^2 Y},$$

but we know $L^2 Y = l(l+1)\hbar^2 Y$ and so substituting for this we have

$$\frac{1}{R}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) - \frac{2mr^2}{\hbar^2}\left(V(r) - E\right) - l(l+1) = 0.$$

This is an equation for the eigenfunctions $R(r)$. The equation depends on l but not m and for each l there will be a set of eigenfunctions which we can label with an index n. Hence we can denote the eigenfunctions $R_{nl}(r)$ and as this equation involves only r the partial derivatives are standard derivatives

$$\frac{d}{dr}\left(r^2\frac{dR_{nl}}{dr}\right) - \frac{2mr^2}{\hbar^2}\left(V(r) - E\right)R_{nl}(r) = l(l+1)R_{nl}(r).$$

To simplify it is useful to let $U_{nl} = rR_{nl}$

$$\frac{d}{dr}\left(r^2\frac{dR_{nl}}{dr}\right) = \frac{d}{dr}\left(r\frac{dU_{nl}}{dr} - U_{nl}\right) = r\frac{d^2U}{dr^2} + \frac{dU_{nl}}{dr} - \frac{dU_{nl}}{dr} = r\frac{d^2U_{nl}}{dr^2}.$$

Substitute this back in

$$\frac{d^2U_{nl}}{dr^2} - \frac{2m}{\hbar^2}\left(V(r) - E\right)U_{nl} - l(l+1)\frac{U_{nl}}{r^2} = 0.$$

Multiplying through by $-\hbar^2/(2m)$ and rearranging we have the **radial equation**

$$\frac{-\hbar^2}{2m}\frac{d^2U_{nl}}{dr^2} + \left[V(r) + \frac{\hbar^2}{2m}\frac{l(l+1)}{r^2}\right]U_{nl} = EU_{nl}$$

which is identical in form to the 1D Schrödinger equation except that potential $V(r)$ is replaced by the **effective potential** $V_{\text{eff}}(r) = V(r) + \hbar^2 l(l+1)/(2mr^2) = V(r) + L^2/(2mr^2)$.

The reason $L^2/(2mr^2)$ behaves like a contribution to the potential can be appreciated by considering an electron on a circular orbit. The force associated with the centripetal acceleration is $F = mv^2/r$, which because $L = mvr$ can be written as $F = L^2/(mr^3)$. A (conservative) force is related to its potential via $F = -dV/dr$ and so integrating we find $V = L^2/(2mr^2)$. Angular momentum acts like a central potential hill.

11.2 The hydrogen atom [Gri 4.2]

We now have everything we need to solve the Schrödinger equation for a real physical system, the hydrogen atom. Here we have 2 particles (electron and proton) rather than one, but as in classical mechanics we can deal with this by just replacing the electron mass by the **reduced mass** $\mu = M_p m_e/(M_p + m_e)$.

The radial Schrödinger equation for the coulomb potential $V(r) = -Ze^2/(4\pi\epsilon_0 r)$ is

$$-\frac{\hbar^2}{2\mu}\frac{d^2U_{nl}}{dr^2} + \frac{l(l+1)\hbar^2}{2\mu r^2}U_{nl} - \frac{Ze^2}{4\pi\epsilon_0 r}U_{nl} = EU_{nl}$$

where $U_{nl}(r) = rR_{nl}(r)$ and $\psi_{nlm}(r, \theta, \phi) = R_{nl}(r)Y_{lm}(\theta, \phi)$. (For hydrogen $Z = 1$, but we include it so that we also get the result for hydrogenic ions such as He$^+$.)

In Appendix: D we show that this equation can be related to a standard differential equation, the **Associated Laguerre Equation** and its solutions written as

$$R_{nl} \propto \rho^l \, e^{-\rho} \, L_{n-l-1}^{2l+1}(2\rho)$$

where $\rho = kr$ with $k = \mu Ze^2/(4\pi\epsilon_0\hbar^2 n)$ for $n = 1, 2, 3...$ so that the Hydrogen wavefunctions are

$$\Psi_{nlm}(r, \theta, \phi) = R_{nl}(r)Y_{lm}(\theta, \phi)$$

and corresponding quantized energies

$$\Rightarrow E = -\frac{\mu}{2\hbar^2}\left(\frac{Ze^2}{4\pi\epsilon_0}\right)^2\frac{1}{n^2}$$

$$\boxed{E = -13.6\frac{Z^2}{n^2}\frac{\mu}{\mu_H}\text{ eV}}$$

This only depends on n, not on l. The properties of Laguerre polynomials require $n \geq l + 1$ so $l \leq n - 1$. These are only features of our $1/r$ potential – other central potentials do not do this though *all* have energies not dependent on m. (For Hydrogen $Z = 1$ and $\mu = \mu_H$, but left in to keep general.)

Degeneracy: For a given n there are n values of l which have the same energy, so the level is n degenerate, but also for each l there are also $2l + 1$ degenerate values of m, so in fact each level is $n^2 = \sum_{l=0}^{n}(2l + 1)$ degenerate (or $2n^2$ degenerate when, later on, we include spin).

Scale length: As ρ depends on n through $\rho = kr$

$$\rho = kr = \frac{\mu Ze^2}{2\pi\epsilon_0\hbar^2\rho_0}\,r = \frac{\mu Ze^2}{4\pi\epsilon_0\hbar^2 n}\,r.$$

Defining $a = (4\pi\epsilon_0\hbar^2)/(\mu Ze^2)$, which for hydrogen is the Bohr radius of 5.29×10^{-11} m^{-1}, we can write our full wavefunction as

$$\psi_{nlm} \propto \left(\frac{r}{2a}\right)^l e^{-r/(an)} L_{n-l-1}^{2l+1}(2r/(an))\, Y_{lm}(\theta, \phi),$$

where the normalization constant is determined by requiring

$$\int_{r=0}^{\infty}\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi} \psi^*(r, \theta, \phi)\psi(r, \theta, \phi)\, r^2 \sin\theta\, d\theta\, d\phi\, dr = 1.$$

Example Radial functions:

$$\begin{aligned}
R_{10}(r) &= 2\, a^{-3/2}\, e^{-r/a} \\
R_{20}(r) &= \frac{1}{\sqrt{2}}\, a^{-3/2}\left(1 - \frac{r}{2a}\right)e^{-r/2a} \\
R_{21}(r) &= \frac{1}{\sqrt{24}}\, a^{-3/2}\frac{r}{a}\, e^{-r/2a}
\end{aligned}$$

The full wavefunction is $\psi_{nlm} = R_{nl}Y_{lm}$ and the radial probability density is

$$P(r)dr = \int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi} R_{nl}^2 Y_{lm}^* Y_{lm} r^2 \sin\theta\, d\theta d\phi\, dr = R_{nl}^2 r^2\, dr.$$

We can visualise these radial distributions and calculate some representative properties.
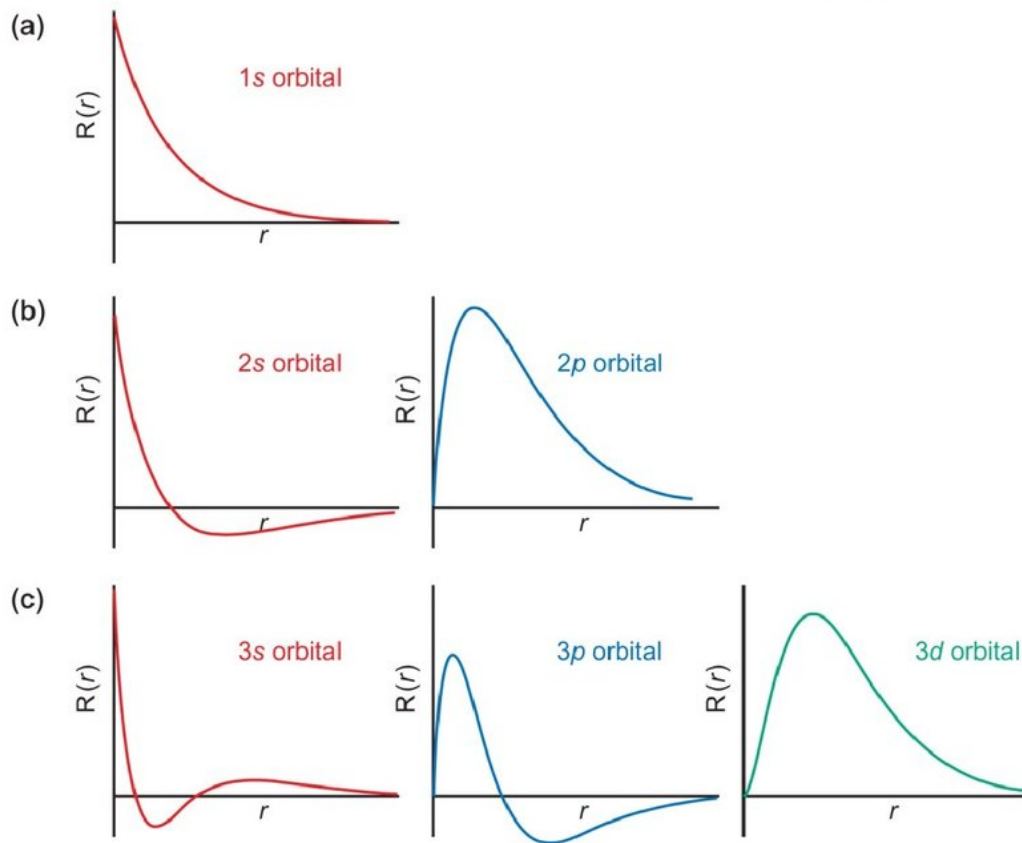
Plot of Radial Wave Function = f(r)

Figure 8: The radial part of the Hydrogen wavefunction. The s,p and d labels are *spectroscopic notation* for the value of l. s$\Rightarrow l = 0$, p$\Rightarrow l = 1$ and d$\Rightarrow l = 2$. (Mnemonic: **S**ober **P**hysicists **D**on't **F**ind **G**iraffes **H**iding **I**n **K**itchens.)

11.3   Ground state of Hydrogen $Z = 1, \mu = \mu_H$

Radial wavefunction R_{10}

```
Plot Laguerre[0,1,2*x] * x**0 * Exp[-x]  x=0 to 10
```

Probability density $|R_{10}|^2 r^2$

```
Plot (Laguerre[0,1,2*x] * x**0 * Exp[-x])**2 * x**2  x=0 to 10
```

Or analytically, we have $R_{10} = \dfrac{2}{a^{3/2}} e^{-r/a}$ and $Y_{00} = \dfrac{1}{\sqrt{4\pi}}$ and so

$$\psi_{100} = R_{10} Y_{00} = \sqrt{\frac{1}{a^3 \pi}} e^{-r/a}$$

From this we can compute expectation values such as

$$\langle r \rangle = \int \int \int \psi_{100}^2 \, r \, r^2 \sin\theta \, dr \, d\theta \, d\phi = 4\pi \frac{1}{\pi a^3} \int r^3 e^{-2r/a} dr$$

which using the standard integral $\int x^p e^{-ax} dx = p!/a^{p+1}$ gives

$$\langle r \rangle = \frac{4}{a^3} \frac{3!}{(2/a)^4} = 3a/2.$$

The classically allowed region is where $E > V$ i.e.

$$E_1 = -\frac{m}{2\hbar^2}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2 > -\frac{e^2}{4\pi\epsilon_0 r}$$

$$r < 2\frac{4\pi\epsilon_0\hbar^2}{me^2} = 2a$$

But being quantum mechanical there is some probability we find the electron at $r \geq 2a$

$$P(r > 2a) = \frac{1}{a^3\pi}\int\int\int e^{-2r/a}r^2\sin\theta dr d\theta d\phi = \frac{1}{a^3\pi}4\pi\int_{2a}^{\infty}r^2 e^{-2r/a}dr = 4\int_2^{\infty}x^2 e^{-2x}dx = 0.24,$$

where the integral is evaluated numerically.

But we could have saved ourselves work on $\langle r\rangle$ by using the fact that the Y_{lm} are already normalised over the θ and ϕ integrals so

$$\langle f(r)\rangle = \int\int\int R_{nl}^* Y_{lm}^* f(r) R_{nl} Y_{lm}\, r^2\sin\theta d\theta d\phi = \int R_{nl}^*(f(r)R_{nl})\, r^2 dr$$

The average potential energy $\langle V\rangle$

$$\begin{aligned}
\langle V\rangle &= \left\langle\frac{e^2}{4\pi\epsilon_0 r}\right\rangle = -\frac{e^2}{4\pi\epsilon_0}\int R_{10}^* r^{-1} R_{10}\, r^2\, dr \\
&= -\frac{e^2}{4\pi\epsilon_0}\frac{4}{a^3}\int e^{-2r/a}\, r\, dr \\
&= -\frac{e^2}{4\pi\epsilon_0}\frac{4}{a^3}\frac{1!}{(2/a)^2} \\
&= -\frac{e^2}{4\pi\epsilon_0 a} = \frac{-\hbar^2}{\mu a^2}
\end{aligned}$$

11.4 Transitions between energy levels [Gri 4.2.2]

If hydrogen in some stationary state ψ_{nlm}, it should be stable, but perturbations (e.g. a collision with another atom/electron/photon) then the electron may undergo a transition to another stationary state – either by absorbing energy or emitting it. Such perturbations are always present, so such transitions – quantum jumps – are constantly occurring. As these transitions occur between levels of fixed energy, then the emitted/absorbed photon has a discrete energy corresponding to energy difference between the two states

$$E_\gamma = E_i - E_f = 13.6\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right)\text{ eV}$$

Transitions to the ground state ($n_f = 1$) give rise to the **Lyman series** of emissions lines with wavelengths given by

$$E = \frac{hc}{\lambda} = 13.6\left(1 - \frac{1}{n^2}\right)\text{ eV with } n \geq 2$$

giving
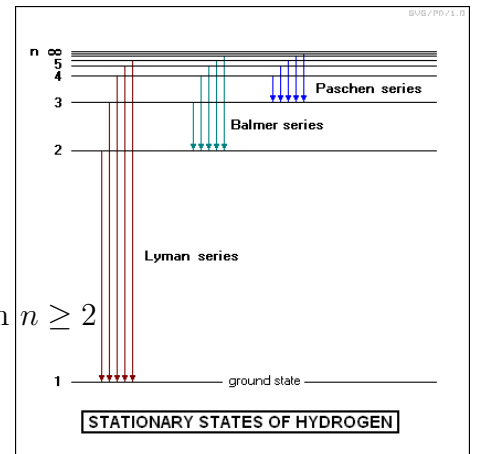
Figure 9: Transitions between Hydrogen energy levels

	Ly-α	Ly-β	Ly-γ	Ly-δ		
$n =$	$2\to 1$	$3\to 1$	$4\to 1$	$5\to 1$	$\ldots$	$\infty\to 1$
λ/nm $=$	121.6	102.6	97.3	95.0	$\ldots$	91.18.

Transitions ending at the first excited state ($n_f = 2$) give rise to the **Balmer series** of emissions lines

Figure 10: Balmer spectral lines.

	H-α	H-β	H-γ	H-δ		
$n =$	$3 \to 2$	$4 \to 2$	$5 \to 2$	$6 \to 2$	$\dots$	$\infty \to 2$
$\lambda/\text{nm} =$	656.3	486.1	434.1	410.2	$\dots$	364.6 .

Any ion with atomic number Z which is ionised so that there is only 1 electron left should be described by the same equations, but with $V(r) = -Ze^2/(4\pi\epsilon_0 r)$.

This affects both atom size and energy as $a = 4\pi\epsilon_0\hbar^2/(mZe^2)$ so this scales as $a = a_H(\mu_H/\mu)(1/Z)$ while $E_n(Z,\mu) = (\mu/\mu_H)Z^2 E_{1H}/n^2$

Helium: Singly ionised Helium, which has more or less the same reduced mass as for hydrogen $\mu_{He} \approx \mu_H \approx m_e$ the ground state energy is $-13.6 \times 4 = -54.4$ eV and length scales are two times smaller than for Hydrogen.

Positronium: Similarly, we could have a different reduced mass μ_Z. E.g. for positronium where the proton is replaced by a positron so the reduced mass $\mu = m_e m_e/(m_e + m_e) = m_e/2$ then $a = a_H/\mu$ so the size scale is twice as large, and as $E \propto \mu$ the energy is smaller by a factor of two.

12 Generalising angular momentum

12.1 Preview of the spin concept

In a classical context, one could consider the angular momentum of a planet as comprising of that from *orbiting* its sun (with a yearly cycle), as well as that due to it *spinning* on its axis (daily cycle). The intrinsic angular momentum of electrons (and other particles) that we call **spin** was named in analogy with this sort of system.

This is potentially misleading, as the electron has *no known internal structure*, and it is therefore not meaningful to think of an electron literally spinning in space (no more than it is to think of changing the orientation in space of a literal, zero-size point).

In quantum mechanics this distinction is fundamental. Orbital angular momentum of an electron about a nucleus we describe in terms of spherical harmonics, but the electron carries another angular momentum unrelated to its position in space, which therefore cannot be described by functions of spatial coordinates such as r, θ, ϕ.

The electron has *intrinsic* angular momentum as well as the *extrinsic* orbital angular momentum.

12.2 Review of angular momentum [Gri 4.3]

Recall (see §9.1.1) that $[L_x, L_y] = i\hbar L_z$, and likewise $[L_y, L_z] = i\hbar L_x$ and $[L_z, L_x] = i\hbar L_y$. As they do not commute, no two components of $\underline{L}$ can be measured simultaneously.

Recall also (see §9.1.2) that $[L^2, L_x] = [L^2, L_y] = [L^2, L_z] = 0$, hence any *one* component of $|\underline{L}|$ can be measured together with L^2 (effectively the total magnitude of $\underline{L}$). We normally choose L_z (this has the simplest mathematical form in spherical polar coordinates).

We found that $L^2 Y_{lm} = l(l+1)\hbar^2 Y_{lm}$ and $L_z Y_{lm} = m\hbar Y_{lm}$ where m takes integer values over $-l, \ldots, l$ — but suppose we didn't know this. Instead, suppose we know only that we have some operator $\underline{L}$ with components L_x, L_y, L_z obeying the commutation relations $[L_x, L_y] = i\hbar L_z$, $[L_y, L_z] = i\hbar L_x$, $[L_y, L_z] = i\hbar L_x$, and $[L^2, L_x] = [L^2, L_y] = [L^2, L_z] = 0$.

This tells us that L^2 and L_z share some common set of eigenstates $f_{\lambda,\mu}$ such that $L^2 f_{\lambda,\mu} = \lambda\hbar^2 f_{\lambda\mu}$, and $L_z f_{\lambda,\mu} = \mu\hbar f_{\lambda,\mu}$.

12.3 Angular momentum ladder operators [Gri 4.3.1]

Similar to the harmonic oscillator ladder operators $a_\pm$ (see §7.5.1), we can construct operators of the form $L_\pm = L_x \pm iL_y$. Considering L_+ specifically, it follows that $[L^2, L_+] = [L^2, L_x + iL_y] = [L^2, L_x] + i[L^2, L_y] = 0$.[10] Hence,

$$0 = [L^2, L_+]f_{\lambda,\mu} = L^2(L_+ f_{\lambda,\mu}) - L_+(L^2 f_{\lambda,\mu}) = L^2(L_+ f_{\lambda,\mu}) - \lambda\hbar^2(L_+ f_{\lambda,\mu})$$
$$\Rightarrow L^2(L_+ f_{\lambda,\mu}) = \lambda\hbar^2(L_+ f_{\lambda,\mu}),$$

and so we have shown the function $L_+ f_{\lambda,\mu}$ is an eigenstate of L^2 with eigenvalue $\lambda\hbar^2$ (but not necessarily proportional to $f_{\lambda,\mu}$ itself, as the μ value can be different)

---

[10] Note that as L_+ is not Hermitian, it does not correspond to an observable, and this result does not imply that L_+ shares common eigenstates with L^2.

Similarly, we determine $[L_z, L_+] = [L_z, L_x] + i[L_z, L_y] = i\hbar L_y + i(-i\hbar L_x) = \hbar L_x + i\hbar L_y = \hbar L_+$, and so, applying this result to the eigenstate $f_{\lambda,\mu}$,

$$\hbar L_+ f_{\lambda,\mu} = [L_z, L_+]f_{\lambda,\mu} = L_z(L_+ f_{\lambda,\mu}) - L_+(L_z f_{\lambda,\mu}) = L_z(L_+ f_{\lambda,\mu}) - \mu\hbar L_+ f_{\lambda,\mu}$$
$$\Rightarrow L_z(L_+ f_{\lambda,\mu}) = (\mu + 1)\hbar(L_+ f_{\lambda,\mu}).$$

Hence, $L_+ f_{\lambda,\mu}$ is also proportional to an eigenstate of L_z with eigenvalue $(\mu+1)\hbar$, i.e., to $f_{\lambda,\mu+1}$. The operator L_+ is a therefore a raising operator — it raises μ by one: $L_+ f_{\lambda,\mu} \propto f_{\lambda,\mu+1}$.

One can equivalently show that L_- is a lowering operator, via $L_z(L_- f_{\lambda,\mu}) = (\mu - 1)\hbar L_- f_{\lambda,\mu}$, and so for a given value of λ there is a ladder of states, with each rung of the ladder separated by one unit of $\hbar$ in L_z. **To ascend the ladder we use L_+, to descend we use L_-.**

12.4 Properties of the ladder operators [Gri 4.3.1]

Experience leads us to conclude there must be a top (and bottom) rung to this ladder — maximum/minimum values of μ which cannot be raised/lowered further by L_+/L_-. Hence, at the top value of $\mu = \mu_{\max}$, $L_+ f_{\lambda,\mu_{\max}} = 0$ (and similarly $L_- f_{\lambda,\mu_{\min}} = 0$).

We now use a useful sleight of hand. We take a product of the raising and lowering operators:

$$L_- L_+ = (L_x - iL_y)(L_x + iL_y) = (L_x^2 + L_y^2 - iL_y L_x + iL_x L_y) = L^2 - L_z^2 + i[L_x, L_y]$$
$$= L^2 - L_z^2 + i^2\hbar L_z = L^2 - L_z^2 - \hbar L_z.$$

Applying this product "on the top rung," we make use of $L_+ f_{\lambda,\mu_{\max}} = 0$. Hence

$$0 = L_-(L_+ f_{\lambda,\mu_{\max}}) = L^2 f_{\lambda,\mu_{\max}} - L_z^2 f_{\lambda,\mu_{\max}} - \hbar L_z f_{\lambda,\mu_{\max}}$$
$$\Rightarrow \lambda\hbar^2 f_{\lambda,\mu_{\max}} = \mu_{\max}^2\hbar^2 f_{\lambda,\mu_{\max}} + \mu_{\max}\hbar^2 f_{\lambda,\mu_{\max}}$$
$$\Rightarrow \lambda = \mu_{\max}(\mu_{\max} + 1),$$

which is consistent with the results of the "Legendre polynomial approach" of Lecture 10.

This procedure can be repeated with $L_+ L_-$ applied to $f_{\lambda,\mu_{\min}}$, yielding $\lambda = \mu_{\min}(\mu_{\min} - 1)$, implying in turn that $\mu_{\max}(\mu_{\max} + 1)$ must be equal to $\mu_{\min}(\mu_{\min} - 1)$. Hence, either $\mu_{\min} = \mu_{\max}+1$ (clearly absurd as this would make the minimum value of μ greater than its maximum), or $\mu_{\min} = -\mu_{\max}$ (also consistent with Lecture 10).

The number of rungs on the ladder must be an integer N, i.e., $-\mu_{\max} + N = \mu_{\max}$; hence $\mu_{\max} = N/2$. This tells us that μ can take integer values — consistent with Lecture 10, which told us that

$$L^2 Y_{lm} = l(l + 1)\hbar^2 Y_{lm} \qquad\qquad l = 0, 1, \ldots;$$
$$L_z Y_{lm} = m\hbar Y_{lm} \qquad\qquad m = -l, -l + 1, \ldots, l - 1, l;$$

or μ can alternatively take *half-integer values*, which is completely new.

Considering only the properties of the commutators as a starting point, the μ associated with a general angular momentum can also take half-integer values.

12.5 Overview of general angular momentum $\underline{J}$

- Any vector $\underline{J}$ is defined to be an angular momentum if its operator components J_x, J_y, J_z satisfy the commutation relations $[J_x, J_y] = i\hbar J_z$, $[J_y, J_z] = i\hbar J_x$, $[J_z, J_x] = i\hbar J_y$, and if its magnitude squared $J^2 = J_x^2 + J_y^2 + J_z^2$ commutes with its components such that $[J^2, J_x] = [J^2, J_y] = [J^2, J_z] = 0$.

- There are then common eigenfunctions of J^2 and J_z called f_{j,m_j}, defined to have eigenvalues $j(j+1)\hbar^2$ and $m_j\hbar$, respectively, i.e., $J^2 f_{j,m_j} = j(j+1)\hbar^2 f_{j,m_j}$ and $J_z f_{j,m_j} = m_j\hbar f_{j,m_j}$.

- There are also ladder operators $J_\pm = J_x \pm iJ_y$, such that J_+ raises m_j by one, and J_- lowers it by one. Hence,

$$J_z J_+ f_{j,m_j} = (m_j + 1)\hbar J_+ f_{j,m_j}, \qquad J_z J_- f_{j,m_j} = (m_j - 1)\hbar J_- f_{j,m_j}.$$

- This cannot go on indefinitely, as we know that $J^2 = J_x^2 + J_y^2 + J_z^2 \geq J_z^2$, and so $j(j+1) \geq m_j^2$. Hence, there is a top value of $m_j = m_{\max}$, and there is also a bottom value $m_j = m_{\min}$.

- Since we are going up and down in integer steps, $m_{\max} - m_{\min} = N$ where $N = 0, 1, 2, \ldots$.

- $J_-(J_+ f_{j,m_{\max}}) = 0$ implies $j(j + 1) = m_{\max}(m_{\max} + 1)$, and $J_+(J_- f_{j,m_{\min}}) = 0$ implies $j(j + 1) = m_{\min}(m_{\min} - 1)$, which shows that $m_{\min} = -m_{\max}$. Hence, $m_{\max} = j$ and $-j + N = j$, where N is an integer, implying that $2j = N$ and therefore that $j = 0, 1/2, 1, 3/2, \ldots$.

- Formally:

$$J^2 f_{j,m_j} = j(j + 1)\hbar^2 f_{j,m_j} \qquad j = 0, 1/2, 1, 3/2, \ldots;$$
$$J_z f_{j,m_j} = j\hbar f_{j,m_j} \qquad m_j = -j, -j + 1, -j + 2, \ldots, 0, \ldots, j - 1, j.$$

A *general* angular momentum can have integer *or half-integer* values of j, with m_j running up to $\pm j$.

13 Spin [Gri 4.4]

13.1 The spin operator

We define $\underline{S}$ as an angular momentum spin operator, with S^2 eigenvalues $s(s+1)\hbar^2$ and S_z eigenvalues $m_s\hbar$. They *do not have spherical harmonics as eigenfunctions*, and their eigenstates are not functions of spatial coordinates (such as θ, ϕ).

Every elementary particle has a specific, immutable value of s which is its intrinsic **spin**; Fermions (such as electrons) have half-integer spin. This is in sharp contrast to orbital angular momentum l, which can take any allowed value $l = 0, \ldots, n-1$ and is *not* fixed — it can change as the system is perturbed.

Electrons are elementary fermions with $s = 1/2$, and so exist in only one of two spin eigenstates: spin up, $m_s = +1/2$, eigenstate χ_+; or spin down, $m_s = -1/2$, eigenstate χ_-.

There is good experimental evidence for this — the Stern–Gerlach experiment. In 1922 they took silver atoms ($Z = 47$), which have a single outer electron in the $n = 5$, $l = 0$, $m = 0$ level. With $l = 0$ the electron has zero orbital angular momentum and therefore produces no current loop; it should therefore not interact with an external magnetic field.

Stern and Gerlach directed their beam of silver atoms through an inhomogeneous magnetic field, which separated the beam into two distinct parts. If the electron had an intrinsic magnetic dipole, it would experience a force proportional to the field gradient ($U = -\boldsymbol{\mu}.\mathbf{B} = -\mu_z B_z$ implying force $F_z = -\partial U/\partial z = \mu_z \partial B_z/\partial z$). This would give a continuous smear at the detectors if the dipole could be oriented in any direction (in 1927 this effect was reproduced using hydrogen by Phipps and Taylor).
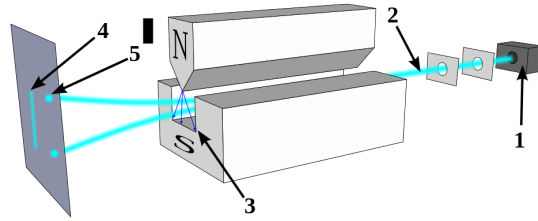

Figure 11: Stern–Gerlach experiment. (By Tatoute - Own work, CC BY-SA 4.0,https://commons.wikimedia.org)

To split into two requires that there are only two allowed directions.

13.2 Electron orbiting in a magnetic field [Gri 4.4.2]

If we consider an electron moving with speed v in a circular orbit of radius r, it sweeps out an area $A = \pi r^2$, and produces a current $I = -ev/2\pi r$. This implies a magnetic moment

$$\mu_l = IA = -\frac{evr}{2} = -\frac{e}{2m_e}m_e vr = -\frac{e}{2m_e}L,$$

where L is the magnitude of the electron orbital angular momentum. In vector form

$$\underline{\mu}_l = -\frac{e}{2m_e}\underline{L} = -\frac{\mu_B}{\hbar}\underline{L},$$

where $\mu_B = e\hbar/2m_e$ is called the **Bohr magneton**, a natural unit of microscopic magnetic moment (value 9.27×10^{-24} J T^{-1} or 5.79×10^{-5} eV T^{-1}).

It is therefore reasonable to identify $-(\mu_B/\hbar)\underline{L}$ as the quantum mechanical magnetic moment operator associated with orbital angular momentum. It follows that the operator for the z-component of the magnetic moment is

$$(\mu_l)_z = -\frac{\mu_B}{\hbar}L_z. \qquad (3)$$

An ideal measurement of the quantity $(\mu_l)_z$ must yield one of the eigenvalues of the corresponding operator.

Hence, for a hydrogen atom with orbital angular momentum quantum number l, the possible values of the quantity $(\mu_l)_z$ are $-m_l\mu_B$, where m_l is an integer value in the range from $-l$ to $+l$.

13.3 Hydrogen atom in a magnetic field

Now consider a hydrogen atom in a z-directed magnetic field $\underline{B} = (0, 0, B_z)$. A classical model would give the associated magnetic potential energy as

$$V = -(\mu_l)_z B_z = m\mu_b B_z.$$

In classical magnetism, a uniform magnetic field creates a torque, but no translational force, on an object with a magnetic moment. A translational force can be produced by applying a spatially varying field — as force is the derivative of potential $F_z = -\partial V/\partial z$ we might expect the force on a hydrogen atom to be

$$F_z = (\mu_l)_z \frac{\partial B_z}{\partial z},$$

where $(\mu_l)_z$ is equal to one of the discrete values $-m_l\mu_B$.

If we have the outer electron in the $l = 0$ state, then $m = 0$, and so there should be no effect — but the beam splits into two in the Stern–Gerlach experiment.

Hence, there is a magnetic moment which _is not associated with orbital angular momentum_, and can take 2 possible values rather than the continuum of values you might expect with a randomly aligned spin dipole.

13.4 Spin magnetic moment $\underline{\mu}_s$

We are thus motivated to associate a magnetic moment $\underline{\mu}_s$ with the spin angular momentum $\underline{S}$. We assume

$$\underline{\mu}_s = -g_s\frac{e}{2m_e}\underline{S} = -g_s\frac{\mu_B}{\hbar}\underline{S} \qquad \text{(moment)}$$

where g_s is called the spin g-factor. The z-directed force is then described by

$$F_z = (\mu_s)_z\frac{\partial B_z}{\partial z} = -g_s m_s \mu_B \frac{\partial B_z}{\partial z}. \qquad \text{(force)}$$

Two lines are seen in the experiment, implying that F_z has two possible values for each hydrogen atom — in turn suggesting that m_s can have two values. If we assume the allowed values of m_s must range from $-s$ to $+s$ in unit steps, in analogy with the relation between m_l and l in the orbital case, then we must take $s = 1/2$ and $m_s = \pm 1/2$.

Since the spin is assumed to be an intrinsic property of the electron, we take the picture to be valid for all electrons, and not just those in hydrogen atoms. Because $s = 1/2$ we refer to the electron as a spin-1/2 particle, which can exist in the states $m_s = 1/2$ or $S_z = +\hbar/2$ (called spin up) and $m_s = -1/2$ or $S_z = -\hbar/2$ (spin down).

Using $m_s = \pm 1/2$, and the experimentally determined force, we find $g_s \approx 2$.[11] Hence, Eq. (moment) yields

$$(\mu_s)_z = -g_s m_s \mu_B \approx \mp \mu_B.$$

13.5 Formalism: Dirac notation and linear algebra [Gri 4.4.1]

In Dirac notation we denote the spin up eigenstate $|\chi_+\rangle$, and the spin down eigenstate $|\chi_-\rangle$. These are not describable as functions of spatial coordinates; there *are no* relevant spatial coordinates and spin is an *intrinsic* property of the electron.

Given that there are only two eigenstates, however, one can describe them as two-dimensional unit vectors. Hence,

$$|\chi_+\rangle \quad \text{corresponds to} \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |\chi_-\rangle \quad \text{corresponds to} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Given also that $S_z|\chi_+\rangle = (+\hbar/2)|\chi_+\rangle$ and $S_z|\chi_-\rangle = (-\hbar/2)|\chi_+\rangle$, one can similarly say that[12]

$$S_z \quad \text{corresponds to} \quad \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

i.e., the Hermitian **operator** S_z is described by a Hermitian **matrix**. Finally, as can be seen from the identity $\langle\chi_+|\chi_+\rangle = \langle\chi_-|\chi_-\rangle = 1$,

$$\langle\chi_+| \quad \text{corresponds to} \quad \begin{pmatrix} 1 & 0 \end{pmatrix}, \qquad \langle\chi_-| \quad \text{corresponds to} \quad \begin{pmatrix} 0 & 1 \end{pmatrix}.$$

Hence, the kets are equivalent to column vectors, bras to row vectors, and operators to square matrices. These statements are very general, and mean that products such as $\langle\psi_n|A|\psi_m\rangle$ can be seen just as two successive matrix multiplications.

Also, taking the complex conjugate of such an inner product can be seen as taking the Hermitian conjugate of three matrices: $(ABC)^\dagger = C^\dagger B^\dagger A^\dagger$, where A is a row vector/bra (Hermitian conjugate is a column vector/ket) and C is a column vector/ket (Hermitian conjugate is a row vector/bra). This is equivalent to $\langle\psi_n|A|\psi_m\rangle^* = \langle\psi_m|A^\dagger|\psi_n\rangle$, which simplifies to $\langle\psi_n|A|\psi_m\rangle^* = \langle\psi_m|A|\psi_n\rangle$ if A is Hermitian.

Bra–ket inner products were introduced in Lecture 4 effectively as a shorthand for integrals [e.g. $\langle\psi_1|\psi_2\rangle = \int \psi_1^*(x)\psi_2(x)\,dx$], however are more general than this as there are times when there is no spatial coordinate to integrate over (e.g. spin).

---

[11]The electron spin g-factor has been measured to extraordinary precision: $g_s/2 = 1.001\,159\,652\,180\,85 \pm 0.000\,000\,000\,000\,076$ [B. Odom, D. Hanneke, B. D'Urso, G. Gabrielse, Phys. Rev. Lett. **97**, 030801 (2006)] and agrees precisely with the prediction of QED.

[12]One can also show that S_x corresponds to $\frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and S_y to $\frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$.

14 Time-independent perturbation theory [Gri 6.1]

14.1 Small modification to a known Hamiltonian

Suppose we have solved the time independent Schrödinger equation for some Hamiltonian H^0. We have the eigenfunctions ψ_n^0 corresponding to each energy level n, so that $H^0 \psi_n^0 = E_n^0 \psi_n^0$. These eigenfunctions are orthonormal, hence $\int \psi_n^{0*} \psi_m^0 \, dx = \delta_{nm}$.

We now **perturb** the system slightly (e.g. putting a small bump within a square well potential) and want to find the eigenfunctions and eigenvalues associated with a new Hamiltonian H, i.e., $H \psi_n = E_n \psi_n$. In general we cannot solve this exactly, but we can use **perturbation theory** to obtain an approximate solution to the perturbed problem by building on known solutions of the unperturbed system.

We write the new Hamiltonian as the sum of two terms: $H = H^0 + \lambda H'$, where H' is the perturbation and the superscript 0 denotes the unperturbed case. We use a dimensionless "small parameter" λ to keep track of first order and second order terms, etc..

14.2 Power expansion

We therefore write E_n and ψ_n as power series in λ

$$E_n = E_n^0 + \lambda E_n^1 + \lambda^2 E_n^2 + \cdots , \qquad\qquad \psi_n = \psi_n^0 + \lambda \psi_n^1 + \lambda^2 \psi_n^2 + \cdots ,$$

where E_n^1 is the first order correction to the nth eigenenergy and ψ_n^1 the first order correction to the nth eigenfunction; superscript 2 denotes the second order corrections, etc..

Expanding $H\psi_n = E_n \psi_n$ then results in

$$(H^0 + \lambda H')(\psi_0 + \lambda \psi_n^1 + \lambda^2 \psi_n^2 + \dots) = (E_n^0 + \lambda E_n^1 + \lambda^2 E_n^2 + \dots)(\psi_0 + \lambda \psi_n^1 + \lambda^2 \psi_n^2 + \dots),$$

which can be restructured to collect together powers of λ

$$H^0 \psi_0 + \lambda(H^0 \psi_n^1 + H' \psi_n^0) + \lambda^2(H^0 \psi_n^2 + H' \psi_n^1) + \cdots = \\ E_n^0 \psi_n^0 + \lambda(E_n^0 \psi_n^1 + E_n^1 \psi_n^0) + \lambda^2(E_n^0 \psi_n^2 + E_n^1 \psi_n^1 + E_n^2 \psi_n^0) + \cdots . \quad \text{(expansion)}$$

14.3 First order correction to the energy

Essentially by definition, Eq. (expansion) to zeroth order is just $H^0 \psi_n^0 = E_n^0 \psi_n^0$ (this is equivalent to having $\lambda = 0$, i.e., no perturbation at all!). However, to first order in λ,

$$H^0 \psi_n^1 + H' \psi_n^0 = E_n^0 \psi_n^1 + E_n^1 \psi_n^0. \qquad\qquad \text{(1st-order)}$$

With this we then take an inner product with ψ_n^0; to maximise generality we use Dirac notation, which means we do not have to specify e.g. how many spatial dimensions we are integrating over, or what length/area/volume element we are using:

$$\langle \psi_n^0 | H^0 | \psi_n^1 \rangle + \langle \psi_n^0 | H' | \psi_n^0 \rangle = \langle \psi_n^0 | E_n^0 | \psi_n^1 \rangle + \langle \psi_n^0 | E_n^1 | \psi_n^0 \rangle = E_n^0 \langle \psi_n^0 | \psi_n^1 \rangle + E_n^1.$$

H^0 is Hermitian, and so $\langle \psi_n^0 | H^0 | \psi_n^1 \rangle \equiv \langle \psi_n^0 | H^0 \psi_n^1 \rangle \equiv \langle H^0 \psi_n^0 | \psi_n^1 \rangle = E_n^0 \langle \psi_n^0 | \psi_n^1 \rangle$, leading to a cancellation and leaving us with

$$\boxed{E_n^1 = \langle \psi_n^0 | H' | \psi_n^0 \rangle.}$$

As a comparison, let us re-express these calculations in terms of integrals, considering specifically a particle moving in one spatial dimension. This means taking Eq. (1st-order), multiplying by $(\psi_n^0)^*$ and then integrating over x:

$$\int (\psi_n^0)^* H^0 \psi_n^1 \, dx + \int (\psi_n^0)^* H' \psi_n^0 \, dx = \int (\psi_n^0)^* E_n^0 \psi_n^1 \, dx + \int (\psi_n^0)^* E_n^1 \psi_n^0 \, dx$$

$$= E_n^0 \int (\psi_n^0)^* \psi_n^1 dx + E_n^1.$$

H^0 is Hermitian, and so

$$\int (\psi_n^0)^* H^0 \psi_n^1 \, dx = \int [H^0 \psi_n^0]^* \psi_n^1 \, dx = E_n^0 \int (\psi_n^0)^* \psi_n^1 \, dx,$$

leading to a cancellation and leaving us with

$$E_n^1 = \int (\psi_n^0)^* H' \psi_n^0 \, dx \,.$$

Either way, the first order correction to an energy eigenvalue is the expectation value of the perturbation using the unperturbed eigenfunctions.

14.4   Example: 1D square well with delta function

With the Kronecker delta, $\delta_{mn} = 1$ if $n = m$ or 0 otherwise. The **Dirac delta** is the continuum version. It is *defined* through

$$\int f(x)\delta(x - x_0) \, dx = f(x_0),$$

i.e., it collapses any integral to the value of the function at the point x_0 marked by the delta function. The delta function itself integrates to unity — $\int \delta(x - x_0) \, dx = 1$.

Hence, if we have $H' = \lambda\delta(x - a/2)$ (perturbing a square well with walls at $x = 0$ and $x = a$), then

$$E_n^1 = \int \psi_n^{0*} H' \psi_n^0 \, dx = \frac{2}{a} \int_0^a \sin(n\pi x/a)\, \lambda\delta(x - a/2)\, \sin(n\pi x/a) \, dx = \frac{2\lambda}{a} \sin^2(n\pi/2).$$

This is $= 0$ if n is even (no correction to even E_n^0), but $E_n^1 = 2\lambda/a$ for odd n. For even n the wavefunction is zero at $x = a/2$ and the perturbation has no effect, while for odd n it peaks and the energies are shifted.

14.5   First order correction to the wavefunction

This can be calculated by first writing $\psi_n^1 = \sum_{n \neq l} c_{nl}\psi_l^0$, substituting this into Eq. (1st-order) and taking the inner product with ψ_l^0. This gives

$$c_{nl} = -\frac{\langle \psi_l^0 | H' | \psi_n^0 \rangle}{(E_n^0 - E_l^0)}$$

and hence ψ_n^1. Then to first order the perturbed wavefunction is $\psi_n \approx \psi_n^0 + \lambda\psi_n^1$.

These are not generally as useful as first-order corrections to the energies, but can be used to determine 2nd order corrections to the energy...

Explore perturbations of the infinite square well QuVis:PerturbSqWell[13]

---

[13]https://www.st-andrews.ac.uk/physics/quvis/simulations_html5/sims/perturbationGame/perturbationGame.html

15 Degenerate perturbation theory [Gri 6.2]

15.1 Degeneracy is different

So far we have assumed that $(E_n - E_m) \neq 0$, however it is not unusual for there to be energy degeneracy, for example in the energy levels of hydrogen! We therefore need to know how to deal with degenerate energy levels.

15.2 Twofold degeneracy [Gri 6.2.1]

Suppose we have an energy level where there are exactly 2 states ψ_a^0 and ψ_b^0 which give the same energy E^0. Hence, any linear combination $\psi^0 = \alpha\psi_a^0 + \beta\psi_b^0$ also gives the same energy E^0; e.g., in hydrogen, for $l = 0, m = 0, m_s = \pm 1/2$.

Typically the perturbation $\lambda H'$ breaks the degeneracy, so that E^0 splits into two values, with the difference in energy increasing as λ increases. If we ramp down the pertubation, the upper state tends back to a unique α, β while the lower state tends back to another unique α, β. We want to find these 'good' unperturbed states!

We have the same expression (1st-order) for the first order correction

$$H^0\psi^1 + \lambda H'\psi^0 = E^0\psi^1 + \lambda E^1\psi^0$$

but now we multiply by the conjugate of *one* of our states ψ_a^{0*} (expressed in Dirac notation as $\langle\psi_a^0|$) and integrate:

$$\langle\psi_a^0|H^0|\psi^1\rangle + \langle\psi_a^0|\lambda H'|\psi^0\rangle = \langle\psi_a^0|E^0|\psi^1\rangle + \langle\psi_a^0|\lambda E^1|\psi^0\rangle$$
$$\Rightarrow \langle H^0\psi_a^0|\psi^1\rangle + \lambda\langle\psi_a^0|H'|\psi^0\rangle = E^0\langle\psi_a^0|\psi^1\rangle + \lambda E^1\langle\psi_a^0|\psi^0\rangle$$
$$\Rightarrow \langle\psi_a^0|H'|\psi^0\rangle = E^1\langle\psi_a^0|\psi^0\rangle$$

Replacing ψ^0 using $\psi^0 = \alpha\psi_a^0 + \beta\psi_b^0$

$$\Rightarrow \langle\psi_a^0|H'(\alpha|\psi_a^0\rangle + \beta|\psi_b^0\rangle) = E^1\langle\psi_a^0|(\alpha|\psi_a^0\rangle + \beta|\psi_b^0\rangle)$$
$$\Rightarrow \alpha\langle\psi_a^0|H'|\psi_a^0\rangle + \beta\langle\psi_a^0|H'|\psi_b^0\rangle = E^1\alpha\langle\psi_a^0|\psi_a^0\rangle + E^1\beta\langle\psi_a^0|\psi_b^0\rangle$$
$$= E^1\alpha.$$

We can write the final result more compactly as

$$\alpha W_{aa} + \beta W_{ab} = \alpha E^1, \text{ where } W_{ij} = \langle\psi_i^0|H'|\psi_j^0\rangle = \int \psi_i^{0*}H'\psi_j^0 \, dx,$$

and i, j is one of a, b. We could have multiplied by ψ_b^{0*} instead, in which case we would get

$$\alpha W_{ba} + \beta W_{bb} = \beta E^1.$$

These can be assembled into a matrix eigenvalue equation, such that

$$\begin{pmatrix} W_{aa} & W_{ab} \\ W_{ba} & W_{bb} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E^1 \begin{pmatrix} \alpha \\ \beta \end{pmatrix},$$

which can be solved in the standard way, by subtracting

$$\begin{pmatrix} W_{aa} - E^1 & W_{ab} \\ W_{ba} & W_{bb} - E^1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and noting that the only non-trivial solutions are where the determinant of the 2x2 matrix is zero (so it can't be inverted). Hence,

$$(W_{aa} - E^1)(W_{bb} - E^1) - W_{ab}W_{ba} = 0.$$

However, as $W_{ab} = W_{ba}^*$, this becomes

$$(E^1)^2 - (W_{aa} + W_{bb})E^1 + (W_{aa}W_{bb} - |W_{ab}|^2) = 0,$$

which is just a quadratic, and so has 2 solutions:

$$\begin{aligned}
E_\pm^1 &= \frac{1}{2}\left[W_{aa} + W_{bb} \pm \sqrt{(W_{aa} + W_{bb})^2 - 4(W_{aa}W_{bb} - |W_{ab}|^2)}\right] \\
&= \frac{1}{2}\left[W_{aa} + W_{bb} \pm \sqrt{(W_{aa} - W_{bb})^2 + 4|W_{ab}|^2}\right].
\end{aligned}$$

If two states, ψ_a and ψ_b are degenerate, both having energy E^0, any linear combination $\psi = \alpha\psi_a + \beta\psi_b$ also has energy E^0. A small perturbation, H', causes a small change in energy, and the first order approximation for this, E^1, is given by the solution of a matrix equation.

15.3 Twofold degeneracy: Example

Suppose the H' is such that $\langle \psi_a^0|H'|\psi_b^0\rangle = \langle \psi_b^0|H'|\psi_a^0\rangle = \kappa$ (i.e., κ is real) while $\langle \psi_a^0|H'|\psi_a^0\rangle = \langle \psi_b^0|H'|\psi_b^0\rangle = 1$. The matrix equation for E^1 is then

$$\begin{pmatrix} 1 & \kappa \\ \kappa & 1 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E^1\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \Rightarrow \begin{pmatrix} 1 - E^1 & \kappa \\ \kappa & 1 - E^1 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0,$$

and the only non-trivial solution (when the determinant is zero) is given by $(1 - E^1)^2 - \kappa^2 = 0$, hence $1 - E_\pm^1 = \mp\kappa$ or $E_\pm^1 = 1 \pm \kappa$.

We can therefore plot the energies splitting as we turn up the perturbation $\lambda H'$ by increasing λ from 0. The energies go from both being E^0 to splitting into $E^0 + \lambda E_+^1$ and $E^0 + \lambda E_-^1$.

To find the corresponding wavefunctions we can solve the eigenvector equation separately for $E_\pm^1$ by substituting into our matrix. Hence, for $E_+^1 = 1 + \kappa$ we have

$$\begin{pmatrix} 1 - (1 + \kappa) & \kappa \\ \kappa & 1 - (1 + \kappa) \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0 \Rightarrow \begin{pmatrix} -\kappa & \kappa \\ \kappa & -\kappa \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0,$$

which implies $-\kappa\alpha + \kappa\beta = 0$ and thus $\alpha = \beta$. We need the standard normalisation conditions of $\alpha^2 + \beta^2 = 1$ and so $\psi_+ = (\psi_a^0 + \psi_b^0)/\sqrt{2}$.

Similarly for $E_-^1 = 1 - \kappa$ we get

$$\begin{pmatrix} \kappa & \kappa \\ \kappa & \kappa \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0,$$

i.e., $\kappa\alpha + \kappa\beta = 0$ or $\alpha = -\beta$ so that $\psi_- = (\psi_a^0 - \psi_b^0)/\sqrt{2}$.

These wavefunctions are special combinations of our original wavefunctions in that they are wavefunctions that "follow" the perturbation.' If we had chosen these to start with them,

rather than ψ_a^0 and ψ_b^0, the matrix elements would have been:

$$
\begin{aligned}
W_{++} &= \langle\psi_+^0|H'|\psi_+^0\rangle = (\langle\psi_a^0| + \langle\psi_b^0|)H'(|\psi_a^0\rangle + |\psi_b^0\rangle)/2 \\
&= (\langle\psi_a^0|H'|\psi_a^0\rangle + \langle\psi_a^0|H'|\psi_b^0\rangle + \langle\psi_b^0|H'|\psi_a^0\rangle + \langle\psi_b^0|H'|\psi_b^0\rangle)/2 \\
&= (1 + \kappa + \kappa + 1)/2 = 1 + \kappa, \\
W_{--} &= \langle\psi_-^0|H'|\psi_-^0\rangle = (\langle\psi_a^0| - \langle\psi_b^0|)H'(|\psi_a^0\rangle - |\psi_b^0\rangle)/2 \\
&= (\langle\psi_a^0|H'|\psi_a^0\rangle - \langle\psi_a^0|H'|\psi_b^0\rangle - \langle\psi_b^0|H'|\psi_a^0\rangle + \langle\psi_b^0|H'|\psi_b^0\rangle)/2 \\
&= (1 - \kappa - \kappa + 1)/2 = 1 - \kappa. \\
W_{+-} = W_{-+}^* &= \langle\psi_+^0|H'|\psi_-^0\rangle = (\langle\psi_a^0| + \langle\psi_b^0|)H'(|\psi_a^0\rangle - |\psi_b^0\rangle)/2 \\
&= (\langle\psi_a^0|H'|\psi_a^0\rangle - \langle\psi_a^0|H'|\psi_b^0\rangle + \langle\psi_b^0|H'|\psi_a^0\rangle - \langle\psi_b^0|H'|\psi_b^0\rangle)/2 \\
&= (1 - \kappa + \kappa - 1)/2 = 0,
\end{aligned}
$$

15.4 Twofold degeneracy: Easier solutions?

If we had only chosen ψ_+ and ψ_- as our original wavefunctions then we would have had a much simpler equation to solve — it being diagonal, we could just read off the solutions.

$$
\begin{pmatrix} W_{++} & 0 \\ 0 & W_{--} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E^1 \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \Rightarrow \begin{pmatrix} W_{++} - E^1 & 0 \\ 0 & W_{--} - E^1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0,
$$

which has solutions $(W_{++} - E^1)\alpha = 0$, i.e., $E_+^1 = W_{++} = \langle\psi_+|H'\psi_+\rangle$, and $(W_{--} - E^1)\alpha = 0$, i.e., $E_-^1 = W_{--} = \langle\psi_-|H'\psi_-\rangle$.

If we feed each of these in turn into the matrix, then we get for $E^1 = W_{++}$

$$
\begin{pmatrix} 0 & 0 \\ 0 & W_{--} - W_{++} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0
$$

so the first line is zero by construction, and the second line is $(W_{--} - W_{++})\beta = 0$, hence $\beta = 0$, and the wavefunction is ψ_+. Simlarly, for $E^1 = W_{--}$ we get $(W_{++} - W_{--})\alpha = 0$, hence $\alpha = 0$, and the wavefunction is ψ_-.

These are the wavefunctions which diagonalise the perturbation, and our solutions are simply the equations we would have picked if we had not bothered thinking about the levels being degenerate.

16 Degenerate perturbation theory II [Gri 6.2]

16.1 Link to non-degenerate perturbation theory

When considering twofold degeneracy, could we choose the 'right' pair of degenerate states in advance? We know that we want the off-diagonal terms in the '$W - E$' matrix to be zero, so that we just solve

$$\begin{pmatrix} W_{aa} - E^1 & 0 \\ 0 & W_{bb} - E^1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

i.e., we want $\langle \psi_a^0 | H' | \psi_b^0 \rangle = 0$. The obvious way this will be true is if ψ_b^0 is also an eigenfunction of H', as then $H' \psi_b^0 = b \psi_b^0$, meaning that $\langle \psi_a^0 | H' | \psi_b^0 \rangle = b \langle \psi_a^0 | \psi_b^0 \rangle = 0$.

What we can do is find some Hermitian operator A which shares common eigenfunctions with H^0, fulfilled when $[A, H^0] = 0$ — thus ψ_a^0 and ψ_b^0 are also eigenfunctions of A — but where the eigenvalues are distinct (unlike for H^0 where they are degenerate): $A\psi_a = \mu\psi_a$ and $A\psi_b = \nu\psi_b$ for $\mu \neq \nu$.

If this operator also commutes with the perturbation, i.e., $[A, H'] = 0$, then ψ_a and ψ_b are the states where the non-diagonal terms in the matrix $W_{ab} = \langle \psi_a^0 | H' | \psi_b^0 \rangle = 0$ and so we can use non-degenerate perturbation theory to work it out. To show this explicitly, $\langle \psi_a | [A, H'] | \psi_b \rangle = 0$ implies that

$$\langle \psi_a | AH' | \psi_b \rangle - \langle \psi_a | H'A | \psi_b \rangle = \langle A\psi_a | H' | \psi_b \rangle - \langle \psi_a | H'\nu | \psi_b \rangle = \mu \langle \psi_a | H' | \psi_b \rangle - \nu \langle \psi_a | H' | \psi_b \rangle = 0.$$

Hence, $(\mu - \nu)W_{ab} = 0$, which, given that $\mu - \nu \neq 0$, means $W_{ab} = 0$.

16.2 Higher-order degeneracy [Gri 6.2.2]

This generalises easily to n-fold degeneracy. We form the matrix elements $W_{ij} = \langle \psi_i^0 | H' | \psi_j^0 \rangle$, where i, j go from $1, 2, \ldots, n$. We then get an $n \times n$ matrix, with n separate roots (some of which may be zero, e.g. if the matrix is very sparse).

16.3 Example: 3D square well

Consider $V(x, y, z) = 0$ for $0 < x < a$, $0 < y < a$, and $0 < z < a$, and otherwise its $V(x, y, z) = \infty$. This has eigenfunctions and eigenvalues:

$$\psi_{n_x, n_y, n_x}^0(x, y, z) = \left(\frac{2}{a}\right)^{3/2} \sin(n_x \pi x/a) \sin(n_y \pi y/a) \sin(n_z \pi z/a), \qquad \text{(3D-SqWell)}$$

$$E_{nx, ny, nz}^0 = \frac{\pi^2 \hbar^2}{2\mu a^2}(n_x^2 + n_y^2 + n_z^2).$$

The *ground state* is non-degenerate, but the first excited state is *triply degenerate*, as $\psi_{211}, \psi_{121}, \psi_{112}$ all have the same energy.

We introduce a perturbation: an additional potential energy equal to V_0 for $0 < x < a/2$ and $0 < y < a/2$ (for all values of z).

We now consider the first excited state, which is 3-fold degenerate. Letting $\psi_1 = \psi_{211}$, $\psi_2 = \psi_{121}$ and $\psi_3 = \psi_{112}$, we can calculate each matrix element.

$$W_{11} = V_0 \left(\frac{2}{a}\right)^3 \int_0^{a/2} \sin^2(2\pi x/a)\, dx \int_0^{a/2} \sin^2(\pi y/a)\, dy \int_0^a \sin^2(\pi z/a)\, dz,$$

where the integrals can be evaluated using

$$\frac{2}{a} \int_0^{a/2} \sin^2(n\pi x/a)\, dx = \frac{2}{a} \left[\frac{x}{2} - \frac{\sin(2n\pi x/a)}{4n\pi/a} \right]_0^{a/2} = \frac{2}{a} \left(\frac{a}{4} \right) = \frac{1}{2},$$

(for $n = 1$ or 2) — a result that follows trivially from $\langle \sin^2 \theta \rangle = 1/2$ (over any integral number of half cycles). Hence, $W_{11} = V_0 \times (1/2) \times (1/2) \times 1 = V_0/4$, and similarly $W_{22} = W_{33} = V_0/4$. The off-diagonal matrix elements are a little trickier:

$$W_{12} = V_0 \left(\frac{2}{a} \right)^3 \int_0^{a/2} \sin(2\pi x/a) \sin(\pi x/a)\, dx \int_0^{a/2} \sin(2\pi y/a) \sin(\pi y/a)\, dy \int_0^a \sin^2(\pi z/a)\, dz,$$

where

$$\frac{2}{a} \int_0^{a/2} \sin(2\pi x/a) \sin(\pi x/a)\, dx = \frac{2}{a} \left[\frac{\sin(\pi x/a)}{2\pi/a} - \frac{\sin(3\pi x/a)}{6\pi/a} \right]_0^{a/2} = \frac{2}{a} \left(\frac{a}{2\pi} + \frac{a}{6\pi} \right) = \frac{4}{3\pi},$$

where we used $2 \sin A \sin B = \cos(A - B) - \cos(A + B)$. Hence, $W_{12} = V_0 \times (4/3\pi) \times (4/3\pi) \times 1 = 16V_0/9\pi^2 = \kappa V_0/4$, where we have defined $\kappa = 64/9\pi^2$ for convenience. Noting that (essentially by definition of orthogonality)

$$\frac{2}{a} \int_0^a \sin(2\pi z/a) \sin(\pi z/a)\, dz = \frac{2}{a} \left[\frac{\sin(\pi x/a)}{2\pi/a} - \frac{\sin(3\pi x/a)}{6\pi/a} \right]_0^a = 0,$$

we deduce that $W_{13} = W_{23} = 0$. Noting further that $W_{ij} = W_{ji}^*$, we can determine the full matrix equation

$$\begin{pmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{22} & W_{23} \\ W_{31} & W_{32} & W_{33} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \frac{V_0}{4} \begin{pmatrix} 1 & \kappa & 0 \\ \kappa & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = E_2^1 \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 1 & \kappa & 0 \\ \kappa & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \frac{4E_2^1}{V_0} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}.$$

We let $w = 4E^1/V_0$, and subtract, to get

$$\begin{pmatrix} 1-w & \kappa & 0 \\ \kappa & 1-w & 0 \\ 0 & 0 & 1-w \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

where the non-trivial solution (i.e., $\alpha, \beta, \gamma \neq 0$) is when the determinant of the matrix is zero.

16.4   Formalism: Meaning of the vector components

The state space being considered in this example is the space of states which are *any superposition* of the states $\psi_{211}, \psi_{121}, \psi_{112}$, defined as wavefunctions in x, y, and z in Eq. (3D-SqWell). We express such a general superposition as $\alpha\psi_{211} + \beta\psi_{121} + \gamma\psi_{112}$

The *vector components* α, β, and γ are the *amplitudes* of the states ψ_{211}, ψ_{121}, and ψ_{112} in a general superposition state.

16.5 Eigenvalues and eigenvectors

The determinant is given by

$$(1-w)\begin{vmatrix} 1-w & \kappa \\ \kappa & 1-w \end{vmatrix} = 0$$

$$\Rightarrow (1-w)[(1-w)^2 - \kappa^2] = 0$$

The roots are therefore: $1-w = 0$, i.e., $w = 1$; and $(1-w)^2 = \kappa^2$, i.e., $1-w = \pm\kappa$ so $4E^1/V_0 = 1$ and $4E^1/V_0 = 1 \mp \kappa$. Hence, $E^1 = V_0/4$, $V_0/4 - 16V_0/(9\pi^2)$ and $V_0/4 + 16V_0/(9\pi^2)$.

We now substitute each value of w back into the matrix — first $w = 1$:

$$\begin{pmatrix} 0 & \kappa & 0 \\ \kappa & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

meaning that $\kappa\beta = 0$ and $\kappa\alpha = 0$ ($\Rightarrow \alpha = \beta = 0$, as $\kappa \neq 0$) and the only nonzero component is γ. Hence, the eigenvector for $w = 1$ is just $|\psi_3\rangle$ (which is just ψ_{112}). Similarly, for $w = 1 - \kappa$:

$$\begin{pmatrix} \kappa & \kappa & 0 \\ \kappa & \kappa & 0 \\ 0 & 0 & \kappa \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

so $\kappa\gamma = 0$ and $\kappa\alpha + \kappa\beta = 0 \Rightarrow \alpha = -\beta$. Hence, the eigenstate is $(|\psi_2\rangle - |\psi_1\rangle)/\sqrt{2}$, which is the same as $(\psi_{121} - \psi_{211})/\sqrt{2}$. For $w = 1 + \kappa$ the eigenstate is $(|\psi_2\rangle + |\psi_1\rangle)/\sqrt{2}$ [corresponds to $(\psi_{121} + \psi_{211})/\sqrt{2}$].

It can be checked that with this basis all off-diagonal terms $W_{ij} = 0$ for $j \neq i$. Letting $|\psi_c\rangle = |\psi_3\rangle$, $|\psi_b\rangle = (|\psi_2\rangle - |\psi_1\rangle)/\sqrt{2}$, and $|\psi_a\rangle = (|\psi_2\rangle + |\psi_1\rangle)/\sqrt{2}$:

$$W_{cc} = W_{33} = \frac{V_0}{4},$$

$$W_{cb} = \langle\psi_c|H'|\psi_b\rangle = \frac{1}{\sqrt{2}}(\langle\psi_3|H'|\psi_2\rangle - \langle\psi_3|H'|\psi_1\rangle) = 0,$$

$$W_{ca} = \langle\psi_a|H'|\psi_c\rangle = \frac{1}{\sqrt{2}}(\langle\psi_3|H'|\psi_2\rangle + \langle\psi_3|H'|\psi_1\rangle) = 0,$$

$$W_{bb} = \langle\psi_b|H'|\psi_b\rangle = \frac{1}{2}(\langle\psi_2|H'|\psi_2\rangle - \langle\psi_1|H'|\psi_2\rangle - \langle\psi_2|H'|\psi_1\rangle + \langle\psi_1|H'|\psi_1\rangle)$$

$$= \frac{1}{2}\left(2\frac{V_0}{4} - 2\frac{\kappa V_0}{4}\right) = (1-\kappa)\frac{V_0}{4},$$

$$W_{ba} = \langle\psi_b|H'|\psi_a\rangle = \frac{1}{2}(\langle\psi_2|H'|\psi_2\rangle - \langle\psi_1|H'|\psi_2\rangle + \langle\psi_2|H'|\psi_1\rangle - \langle\psi_1|H'|\psi_1\rangle) = 0,$$

$$W_{aa} = \langle\psi_a|H'|\psi_a\rangle = \frac{1}{2}(\langle\psi_2|H'|\psi_2\rangle + \langle\psi_1|H'|\psi_2\rangle + \langle\psi_2|H'|\psi_1\rangle + \langle\psi_1|H'|\psi_1\rangle) = \frac{V_0}{4}(1+\kappa).$$

Hence the matrix equation is given, in terms of the $|\psi_a\rangle$, $|\psi_b\rangle$, $|\psi_c\rangle$ basis, by

$$\begin{pmatrix} (1+\kappa)V_0/4 - E^1 & 0 & 0 \\ 0 & (1-\kappa)V_0/4 - E^1 & 0 \\ 0 & 0 & V_0/4 - E^1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0,$$

where we have $E^1 = (1+\kappa)V_0/4, (1-\kappa)V_0/4, V_0/4$. Had we chosen $|\psi_a\rangle$, $|\psi_b\rangle$ and $|\psi_c\rangle$ at the start we could have used non-degenerate perturbation theory with $E_a^1 = \langle\psi_a|H'|\psi_a\rangle$, etc..

17 Degenerate perturbation theory and hydrogen [Gri 6.3]

17.1 Degeneracy in hydrogen

Recall that each energy level n in Hydrogen is n^2 degenerate (without spin) or $2n^2$ degenerate once we count spin as well, bringing in an additional quantum number $m_s = \pm 1/2$.

17.2 Spin-orbit coupling [Gri 6.3.2]

The perturbation to the potential from the magnetic dipole moment generated by electron spin, μ_s, is $H'_{so} = -\underline{\mu}_s \cdot \underline{B}$, where $\underline{B}$ is the external magnetic field. As described in lecture 13, $\underline{\mu}_s = -g_s(e/2m_e)\underline{S}$, where the electron spin g-factor g_s is ≈ 2. Hence, we let $\underline{\mu}_s = -e/m_e\underline{S}$.

For orbitals with angular momentum (i.e., all except $l = 0$), viewed classically one can think of the electron orbiting the nucleus — but from the electron's point of view it is the positive nucleus which orbits around it. This sets up a magnetic field from the "current loop" of magnitude $B = \mu_0 I/2r$ (with the effective current $I = e/P$, where $P = 2\pi r/v$ is the orbital period, v its speed, and r its radius).

The magnitude of the orbital angular momentum can be related to the B-field via

$$L = r m_e v = \frac{r m_e 2\pi r}{P} = \frac{2\pi m_e r^2}{P} = \frac{2\pi m_e r^2 I}{e} = \frac{4\pi m_e r^3 B}{\mu_0 e}.$$

More fully, $\underline{B}$ and $\underline{L}$ are vectors, both pointing in the same direction. Noting that $\epsilon_0 \mu_0 = 1/c^2$,

$$\underline{L} = \frac{4\pi \epsilon_0 c^2 m r^3}{e} \underline{B}.$$

Hence,

$$H'_{so} = -\underline{\mu}_s \cdot \underline{B} = \frac{e}{m} \underline{S} \cdot \frac{e}{4\pi \epsilon_0 c^2 m r^3} \underline{L} = \frac{e^2}{4\pi \epsilon_0 c^2 m^2 r^3} \underline{S} \cdot \underline{L}.$$

However, it turns out there is a significant effect due to special relativity,[14] resulting in the correct perturbation term being approximately half in magnitude:

$$H'_{so} = \frac{e^2}{8\pi \epsilon_0 c^2 m^2 r^3} \underline{S} \cdot \underline{L}.$$

17.3 Commutation properties

Our Hamiltonian therefore has an extra term proportional to $\underline{S} \cdot \underline{L} = S_x L_x + S_y L_y + S_z L_z$. Recall that, if the perturbation H'_{so} commutes with H^0, we can use $E_i^1 = \langle \psi_i^0 | H'_{so} | \psi_i^0 \rangle$, even with degenerate levels.

We might consider ψ_{nlmm_s}, i.e., use the joint eigenfunctions of H^0, L^2, L_z, S^2, S_z labelled by our standard quantum numbers n, m, l, m_s (where $m_s = \pm 1/2$, as $s = 1/2$). These operators however do not commute with the perturbation $\underline{S} \cdot \underline{L}$ (because $\underline{S} \cdot \underline{L}$ contains terms with L_x and L_y, which do not commute with L_z, and terms S_x, S_y which do not commute with S_z).

However, if we consider the **total angular momentum** $\underline{J} = \underline{L} + \underline{S}$, then J^2, J_z do commute with all our original operators *and* with the perturbation.

---

[14]This is called *Thomas precession*, and takes into account relativistic time dilation between the electron and the nucleus.

Considering e.g. $J^2 = (\underline{L} + \underline{S}) \cdot (\underline{L} + \underline{S}) = L^2 + S^2 + 2\underline{L} \cdot \underline{S}$ (as $\underline{S} \cdot \underline{L} = \underline{L} \cdot \underline{S}$, since $\underline{S}$ and $\underline{L}$ commute), it follows that

$$
\begin{aligned}
[J^2, \underline{L}.\underline{S}] &= [L^2 + S^2 + 2\underline{L} \cdot \underline{S}, \ \underline{L} \cdot \underline{S}] \\
&= [L^2, L_x S_x + L_y S_y + L_z S_z] + [S^2, L_x S_x + L_y S_y + L_z S_z] + 2[\underline{L} \cdot \underline{S}, \underline{L} \cdot \underline{S}] \\
&= [L^2, L_x]S_x + [L^2, L_y]S_y + [L^2, L_z]S_z + [S^2, S_x]L_x + [S^2, S_y]L_y + [S^2, S_z]L_z \\
&= 0,
\end{aligned}
$$

as all components of an angular momentum commute with total square of that angular momentum.

Total angular momentum J is conserved, but the S and L parts of it are not. It therefore makes sense to think about joint eigenfunctions of the operators H^0, L^2, J^2, S^2, J_z.

17.4   Adding angular momenta

Consider the relatively simple case of $J_z = L_z + S_z$. L_z has values $m_l \hbar$, where $m_l = -l, -l + 1, \ldots, 0, \ldots, l-1, l$, and l is the quantum number from the eigenvalues of L^2, given by $l(l+1)\hbar^2$. S_z likewise has values $m_s \hbar$ where $m_s = -s...s$, where s is the quantum number from the eigenvalues of S^2, which are $s(s+1)\hbar^2$. Here $s = 1/2$ for *all* electrons, and so $m_s = -1/2$ or $+1/2$

Denoting $J_z = m_J \hbar$, the maximum value of m_J is therefore the sum of the maximum values of m_l and s_l, i.e., max $m_J = l + s$. Similarly, min $m_J = -l - s$. Hence, m_J runs from $-l - s$ to $l + s$ in integer steps, so it must correspond to a value $j = l + s$ drawn from an eigenvalue $j(j+1)\hbar^2$ of the operator J^2.

A way of visualising this (see Lecture 10) is to draw $\underline{L}$ as a vector in L_x, L_y, L_z space, with length $\sqrt{l(l+1)}\hbar$, and projection $m_l \hbar$ onto the L_z axis. It can be oriented anywhere around L_x, L_y forming a cone of possible values for each m_l; hence, the full sequence corresponds to cones with $m_l = -l$, $m_l = -l+1$, and so on up to $m_l = l$ ($2l + 1$ possible cones, each with their own m_l value, for each l).

We can similarly draw $\underline{S}$ in the S_x, S_y, S_z space as a vector of length $\sqrt{s(s+1)}\hbar$ (cones with $S_z = -s$ up to s). Within this picture the maximum possible magnitude for $\underline{J}$ is obtained by taking the maximal m_l and m_s and have their directions aligned as closely as possible. This gives a vector $\underline{J}$ of length $\approx (l + s)\hbar$. The minimal $\underline{J}$ is when we have the maximum m_l and the minimum m_s, and have them as *anti-aligned* as possible, giving a vector $\underline{J}$ of length $\approx |l - s|\hbar$.

From this we can deduce that when we add orbital and spin angular momenta, to make
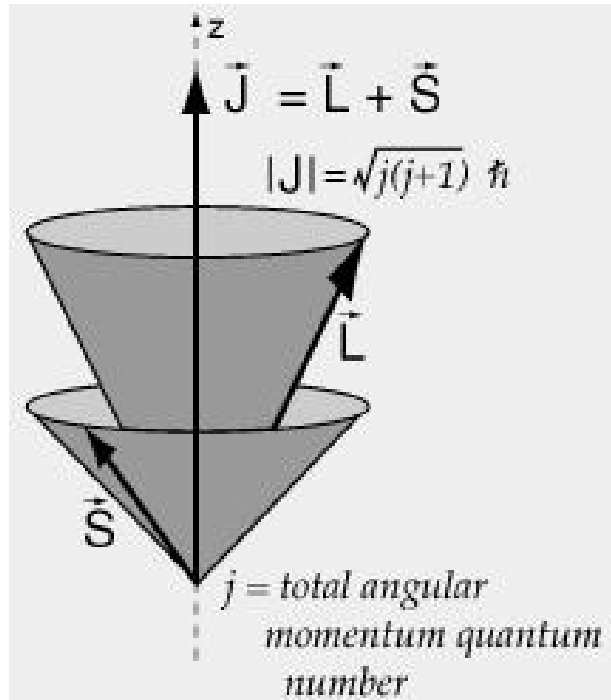

Figure 12: Angular momentum addition in the Vector Model

m_l	m_s	$m_j = m_l + m_s$
1	1/2	3/2
1	-1/2	1/2
0	1/2	1/2
0	-1/2	-1/2
-1	1/2	-1/2
-1	-1/2	-3/2

Table 1: Possible values of m_j resulting from $l = 1$ and $s = 1/2$

$\underline{J} = \underline{L} + \underline{S}$ then j can take values $|l - s| \leq j \leq l + s$ in integer steps, and for each one of these then m_j takes values $-j, \ldots, j$ in integer steps.

In our specific case $s = 1/2$ and there is only one integer between $|l - s| \leq j \leq l + s$. Hence, j takes values $|l - 1/2|$ and $l + 1/2$, so in general two different values of j can be taken. In the special case of $l = 0$ these are both the same, with $j = 1/2$. for $l \neq 0$ then $j = l - 1/2$ and $l + 1/2$. And since this is a general angular momentum, we also know that J^2 has eigenvalues $j(j + 1)\hbar^2$, with the corresponding $-j \leq m_j \leq j$ taking values in integer steps.

17.5   Example of adding angular momentum

We consider $l = 1$ (so this means $n \geq 2$) and $s = 1/2$. We could have $m_l = -1, 0, 1$ and for each state we can have $m_s = \pm 1/2$. This tells us that the eigenvalues of $J_z = L_z + S_z$ are $m_j \hbar$ where $m_j = m_l + m_s$. We enumerate this in Table 1.

Here we notice something interesting — there are two ways to get $m_j = 1/2$ or $-1/2$, but only 1 way to get $m_j = 3/2$ or $-3/2$. We get more insight by writing this out in j as well. We know that $l = 1$ and $s = 1/2$ can give us total $j = |l - s| = 1/2$ or $l + s = 3/2$, and so can order these values as in Table 2.

j	m_j
3/2	3/2
3/2	1/2
3/2	-1/2
3/2	-3/2
1/2	1/2
1/2	-1/2

Table 2: Possible values of m_j resulting from $j = |l - s|$ and $j = l + s$

We see that the 'extra' two states come from the fact that j can have $j = l + s, \ldots, |l - s|$, i.e., that $j = 3/2$ or $1/2$. Had we considered only $j = l + s = 3/2$ we wouldn't have had enough! Hence, if we consider $l = 1$ and $s = 1/2$, then in general we have a $4/6 = 2/3$ chance to have $j = 3/2$ and a $1/3$ chance to have $j = 1/2$.

18 Degenerate perturbation theory and hydrogen II [Gri 6.3]

18.1 Meaning of the spin–orbit term

Recalling (from lecture 17) that $J^2 = (\underline{L} + \underline{S}).(\underline{L} + \underline{S}) = L^2 + S^2 + 2\,\underline{L} \cdot \underline{S}$. Hence,

$$\underline{L} \cdot \underline{S} = \frac{1}{2}(J^2 - L^2 - S^2),$$

and the eigenvalues of $\underline{L} \cdot \underline{S}$ are therefore the eigenvalues of $(J^2 - L^2 - S^2)/2$, i.e.,

$$\frac{\hbar^2}{2}[j(j+1) - l(l+1) - s(s+1)] = \frac{\hbar^2}{2}\left[j(j+1) - l(l+1) - \frac{3}{4}\right],$$

where we have used the fact that the spin $s = 1/2$.

Hence, the first-order correction to the energy eigenvalues [where $H'_{\text{so}} = (e^2/8\pi\epsilon_0 c^2 m^2 r^3)\underline{S} \cdot \underline{L}$] is given by (we label the unperturbed eigenstates by the quantum numbers n, l, j):

$$E^1_{nlj,\text{so}} = \langle \psi_{nlj}|H'_{\text{so}}|\psi_{nlj}\rangle = \frac{e^2}{8\pi\epsilon_0 c^2 m^2}\frac{\hbar^2}{2}\left[j(j+1) - l(l+1) - \frac{3}{4}\right]\langle \psi_{nlj}|r^{-3}|\psi_{nlj}\rangle.$$

Because the $|\psi_{nlj}\rangle$ are, by construction, eigenstates of $\underline{S} \cdot \underline{L}$, this operator quantity can be replaced by its appropriate eigenvalues, which can then (they are just numbers) be "pulled outside" the expectation value expression.

We are left with evaluating the expectation value of r^{-3}, for the Hydrogen wavefunctions from Lecture 11 which is given by $\langle r^{-3}\rangle = [l(l+1/2)(l+1)n^3 a^3]^{-1}$. Hence,[15]

$$E^1_{nlj,\text{so}} = \frac{e^2\hbar^2}{16\pi\epsilon_0 c^2 m^2}\frac{j(j+1) - l(l+1) - 3/4}{l(l+1/2)(l+1)n^3 a^3}.$$

However, from the definition of Bohr radius a, we can set $e^2/(4\pi\epsilon_0) = \hbar^2/(am)$, giving

$$E^1_{nlj,\text{so}} = \frac{\hbar^4}{4a^4 n^3 m^3 c^2}\frac{j(j+1) - l(l+1) - 3/4}{l(l+1/2)(l+1)},$$

and we also have $E^0_n = -\hbar^2/(2ma^2 n^2)$, meaning that we can set $\hbar^4/(4m^2 a^4 n^4) = (E^0_n)^2$. Hence,

$$E^1_{nlj,\text{so}} = \frac{(E^0_n)^2}{mc^2}\frac{n[j(j+1) - l(l+1) - 3/4]}{l(l+1/2)(l+1)}, \qquad \text{(Spin-Orbit Term)}$$

where it is assumed that $l \neq 0$ (fortunately, as for $l = 0$ the denominator appears to be $= 0$). This **spin-orbit coupling** typically leads to $E^1_n/E^0_n \sim E^0_n/mc^2$.

18.2 Correction to the electron–proton potential energy

However, spin-orbit coupling is not the only thing of importance when it comes to getting the energy levels exactly right. There are two more terms which are important at this level.

The form of the potential energy for the interaction between the electron and the proton assume both to be point charges. However neither particle can have a position defined more accurately than is Compton wavelength, $\lambda_{\text{C}} = \hbar/mc$. We can think of the charge of the electron

---

[15]Note that $a \equiv (4\pi\epsilon_0\hbar^2)/(\mu Z e^2)$ is (with $Z = 1$) the *Bohr radius* (see Lecture 11).

(more correctly the electrostatic interaction) being smeared out over a volume $(\hbar/mc)^3$ (The corresponding smearing for the proton is smaller by a factor m_p/m_e and so can be neglected.) Hence there is an additional correction to the potential energy if the electron is at the same position as the proton — called the **Darwin term**. [16]

$$H_{\mathrm{D}}^1 = \frac{\pi\hbar^2}{2m^2c^2}\frac{e^2}{4\pi\epsilon_0}\delta(\underline{r}).$$

This is independent of (doesn't care about) spin, and commutes with all the angular momentum operators. We can therefore use non-degenerate perturbation theory.

$$E_n^1 = \frac{\pi\hbar^2}{2m^2c^2}\frac{e^2}{4\pi\epsilon_0}\int_V \psi_{nlm}^*\delta(\underline{r})\psi_{nlm}\,dV$$

This only acts at the origin and therefore only acts on states which are nonzero at the origin. Only when $l = 0$ is $R_{nl}(r = 0) \neq 0$, so it acts only on the $l = 0$ — which must also have $m = 0$ (s-state) wavefunctions.

$$E_n^1 = \frac{\pi\hbar^2}{2m^2c^2}\frac{e^2}{4\pi\epsilon_0}|\psi_{n00}(0)|^2.$$

At $r = 0$ all the Laguerre polynomials for $l = 0$ are $L_{n-1}^1(0)$, and so this is just the constant term in the polynomial. Hence, we get $\psi_{n00}(0) = 1/(\pi a^3 n^3)$, and

$$E_n^1 = \frac{\pi\hbar^2}{2m^2c^2}\frac{e^2}{4\pi\epsilon_0}\frac{1}{\pi a^3 n^3} = \frac{\pi\hbar^2}{2m^2c^2}\frac{\hbar^2}{am}\frac{1}{\pi a^3 n^3} = \frac{\hbar^4}{2m^3c^2}\frac{1}{a^4 n^3}$$

Recall that $(E_n^0)^2 = \hbar^4/(4m^2a^4n^4)$, hence

$$E_n^1 = \frac{2}{mc^2}(E_n^0)^2 n. \qquad\qquad \text{(Darwin Term)}$$

18.3 Relativistic corrections to kinetic energy [Gri 6.3.1]

In special relativity, the kinetic energy is given by $E^2 = (T + mc^2)^2 = p^2c^2 + m^2c^4$. Hence,

$$T + mc^2 = mc^2\sqrt{1 + \frac{p^2}{m^2c^2}}.$$

One can carry out a Taylor expansion on the square root, to get

$$T + mc^2 = mc^2\left[1 + \frac{1}{2}\frac{p^2}{m^2c^2} - \frac{1}{8}\left(\frac{p^2}{m^2c^2}\right)^2 + \cdots\right].$$

Hence,

$$T \approx \frac{p^2}{2m} - \frac{1}{8}\frac{p^4}{m^3c^2} = T^0 - \frac{1}{2mc^2}(T^0)^2,$$

where $T^0 = p^2/2m$ and the lowest order relativistic correction is $H_{\mathrm{r}}' = -p^4/8m^3c^2 = -(T^0)^2/2mc^2 = -[H^0 - V(r)]^2/2mc^2$.

---

[16]Note that $\delta(\underline{r} - \underline{r}_0)$ is a *three dimensional Dirac delta function* [evaluated in Eq. (6) at $\underline{r}_0 = 0$, hence the term is proportional to $\delta(\underline{r})$]. This is in some sense a shorthand: $\delta(\underline{r} - \underline{r}_0) = \delta(x - x_0)\delta(y - y_0)\delta(z - z_0)$ in cartesian coordinates. In spherical polar coordinates, $\delta(\underline{r} - \underline{r}_0) = (r^2\sin\theta)^{-1}\delta(r - r_0)\delta(\theta - \theta_0)\delta(\phi - \phi_0)$, which simplifies to $\delta(\underline{r} - \underline{r}_0) = (4\pi r^2)^{-1}\delta(r - r_0)$ if there is no θ or ϕ dependence to the problem at hand.

Each level is $2n^2$ degenerate. But this perturbation commutes (more or less) with H^0 — the perturbation does not act on the spin, and is spherically symmetric and so commutes with L^2 and L_z. Hence, the degenerate states belonging to E_n^0 are not connected to first order to H_r'. so actually we can use non-degenerate perturbation theory (in effect, we have choosen "good" eigenvectors and eigenvalues by using n, l, m and m_s). We therefore have

$$E_n^1 = -\frac{1}{2mc^2} \int \psi_n^{0*} [H^0 - V(r)]^2 \psi_n^0 \, dV \, .$$

We know $H^0 \psi_n^0 = E_n^0 \psi_n^0$, and so[17]

$$E_n^1 = -\frac{1}{2mc^2} \int \psi_n^* [(E_n^0)^2 - 2E_n^0 V(r) + V(r)^2] \psi_n \, dV$$

$$= -\frac{1}{2mc^2} \left[(E_n^0)^2 \int \psi_n^* \psi_n \, dV - 2E_n \int \psi_n^* V(r) \psi_n \, dV + \int \psi_n^* V(r)^2 \psi_n \, dV \right]$$

$$= -\frac{1}{2mc^2} \left[(E_n^0)^2 - 2E_n \int \psi_n^* V(r) \psi_n \, dV + \int \psi_n^* V(r)^2 \psi_n \, dV \right]$$

$$= -\frac{1}{2mc^2} \left[(E_n^0)^2 - 2E_n^0 \langle V(r) \rangle + \langle V(r)^2 \rangle \right] \, .$$

We can get $\langle V \rangle = [-e^2/(4\pi\epsilon_0)]\langle 1/r \rangle$ and then use the virial theorem to get $\langle 1/r \rangle = 1/an^2$. Similarly (but less easily) $\langle V^2 \rangle = [-e^2/(4\pi\epsilon_0)]^2 \langle 1/r^2 \rangle$, and

$$\left\langle \frac{1}{r^2} \right\rangle = \frac{1}{a^2 n^3 (l + 1/2)} \, .$$

We can substitute these expressions in, and also make use of $a = 4\pi\epsilon_0 \hbar^2 / me^2 \Rightarrow e^2/4\pi\epsilon_0 = \hbar^2/am$, to get

$$E^1 = -\frac{(E_n^0)^2}{2mc^2} \left[1 + \left(\frac{-e^2}{4\pi\epsilon_0} \right) \frac{(2/E_n^0)}{an^2} + \left(\frac{-e^2}{4\pi\epsilon_0} \right)^2 \frac{1/(E_n^0)^2}{a^2 n^3 (l + 1/2)} \right]$$

$$= -\frac{(E_n^0)^2}{2mc^2} \left[1 + \frac{\hbar^2}{am} \frac{(2/E_n^0)}{an^2} + \frac{\hbar^4}{a^2 m^2} \frac{(1/E_n^0)^2}{a^2 n^3 (l + 1/2)} \right] \, .$$

Finally, we can also write (again making use of $a = 4\pi\epsilon_0 \hbar^2 / me^2$)

$$E_n^0 = -\frac{m}{2\hbar^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \frac{1}{n^2} = -\frac{m}{2\hbar^2} \frac{\hbar^4}{m^2 a^2} \frac{1}{n^2} = -\frac{\hbar^2}{2ma^2} \frac{1}{n^2} \, ,$$

meaning that we can write down a less cumbersome final expression:

$$E^1 = -\frac{(E_n^0)^2}{2mc^2} \left[1 - 4 + \frac{4n}{(l + 1/2)} \right] = -\frac{(E_n^0)^2}{2mc^2} \left[\frac{4n}{(l + 1/2)} - 3 \right] \, . \qquad \text{(Relativistic Term)}$$

---

[17]Contrary to what is stated in Griffiths, p^4 is indeed, in full generality, a Hermitian operator (see, for example, http://www.reed.edu/physics/faculty/griffiths/QM2b.pdf).

19 Hydrogen fine structure

19.1 Total first-order energy correction

If we add add all the first-order energy correction terms (spin–orbit, Darwin, and relativistic) together, we get

$$E_{nj}^1 = \frac{(E_n^0)^2}{2mc^2}\left(3 - \frac{4n}{j+1/2}\right).$$

It is common to write E_n^0 in another way:

$$E_n^0 = -\frac{\hbar^2}{2ma^2}\frac{1}{n^2} = -\alpha^2\frac{mc^2}{2n^2},$$

where $\alpha = e^2/(4\pi\epsilon_0\hbar c) \approx 1/137$ is the dimensionless **fine structure constant.**[18] Hence, E_n^0 is of order α^2, and E_{nj}^1 is of order α^4. To see this explicitly:

$$E_{nj} = E_n^0 + E_n^0\alpha^2\frac{mc^2}{2n^2 2mc^2}\left[-3 + \frac{4n}{(j+1/2)}\right] = E_n^0\left(1 + \frac{\alpha^2}{n^2}\left[\frac{n}{(j+1/2)} - \frac{3}{4}\right]\right),$$
$$\text{(Fine Structure)}$$

and so these corrections are of order $\alpha^2 \sim 10^4\times$ smaller than the unperturbed energy.

Collectively, these additional terms, and the additional level of structure in the spectral lines they reveal, are known as the *fine structure.*

19.2 Ground state of hydrogen

The ground state of hydrogen has $l = 0$, $j = |l - s|$ and $j = l + s$ so there is only one value for j, which is $j = 1/2$. Hence, there is a relative correction to the ground state energy of

$$\frac{\Delta E_1}{E_1^0} = \frac{\alpha^2}{n^2}\left(\frac{n}{j+1/2} - \frac{3}{4}\right) = \alpha^2\left(1 - \frac{3}{4}\right) = \frac{\alpha^2}{4}.$$

hence $\Delta E_1 = E_1^0\alpha^2/4 = 1.33\times10^{-5}E_1^0 = -1.81\times10^{-4}$ eV. So the ground state energy is *slightly* lower (meaning larger in magnitude, but more negative), at $(-13.6 - 1.81 \times 10^{-4})$ eV.

19.3 Quantum number sets

Our unperturbed energy levels were labelled with n, l, m, m_s, i.e., the quantum numbers corresponding to operators H^0, L^2, L_z, S_z [in principle there is S^2 as well, but as this has quantum number $s = 1/2$ and eigenvalue $s(s + 1)\hbar^2 = 3/4\hbar^2$ for all electrons it is not useful as a label].

Our new quantum number sets label the wavefunctions which follow all these fine structure perturbations. **These are n, l, j, m_j, corresponding to H^0, L^2, J^2, J_z** (again, we in principle have S^2 as well, but it is not useful as a label).

---

[18]The currently recommended value for the inverse fine structure constant is $\alpha^{-1} = 137.035\,999\,139 \pm 0.000\,000\,031$ (see http://physics.nist.gov/cgi-bin/cuu/Value?alph). See also http://physics.nist.gov/cuu/Constants/alpha.html for a brief discussion putting the fine structure constant, and how it may be measured, into context; and T. Aoyama, M. Hayakawa, T. Kinoshita, M. Nio, Phys. Rev. Lett. **109**, 111807 (2012), which effectively uses a form of tenth order perturbation theory in the context of quantum electrodynamics (calculating a total of 12 672 terms) to relate a high-precision measurement of the electron's anomalous magnetic moment to a high precision value of α^{-1}.

$\underline{J} = \underline{L} + \underline{S}$, i.e., is a sum of orbital and spin angular momentum of the electron. As such, it is another angular momentum and so **by definition J^2 has eigenvalues $j(j+1)\hbar^2$ and J_z has eigenvalues $m_j\hbar$.**

The quantum number j takes values from $|l-1/2|$ to $l+1/2$ in integer steps, and so (as $s = 1/2$) j can take 2 values in general: $l - 1/2$ and $l + 1/2$, except for $l = 0$, where $j = 1/2$. **Hence, m_j can take any value from $-j, \ldots, j$ in integer steps.**

19.4 Balmer line energies

The Balmer line transition is from $n = 3$ to $n = 2$. In the analytic solution of the Schrödinger equation used to model the hydrogen atom in Lectures 11, we had $E_n = -13.6/n^2$ eV. This implies the line should have energy

$$E_3 - E_2 = -13.6\left(\frac{1}{3^2} - \frac{1}{2^2}\right) = 1.889 \text{ eV}.$$

We now have an energy which depends on j as well as n, however, and so the $n = 3$ level splits. We can keep track by using n, l, j, m_j, however it is very common to label levels with n and a *letter* for l, such that $l = 0, 1, 2, 3, \ldots$ is $s, p, d, f, \ldots$, and then the value of j as a subscript.[19]

Figure 13: Fine structure energy splitting

Hence, from $n = 3$, where we can have:

- $n = 3, l = 0$ means $j = 1/2$ ($3s_{1/2}$),
- $n = 3, l = 1$ means $j = 1/2$ and $3/2$ ($3p_{1/2}$ and $3p_{3/2}$),
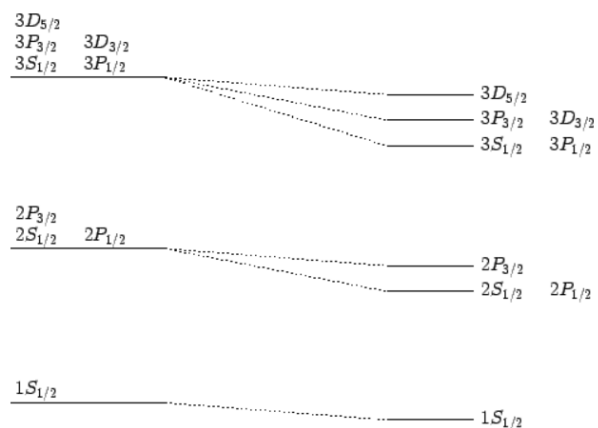- $n = 3, l = 2$ means $j = 3/2$ and $5/2$ ($3d_{1/2}$ and $3d_{5/2}$),

we drop to $n = 2$, where we can have:

- $n = 2, l = 0$ means $j = 1/2$ ($2s_{1/2}$),
- $n = 2, l = 1$ means $j = 1/2$ and $3/2$ ($2p_{1/2}$ and $2p_{3/2}$).

There are therefore 5 different labels for the initial $n = 3$ state, but only 3 with distinct n, j values, i.e., with distinct energies. We began with the level as $2n^2 = 18$ degenerate — we have now split the 18 possible degenerate states into 3 different energies, but there is still a great deal of degeneracy.

Similarly, there are 3 different labels for the final $n = 2$ state, but only 2 distinct n, j, i.e., only 2 distinct energies. Since this level has degeneracy 8 before we considered the perturbations it is also still highly degenerate.

It follows that there are in principle 6 possible different transition energies:

---

[19]The letters s, p, d, f actually stand for sharp, principal, diffuse, and fundamental, referring to the visual impression left by the fine structure of the spectral lines before their cause was clearly understood. Subsequent letters follow the alphabet (g, h, etc.).

- $n = 3, j = 1/2$ ($l = 0$ and 1) $3s_{1/2}$ and $3p_{1/2}$, going to
  $n = 2, j = 1/2$ ($l = 0$ and 1) $2s_{1/2}$ and $2p_{1/2}$;

- $n = 3, j = 1/2$ ($l = 0$ and 1) $3s_{1/2}$ and $3p_{1/2}$, going to
  $n = 2, j = 3/2$ ($l = 1$) $2p_{3/2}$;

- $n = 3, j = 3/2$ ($l = 1$ and 2) $3p_{3/2}$ and $3d_{3/2}$, going to
  $n = 2, j = 1/2$ ($l = 0$ and 1) $2s_{1/2}$ and $2p_{1/2}$;

- $n = 3, j = 3/2$ ($l = 1$ and 2) $3p_{3/2}$ and $3d_{3/2}$, going to
  $n = 2, j = 3/2$ ($l = 1$) $2p_{3/2}$;

- $n = 3, j = 5/2$ ($l = 2$) $3d_{5/2}$, going to
  $n = 2, j = 1/2$ ($l = 0$ and 1) $2s_{1/2}$ and $2p_{1/2}$;

- $n = 3, j = 5/2$ ($l = 2$) $3d_{5/2}$, going to
  $n = 2, j = 3/2$ ($l = 1$) $2p_{3/2}$;

however we in fact only observe two strong spectral lines. Why?

Essentially, the emitted photon carries angular momentum of magnitude $\hbar$, and so to conserve angular momentum we need to have a change in l of $\Delta l = \pm 1$ for a radiative transition. This corresponds to $\Delta j = 0, \pm 1$ for allowed transitions (electric dipole radiation). Other transitions are called **forbidden**[20] — they do happen but with much much less frequency.

Hence, these **selection rules** (due to conservation of angular momentum) remove some of the above possibilities, leaving only:

- $3s_{1/2} \to 2p_{1/2}$ and $3p_{1/2} \to 2s_{1/2}$,
- $3s_{1/2} \to 2p_{3/2}$,
- $3p_{3/2} \to 2s_{1/2}$ and $3d_{3/2} \to 2p_{1/2}$,
- $3d_{3/2} \to 2p_{3/2}$,
- $3d_{5/2} \to 2p_{3/2}$,

i.e., there are 7 possible *labels*, but only 5 possible *energies* (distinct combinations of n, j). These are

- $3, 1/2 \to 2, 1/2$,
- $3, 1/2 \to 2, 3/2$,
- $3, 3/2 \to 2, 1/2$,
- $3, 3/2 \to 2, 3/2$,
- $3, 5/2 \to 2, 3/2$.

Hence there are 5 lines, but as the energy split gets smaller with increasing n they form 2 sets of near degenerate lines $n = 3 \to 2, 1/2$ (2 lines) and $n = 3 \to 2, 3/2$ (3 lines), which is why we see this as a doublet.

---

[20]More precisely, *dipole forbidden.*

19.5 Lamb splitting

In the previous lecture our calculations gave $2p_{1/2}$ and $2s_{1/2}$ as having the same energy — but there is in fact a (very small — of order α^5, i.e., factor 137 smaller than the fine structure) energy difference between $2s_{1/2}$ and $2p_{1/2}$ which comes from the *Lamb shift* (an interaction between the electron and the vacuum fluctuations of the quantised electromagnetic field). It gives a difference between states of the same n, j but different l.

19.6 Hyperfine splitting [Gri 6.5]

There is also *hyperfine splitting*, which gives an energy shift of $E^1_{n,\text{hyperfine}} \propto \alpha^4 m_e/m_p$, i.e., a factor 10 smaller even than the Lamb shift.

This arises because the proton itself has spin, $\underline{S}_p$, giving a magnetic dipole $\underline{\mu}_p = (g_p e/2m_p)\underline{S}_p$, where $g_p \approx 5.59$ is the proton g-factor. In classical electrodynamics, such a dipole sets up a magnetic field

$$\underline{B}_p = \frac{\mu_0}{4\pi r^3}[3(\underline{\mu}_p \cdot \hat{r})\hat{r} - \underline{\mu}_p] + \frac{2\mu_0}{3}\underline{\mu}_p \delta(\underline{r}). \qquad \text{(proton B-field)}$$

The electron also has a magnetic dipole $\underline{\mu}_e = -(g_e e/2m_e)\underline{S}_e$ [recall that $g_e \approx 2$, hence we take $\underline{\mu}_e = -(e/m_e)\underline{S}_e$], and so there is an additional contribution to the potential energy, given by $\hat{H}' = -\underline{\mu}_e . \underline{B}_p$, from the interaction of this with the magnetic field induced by the proton spin.

If we could use our original wavefunctions, we would get $E^1_{nlm} = \langle \psi_{nlm}|H'|\psi_{nlm}\rangle$, and for states with $l = 0$ (and therefore $m_l = 0$) the resulting expectation value arising from the first term in Eq. (proton B-field) goes to zero. In this case $E^1 \propto \underline{S}_e \cdot \underline{S}_p$, and in the same way that we used total angular momentum $\underline{J} = \underline{L} + \underline{S}_e$ we can use total spin $\underline{F} = \underline{S}_e + \underline{S}_p$ to get the energy shift, which is 5.877×10^{-6} eV.[21]

---

[21]This corresponds to a wavelength of 21 cm, which is of enormous astrophysical importance — the reasons being that hydrogen is the most common element in the universe, and that this wavelength falls within the microwave region of the electromagnetic spectrum, which can therefore penetrate large clouds of interstellar dust opaque to visible light.

19.7 Energy levels of other elements

Electrons are fermions – identical, spin-1/2 particles that cannot be in the same state. Thus two electrons in a multi-electron atom cannot have the same quantum numbers. Thus as we progress through elements with increasing atomic number, Z, we are forced to populate energy levels of increasing energy. The periodic table and hence the chemical properties of the elements are largely determined by the occupancy of their outer shell of electrons.

19.7.1 The inert gases

Helium, $Z = 2$, is an inert gas as its two electrons fill the 1s shell ($n = 1$, $l = 0$ and $m_l = -1/2$ or $+1/2$). There is a big jump in energy up the next unoccupied state, 2s ($n = 2$, $l = 0$).

The next inert gas is Neon, $z = 10$. Here the $n = 2$ shell is full. Its electron structure can be denoted,
$$1s^2 2s^2 2p^6.$$

The next inert gas is Argon $Z = 18$ with electron structure
$$1s^2 2s^2 2p^6 3s^2 3p^6$$

which has a full outer p-subshell. (The $3d$ subshell is of significantly higher energy than $3s$ and $3p$ as explained below.)

Then Krypton $Z = 36$ with electron structure
$$1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^6$$

which also has a full outer p-subshell.

19.7.2 The Alkali metals

In contrast the very reactive alkali metals have just one unpaired electron in their outer most shell that is easily removed in chemical reactions.

The electrons structures are:
$$\text{Li} \quad Z = 3 \qquad 1s^2 2s^1,$$
$$\text{Na} \quad Z = 11 \qquad 1s^2 2s^2 2p^6 3s^1,$$
and
$$\text{K} \quad Z = 19 \qquad 1s^2 2s^2 2p^6 3s^2 3p^4 4s^1.$$

Figure 14: Lithium energy levels compared to hydrogen.

We can even attempt to estimate the energy levels of these alkali metals as only the loosely bound outer electron is excited. It doesn't "see" the full charge $+Ze$ of the nucleus as it is screened by the spherical cloud of the inner tightly bound shells of electrons.

If its probability distribution were all at large radii and so did not penetrate into the inner shells then the energy levels would be exactly as in hydrogen
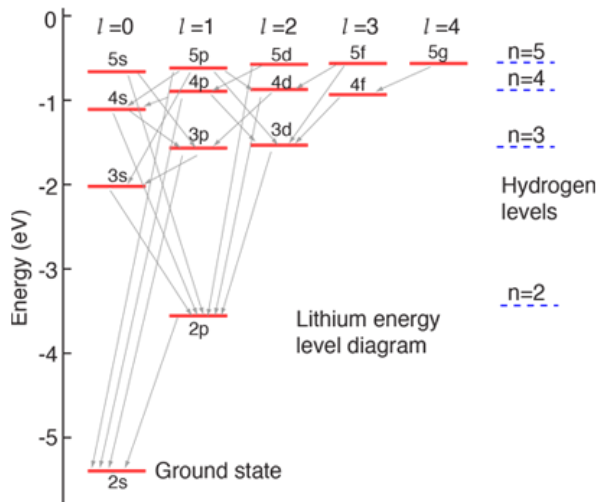$$E = \frac{-13.6}{n^2} \text{ eV}.$$

This is a very good approximation for the states with high l as these are the ones that are most suppressed at small r. At lower l, e.g. $l = 0$ the probability distribution penetrates the core, the electron "sees" more of the nuclear charge and so the state is more tightly bound than in hydrogen. (This is what makes the $4s$ level more tightly bound than $3d$.)

Penetration of 2s and 2p states inside first Bohr radius.
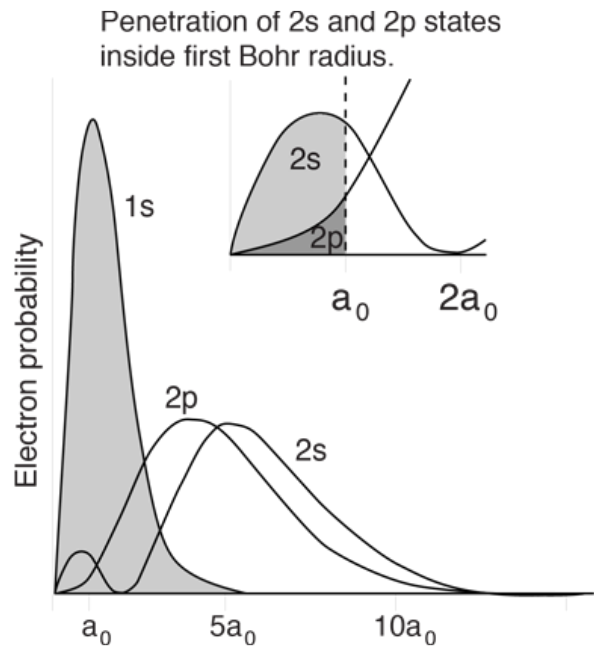
20 Formalities and the Correspondence Principle

20.1 Marginal Probability distributions

Considering a single particle in 1D described by a wavefunction $\psi(x)$, the probability of finding it within dx of x is $|\psi(x)|^2 dx = D(x) dx$ where $D(x)$ is the **probability distribution** in x. The probability we find it within $a < x < b$ is $\int_a^b D(x) dx$ [and $\int_{-\infty}^{\infty} D(x) dx = 1$].

In 3D spherical polars the probability of finding the particle within a volume dV of position (r, θ, ϕ) is $|\psi(r, \theta, \phi)|^2 dV = |\psi(r, \theta, \phi)|^2 r^2 \sin \theta \, dr \, d\theta \, d\phi$. We can then *integrate* (also called marginalizing) to get the distribution function in the coordinate of interest. The radial probability, within dr of r, is $\int_\theta \int_\phi |\psi(r, \theta, \phi)|^2 dV = [\int_\theta \int_\phi |\psi(r, \theta, \phi)|^2 \sin \theta \, d\theta \, d\phi] r^2 \, dr = D(r) dr$.

If we have a wavefunction split into radial wavefunction and spherical harmonic, i.e., $\psi(r, \theta, \phi) = R(r) Y(\theta, \phi)$, we can go a step further. For radial probability, within dr of r is $\int_\theta \int_\phi |\psi(r, \theta, \phi)|^2 dV = \int_\theta \int_\phi |R(r)|^2 |Y(\theta, \phi)|^2 r^2 \sin \theta \, dr \, d\theta \, d\phi = |R(r)|^2 r^2 \, dr = D(r) \, dr$. Similarly, the probability to find the particle within $d\theta$ of θ is $\int_r \int_\phi |R(r)|^2 |Y(\theta, \phi)|^2 r^2 \sin \theta \, dr \, d\theta \, d\phi = \int_\phi |Y(\theta, \phi)|^2 \sin \theta \, d\theta \, d\phi$. As $|Y(\theta, \phi)|^2$ is actually independent of ϕ, this is equal to $2\pi |Y(\theta, \phi)|^2 \sin \theta d\theta = D(\theta) \, d\theta$.

20.2 Time–energy uncertainty principle

One can ask how long it takes a moving wavepacket (in free space) to pass a particular point. In this case $\Delta t = \Delta x / v = m \Delta x / p$, but $E = p^2 / 2m$, and so $\Delta E = 2p \Delta p / 2m$. Hence,

$$\Delta t \Delta E = \frac{m \Delta x}{p} \frac{2p \Delta p}{2m} = \Delta x \Delta p \geq \hbar / 2,$$

and there is a corresponding energy-time uncertainty relation.

Nonetheless, note that position and momentum are both dynamical variables, i.e., measurable characteristics of the system — as is the energy. But time is not. The Δt is *not* the intrinsic dispersion we have if we make a large number measurements of time, it is the time it takes the system to change substantially. Time is not an operator "belonging" to the particle, it is a parameter describing the evolution of the system. "

Nevertheless, there is something to it. A state that only exists for a short time cannot have a definite energy. To have a definite energy, the associated frequency of the state must accurately be defined, and this requires the state to hang around for many cycles — the reciprocal of the required accuracy. Consider spectroscopic linewidths — a long lived transition has narrow energy width, a short one has broad energy width.

What we are really doing is saying that $\Delta t = \Delta \langle B \rangle / (dB/dt)$, where B is some dynamical operator and dB/dt is the rate of change of that dynamical quantity.

20.3 The time evolution of expectation values [Gri 1.5]

For this, we will use Dirac notation. In this context it is a very useful shorthand if we bear in mind that a ket $|\psi\rangle$ can be thought of as a column vector, a bra $\langle \psi| = (|\psi\rangle)^\dagger$ as the Hermitian conjugate of a ket $|\psi\rangle$ (and therefore as a row vector), and Hermitian operators as square, Hermitian matrices.[22]

---

[22]Note that these are vectors in the abstract, linear algebraic sense (any number of dimensions, including infinite dimensions), not in the geometric sense (the 3 spatial dimensions).

Taking the time derivative of the expectation value of a Hermitian operator Q, in general

$$\frac{d\langle Q\rangle}{dt} = \frac{d}{dt}\langle\psi|Q|\psi\rangle = \left(\frac{\partial\langle\psi|}{\partial t}\right)Q|\psi\rangle + \langle\psi|\frac{\partial Q}{\partial t}|\psi\rangle + \langle\psi|Q\left(\frac{\partial|\psi\rangle}{\partial t}\right).$$

However, $H|\psi\rangle = i\hbar\partial|\psi\rangle/\partial t \Rightarrow \partial|\psi\rangle/\partial t = (-iH/\hbar)|\psi\rangle$, and, taking the Hermitian conjugate of both sides, $\partial\langle\psi|/\partial t = \langle\psi|(iH/\hbar)$. Hence,

$$\begin{aligned}
\frac{d\langle Q\rangle}{dt} &= \langle\psi|\frac{i}{\hbar}HQ|\psi\rangle + \langle\psi|\frac{\partial Q}{\partial t}|\psi\rangle + \langle\psi|Q\left(-\frac{i}{\hbar}\right)H|\psi\rangle \\
&= \frac{i}{\hbar}\left(\langle\psi|HQ|\psi\rangle - \langle\psi|QH|\psi\rangle\right) + \langle\psi|\frac{\partial Q}{\partial t}|\psi\rangle \\
&= \frac{i}{\hbar}\langle\psi|(HQ - QH)|\psi\rangle + \langle\psi|\frac{\partial Q}{\partial t}|\psi\rangle \\
&= \frac{i}{\hbar}\langle[H,Q]\rangle + \left\langle\frac{\partial Q}{\partial t}\right\rangle.
\end{aligned}$$

20.4   Formalism: Ehrenfest theorems [Gri 1.5]

If $Q = x$, then

$$\frac{d\langle x\rangle}{dt} = \frac{i}{\hbar}\langle[H,x]\rangle + \left\langle\frac{\partial x}{\partial t}\right\rangle$$

If the Hamiltonian has the form $H = p^2/2m + V(x)$, then $[H,x] = [p^2/2m + V(x), x] = [p^2, x]/2m = -i\hbar p/m$, where we have used $[p^2, x] = ppx - xpp = ppx - pxp + pxp - xpp = p[p,x] + [p,x]p = -2i\hbar p$ as $[p,x] = -i\hbar$. The operator x has no *explicit* dependence on the time t (in the same way a coordinate system used to describe the location of a classical particle typically is not time dependent, unlike the actual location of the particle, which will be time dependent if the particle is moving).

Hence, $\partial x/\partial t = 0$, and we get the Ehrenfest theorem

$$\frac{d\langle x\rangle}{dt} = \frac{i}{\hbar}\left\langle -i\hbar\frac{p}{m}\right\rangle = \frac{\langle p\rangle}{m}. \qquad \text{(Ehrenfest 1)}$$

The other Ehrenfest theorem

$$\frac{d\langle p\rangle}{dt} = -\left\langle\frac{dV}{dx}\right\rangle \qquad \text{(Ehrenfest 2)}$$

can be proven in similar fashion. Combining the two, and making the identification of the negative gradient of the potential with the force, F, we get

$$\frac{d^2\langle x\rangle}{dt^2} = \frac{1}{m}\frac{d\langle p\rangle}{dt} = -\frac{1}{m}\left\langle\frac{dV}{dx}\right\rangle$$

$$\Rightarrow m\frac{d^2\langle x\rangle}{dt^2} = \langle F\rangle,$$

which effectively corresponds to the classical $F = ma$.

20.5   Formalism: Virial theorem

Let us instead let $Q = xp$, which implies

$$\frac{d\langle xp\rangle}{dt} = \frac{i}{\hbar}\langle[H,xp]\rangle + \left\langle\frac{\partial(xp)}{\partial t}\right\rangle$$

It is possible to show that $[H, p] = i\hbar dV/dx$ (See Appendix Appendix: E and note this is also needed to prove Eq. (Ehrenfest 2). Hence, $[H, xp] = [H, x]p + x[H, p] = [(-i\hbar/m)p]p + x[i\hbar dV/dx]$, and

$$\frac{d\langle xp \rangle}{dt} = \frac{i}{\hbar} \left\langle -\frac{i\hbar}{m}p^2 + i\hbar x \frac{dV}{dx} \right\rangle = \left\langle \frac{p^2}{m} - x \frac{dV}{dx} \right\rangle.$$

Hence in equilibrium, such as for a stationary state (energy eigenfunction), expectation values do not change and so $d\langle xp \rangle/dt = 0$ implying

$$\left\langle \frac{p^2}{m} \right\rangle = \left\langle x \frac{dV}{dx} \right\rangle \quad \text{or} \quad \langle T \rangle = \frac{1}{2} \left\langle x \frac{dV}{dx} \right\rangle \qquad \text{(Virial Theorem)}$$

where $T = p^2/(2m)$ is the kinetic energy.

In the case of the harmonic oscillator, where $V = m\omega^2 x^2/2$, then $dV/dx = m\omega^2 x$, and

$$\langle T \rangle = \langle xm\omega^2 x \rangle/2 = \langle m\omega^2 x^2 \rangle/2 = \langle V \rangle, \quad \langle T \rangle = \langle V \rangle.$$

In the case of gravity or electrostatics where $V \propto 1/r$ one can show

$$2\langle T \rangle + \langle V \rangle = 0.$$

In general calculating $\langle T \rangle$ is hard as T is typically portrayed as a second order differential operator, whereas calculating $\langle V \rangle$ (or, more generally, $\langle xdV/dx \rangle$) is easier.[23]

20.6 The Correspondence Principle

The **correspondence principle** (Niels Bohr 1920) states that the behaviour of systems described by the theory of quantum mechanics reproduces classical physics in the limit of large quantum numbers. In other words, it says that for large orbits and for large energies, quantum calculations must agree with classical calculations.

Classical quantities appear in quantum mechanics in the form of expected values of observables, and as such the Ehrenfest theorems (which predict the time evolution of the expected values) is an example of the correspondence principle.

---

[23]The virial theorem in generality refers to properties of very general systems, and has its origins in the work of Rudolf Clausius in the context of classical mechanics, thermodynamics, and statistical mechanics. The term "virial" (for the $\langle xdV/dx \rangle/2$ term) derives from *vis*, the Latin for "force" or "energy."

21 Overview

We can summarise the methodology Quantum Mechanics presented in this course with the following postulates:

21.1 Basic Postulates of Quantum Mechanics

1. For every dynamical system there exists a wavefunction that is a single valued function of the parameters of the system and of time, and from which all possible predictions of the physical properties of the system can be obtained.

2. Every dynamical variable may be represented by a Hermitian Operator whose eigenvalues represent the possible results of carrying out the measurement of the dynamical variable. Immediately after such a measurement the wavefunction of the system will be identical to the eigenfunction corresponding to the eigenvalue obtained as a result of the measurement.

3. The operators representing position and momentum of a particle are $\underline{x}$ and $-i\hbar\underline{\nabla}$ respectively. Operators representing the other dynamical quantities bear the same functional relation to these as do the corresponding classical quantities to the classical position and momentum variables.

4. When a measurement of a dynamical quantity represented by $\hat{Q}$ (a Hermitian operator) is carried out on a system whose wavefunction is ψ then the probability of a result being equal to a particular discrete eigenvalue q_m will be $|a_m|^2$, where $\psi = \sum_n a_n\phi_n$ and the ϕ_n are the eigenfunctions of $\hat{Q}$ corresponding to the eigenvalues q_n. (In the case of continuous eigenvalues the probability of a result between $q(k)$ and $q(k+df)$ is $|a_k|^2$ where $\psi(x) = \int a(k)\phi(k,x)dk$.)

5. The development of the wavefunction with time is governed by the time-dependent Schrödingerequation.

These are sufficient to get a definitive answer to any exam question, but what does it actually mean?

21.2 Interpretation: Probabilism versus determinism [Gri 12]

With all waves, there is intrinsically an uncertainty relationship. Consider playing a note on a musical instrument; it is necessary to play the note for sufficiently long in order to be able to identify its pitch (i.e., its *frequency*), such that a reasonable number of vibration cycles occur over the elapsed time — and to identify the pitch/frequency *perfectly*, in principle the note must be played for an infinitely long time.

It is the same with position x and wavenumber $k = 2\pi/\lambda$ (where λ is the wavelength), and for quantum mechanical particles governed by the Schrödinger wave equation, this carries over into the impossibility of being able to simultaneously exactly specify the position x and the momentum $p = \hbar k$, as expressed by Heisenberg's uncertainty principle. Looking at superpositions of plane waves (waves of definite k, and therefore definite p) it is possible to produce wave solutions where both position and momentum are comparably well-defined around average values, but they are still subject to the same uncertainty principle.

Consider a double slit experiment with massive particles.[24] The particles are clearly wave-like as interference is produced, and this continues to be the case if the rate is reduced such that only particle goes through the slits at any one time.

The wavefunctions give us the probability that the particle ends up at a particular point when it is detected, but we don't know where until it "hits the screen" (or equivalent), as a localised single particle hit. One might ask, just before the electron got to point P "on the screen," where was it? Possible answers could include:
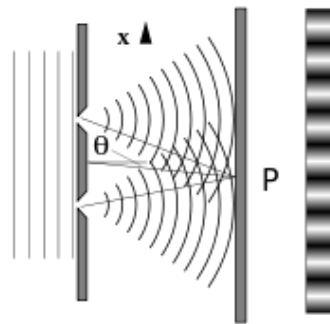

Figure 15: 2-slit interference.

Realist: It was just before P — in which case QM is incomplete because it couldn't tell us this. The position of the particle was never indeterminate, just unknown to the experimenter. ψ is not the whole story — there is some additional information (hidden variables) of which we are ignorant, which is needed to provide a complete description of the particle.

Orthodox (Copenhagen): The particle wasn't really anywhere, it was truly indeterminate. The act of measurement forced the particle to somehow take a stand at some specific position. "Observations not only disturb what is to be measured, they produce it!" (Pascual Jordan). One might then conclude there is something very peculiar about the act of measurement.

Agnostic: Refuse to answer. Effectively, stating that the only way to have known what *would have* been measured as the position of the particle before it was measured would have been *genuinely* to have measured it — and dismissing further speculations as essentially metaphysical.

To this one could add that a literal interpretation of the usual approach in determining the likelihood of measurement outcomes effectively assumes a continuous time evolution governed by the time-dependent Schrödinger equation, followed by a sudden and discontinuous collapse on measurement. As literally instantaneous changes typically have such undesirable requirements as infinite amounts of energy, the question may be whether this is a *genuine* description of the measurement process, or if it is effectively a phenomenological "rule of thumb" (albeit a very accurate and useful one).

21.3 Schrödinger's cat [Gri 12.6]

An interesting way to illustrate these positions is the Schrödinger cat thought experiment. A cat is placed in a box containing a radioactive sample, a Geiger counter connected to a hammer, and a vial of cyanide. If an atom in the sample decays, the Geiger counter will activate the hammer and smash open the vial of cyanide, and the cat dies.

A radioactive sample with a known decay rate is chosen. After a period of time where the sample is equally likely to have decayed as not, we look in the box and check the health of

---

[24]For example electrons, as shown in 1961 by Claus Jönnson [C. Jönnson, Z. Phys. **161**, 454 (1961), Am. J. Phys. **42**, 4 (1974)], or, more recently, quite complex molecules [T. Juffmann, A. Milic, M. Müllneritsch, P. Asenbaum, A. Tsukernik, J. Tüxen, M. Mayor, O. Cheshnovsky, M. Arndt, Nature Nanotechnology **7**, 297 (2012)]

the cat. We will find that the cat is dead or alive. However, it would seem that, according to the orthodox view, the cat is both alive and dead until the box is checked. The system is in a superposition of equally likely states and will remain that way until a measurement is taken, i.e., the system can be represented by the wavefunction

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\text{alive}\rangle + |\text{dead}\rangle). \tag{6}$$

The measurement will also influence future measurements of the system. That is, if you look in the box and see a dead cat then the chances are extremely high that when you look in the box five minutes later the cat will still be dead. This is the "collapse" of the wavefunction. The wavefunction is no longer a good description of the system now that our measurement has forced a state upon it, seemingly putting a great deal of importance on the act of measurement. Without the act of "looking in the box," all systems that can be described by a quantum mechanical interpretation exist in a perpetual limbo of superposition states, according to the orthodox viewpoint. It is the measurement that forces the system to "take a stand," to choose between live cat or dead cat.

The realist position holds that if you get a measurement from a system described by a probability distribution then the system has actually been in that state from the beginning. So, when you look into the box and see a dead cat then the cat was that way prior to your checking the box. The sample decayed at some point prior to looking in the box and the cat has been dead since that time. This belief that the system has been in the state prior to the measurement removes the importance of the act of measurement.

The final stance is the agnostic response. People with this belief avoid the question completely on the basis that it does not matter. The equations work and the philosophical implications are of no consequence. This is perhaps the most practical of approaches, and effectively what experimental physicists actually do in their day-to-day practise.

Schrödinger regarded this all this as patent nonsense! There is something absurd in a macroscopic object like a cat (rather than, say, an electron) being in a linear combination of 2 different states. The easy way around this in the orthodox position is to say that actually it's the triggering of the geiger counter which constitutes the measurement, which gets you out of the apparent necessity for a conscious mind observer. However there still appears to be a superposition state which collapses instantaneously on measurement.

Another caveat is that a cat is a macroscopic thing, not a single atom, and we know that quantum mechanics should give the same answers in the end as classical mechanics when we go from quantum scales to macroscopic scales in order to conform both to our everyday experience and an enormous amount of detailed experimental evidence. There is still a question when "macroscopic" is definitively reached, however one take is that even if one could prepare such a superposition state, it would quickly *decohere* due to multiple interactions between multiple atoms in the cat.[25]

Bell's theorem gave a way to *test* between local realism and orthodoxy, and to date every single test has supported quantum mechanics. To this one should qualify that: *absolute certainty* about a physical theory is not really possible in the same way that in can be in, e.g., pure mathematics; and that it is possible to formulate "realistic" theories the predictions of which are in complete agreement with those of quantum mechanics, but they must be *nonlocal* — effectively *requiring* what Einstein once famously derided as a "spooky action at a distance."

---

[25]See some very readable works by Wojciech Zurek on the subject [W. Zurek, Physics Today **44**, 36 (1991), arXiv:quant-ph/0306072]

This seeming "spooky action at a distance" can be seen, for example, in a possible experiment where for example a state of two photons of opposite direction and spin is formed. The spin of the photons can be measured at some distance from the point of their original formation, and both have an equal chance of having spin $m_s = \pm 1$, however once the spin of one is measured, this directly implies that the other photon *must have* opposite spin (the original state is known as an *entangled state* of two photons), and that this appears to be effectively *instantly known.*[26]

Some related illustrations:

QuVis:Entanglement[27]; QuVis:QMvsHiddenVariables1[28]; QuVis:QMvsHiddenVariables2[29].

21.4 Conclusions

The wavefunction represents the state of a quantum mechanical particle or system. Such particles do not *in general* possess such specific dynamical properties as position, momentum energy etc., until the act of measurement intervenes. The probability of getting a particular value of a specific property is determined by the statistical interpretation of the wavefunction. The wavefunction is taken to collapse on measurement, and so an immediately repeated measurement is certain to yield the same result.

---

[26]See, for example, A. Aspect, et al., Phys. Rev. Lett. **47**, 460 (1981); S. Gröblacher et al., Nature **446**, 871 (2007); B. Hensen et al., Nature **526**, 682 (2015)

[27]`https://www.st-andrews.ac.uk/physics/quvis/simulations_html5/sims/entanglement/entanglement.html`

[28]`https://www.st-andrews.ac.uk/physics/quvis/simulations_html5/sims/quantum-versus-hv1/quantum-versus-hv1.html`

[29]`https://www.st-andrews.ac.uk/physics/quvis/simulations_html5/sims/quantum-versus-hv2/quantum-versus-hv2.html`

Appendix: A   The Uncertainty Principle

Proof that for Hermitian operators A and B with commutator if $[A, B] = i\hbar$ their errors must satisfy the uncertainty relation

$$\boxed{\Delta A \, \Delta B \geq \hbar/2.}$$

Consider

$$
\begin{aligned}
I(\lambda) &= \langle A\Psi + i\lambda B\Psi | A\Psi + i\lambda B\Psi \rangle \\
&= \int (A\Psi + i\lambda B\Psi)^*(A\Psi + i\lambda B\Psi)\, dx \geq 0 \quad \text{if } \lambda \text{ is real}
\end{aligned}
$$

Expanding out

$$
\begin{aligned}
I(\lambda) &= \langle A\Psi | A\Psi \rangle + \langle A\Psi | i\lambda B\Psi \rangle + \langle i\lambda B\Psi | A\Psi \rangle + \langle i\lambda B\Psi | i\lambda B\Psi \rangle \\
&= \langle A\Psi | A\Psi \rangle + i\lambda \langle A\Psi | B\Psi \rangle - i\lambda \langle B\Psi | A\Psi \rangle + \lambda^2 \langle B\Psi | B\Psi \rangle
\end{aligned}
$$

Now as A and B are Hermitian

$$
\begin{aligned}
\langle A\Psi | B\Psi \rangle &= \langle \Psi | AB\Psi \rangle \text{ and} \\
\langle B\Psi | A\Psi \rangle &= \langle \Psi | BA\Psi \rangle
\end{aligned}
$$

Hence

$$\langle A\Psi | B\Psi \rangle - \langle B\Psi | A\Psi \rangle = \langle \Psi | (AB - BA)\Psi \rangle = \langle \Psi | [A, B]\Psi \rangle = \langle \Psi | i\hbar\Psi \rangle = i\hbar \text{ assuming } \Psi \text{ is normalized.}$$

Subbing in

$$I(\lambda) = \langle A\Psi | A\Psi \rangle - \hbar\lambda + \langle B\Psi | B\Psi \rangle \lambda^2 \geq 0$$

For a quadratic $a\lambda^2 + b\lambda + c \geq 0$ for all a λ requires that it has a at most one real root and so $4ac \geq b^2$, i.e.

$$4 \langle A\Psi | A\Psi \rangle \langle B\Psi | B\Psi \rangle \geq \hbar^2$$

Similarly

$$4 \langle C\Psi | C\Psi \rangle \langle D\Psi | D\Psi \rangle \geq \hbar^2$$

where $C = A - \langle A \rangle$ and $D = B - \langle B \rangle$ as

$$[C, D] = [A, B] - [A, \langle B \rangle] - [\langle A \rangle, B] + [\langle A \rangle, \langle B \rangle] = [A, B] = i\hbar$$

Hence

$$4(\Delta A)^2(\Delta B)^2 \geq \hbar^2$$

as

$$(\Delta A)^2 = \int |(A - \langle A \rangle)\Psi|^2 = \langle (A - \langle A \rangle)\Psi | (A - \langle A \rangle)\Psi \rangle = \langle C\Psi | C\Psi \rangle.$$

Thus

$$\boxed{\Delta A \, \Delta B \geq \frac{\hbar}{2} \quad \text{if} \quad [A, B] = i\hbar.}$$

Appendix: B   The Laplacian in various coordinate systems

It is important to realise that the Laplacian operator,

$$\boxed{\nabla^2 = \frac{\partial}{\partial x^2} + \frac{\partial}{\partial y^2} + \frac{\partial^2}{\partial z^2}} \,,$$

has different forms in different coordinate systems. It is useful to go through these derivations once so that the forms aren't a mystery, but you don't need to remember the proofs. The important thing is to remember to look up the boxed equations when needed.

The Laplacian in circular polars

The circular polar and Cartesian coordinate systems are related via

$$
\begin{aligned}
x &= r\cos\theta, & y &= r\sin\theta, \\
r^2 &= x^2 + y^2, & \theta &= \tan^{-1} y/x.
\end{aligned}
\tag{7}
$$

We can therefore use the **chain rule** to re-express

$$\left.\frac{\partial}{\partial x}\right|_y = \left.\frac{\partial r}{\partial x}\right|_y \left.\frac{\partial}{\partial r}\right|_\theta + \left.\frac{\partial \theta}{\partial x}\right|_y \left.\frac{\partial}{\partial \theta}\right|_r$$

then using

$$r^2 = x^2 + y^2 \quad \Rightarrow \quad 2r\left.\frac{\partial r}{\partial x}\right|_y = 2x$$

and

$$\tan\theta = y/x \quad \Rightarrow \quad (1+\tan^2\theta)\left.\frac{\partial \theta}{\partial x}\right|_y = -y/x^2 \quad \Rightarrow \quad \left.\frac{\partial \theta}{\partial x}\right|_y = \frac{-y}{x^2+y^2}$$

we get

$$\left.\frac{\partial}{\partial x}\right|_y = \frac{x}{r}\frac{\partial}{\partial r} - \frac{y}{x^2+y^2}\frac{\partial}{\partial \theta} = \left.\frac{\partial}{\partial x}\right|_y = \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}$$

Similarly

$$\left.\frac{\partial}{\partial y}\right|_x = \left.\frac{\partial r}{\partial y}\right|_x \left.\frac{\partial}{\partial r}\right|_\theta + \left.\frac{\partial \theta}{\partial y}\right|_x \left.\frac{\partial}{\partial \theta}\right|_r = \left.\frac{\partial}{\partial y}\right|_x = \sin\theta\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial \theta}$$

Now we can put these together to form

$$
\begin{aligned}
\nabla^2 &= \left(\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\right)\left(\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\right) \\
&+ \left(\sin\theta\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial \theta}\right)\left(\sin\theta\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial \theta}\right) \\
&= \cos^2\theta\frac{\partial}{\partial r^2} - \frac{2\sin\theta\cos\theta}{r}\frac{\partial^2}{\partial r\partial\theta} + \frac{\sin^2\theta}{r^2}\frac{\partial^2}{\partial\theta^2} + \frac{\cos\theta\sin\theta}{r^2}\frac{\partial}{\partial\theta} + \frac{\sin^2\theta}{r}\frac{\partial}{\partial r} + \frac{\cos\theta\sin\theta}{r^2}\frac{\partial}{\partial\theta} \\
&+ \sin^2\theta\frac{\partial}{\partial r^2} + \frac{2\sin\theta\cos\theta}{r}\frac{\partial^2}{\partial r\partial\theta} + \frac{\cos^2\theta}{r^2}\frac{\partial^2}{\partial\theta^2} - \frac{\cos\theta\sin\theta}{r^2}\frac{\partial}{\partial\theta} + \frac{\cos^2\theta}{r}\frac{\partial}{\partial r} - \frac{\cos\theta\sin\theta}{r^2}\frac{\partial}{\partial\theta} \\
&= \frac{\partial}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}
\end{aligned}
$$

$$\boxed{\nabla^2 = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}}$$

The Laplacian in Cylindrical Polars

In cylindrical polars we have the normal r and θ of circular polars plus a linear z coordinate. Hence

$$\nabla^2 = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2} + \frac{\partial^2}{\partial z^2}$$

The Laplacian in Spherical Polars

Spherical polar coordinates, r, θ, ϕ, cylindrical polar coordinates r_c, θ_c, z and Cartesian coordinates x, y, z are related by

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

The Laplacian in cylindrical polars had above and follows directly from our result for circular polars with just an extra z-derivative term

$$\nabla^2 = \frac{1}{r_c}\frac{\partial}{\partial r_c}\left(r_c\frac{\partial}{\partial r_c}\right) + \frac{1}{r_c^2}\frac{\partial^2}{\partial\theta_c^2} + \frac{\partial^2}{\partial z^2},$$

or expanding out the first term

$$\nabla^2 = \frac{\partial^2}{\partial r_c^2} + \frac{1}{r_c}\frac{\partial}{\partial r_c} + \frac{1}{r_c^2}\frac{\partial^2}{\partial\theta_c^2} + \frac{\partial^2}{\partial z^2}.$$

Now as r_c and z are related to r and θ in exactly the same way as Cartesian and circular polar coordinates, i.e.

$$
\begin{aligned}
r_c &= r\sin\theta \\
z &= r\cos\theta
\end{aligned}
$$

we have that

$$\frac{\partial^2}{\partial r_c^2} + \frac{\partial^2}{\partial z^2} = \frac{\partial}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}.$$

Also (from circular polar result above)

$$\frac{\partial}{\partial r_c} = \sin\theta\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial\theta}.$$

Putting this together along with $\theta_c \equiv \phi$ and $r_c = r\sin\theta$ we have

$$\nabla^2 = \frac{\partial}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2} + \frac{\cos\theta}{r^2\sin\theta}\frac{\partial}{\partial\theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2},$$

which can be rewritten more simply as

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

Appendix: C   The Hermitian Property

The expectation value $\langle Q \rangle$ is determined from its operator $\hat{Q}$ using

$$\langle Q \rangle = \langle \psi | \hat{Q} \psi \rangle.$$

If Q is real then $\langle Q \rangle = \langle Q \rangle^*$ and this implies

$$\langle \psi | \hat{Q} \psi \rangle = \langle \psi | \hat{Q} \psi \rangle^* \quad \text{or equivalently} \quad \int \psi^* \hat{Q} \psi \, dx = \int \psi \hat{Q}^* \psi^* \, dx$$

for any function ψ. If the operator satisfies this condition we say it is **hermitian**.
Below we show that being hermitian also implies

$$\langle f | \hat{Q} g \rangle = \langle g | \hat{Q} f \rangle^* \quad \text{or equivalently} \quad \int f^* \hat{Q} g \, dx = \int g \hat{Q}^* f^* \, dx$$

for any pair of functions f and g. (Note that the first form is analogous to a hermitian matrix $Q_{fg} = Q_{gf}^*$).

Proof:

First writing down the hermitian property $\langle \psi | \hat{Q} \psi \rangle = \langle \psi | \hat{Q} \psi \rangle^*$ for the case of $\psi = f + g$

$$\langle (f + g) | \hat{Q} (f + g) \rangle = \langle (f + g) | \hat{Q} (f + g) \rangle^*$$

which expands to

$$\cancel{\langle f | \hat{Q} f \rangle} + \langle f | \hat{Q} g \rangle + \langle g | \hat{Q} f \rangle + \cancel{\langle g | \hat{Q} g \rangle} = \cancel{\langle f | \hat{Q} f \rangle^*} + \langle f | \hat{Q} g \rangle^* + \langle g | \hat{Q} f \rangle^* + \cancel{\langle g | \hat{Q} g \rangle^*}.$$

Terms cancel as we know the hermitian property applies to f, $\langle f | \hat{Q} f \rangle = \langle f | \hat{Q} f \rangle^*$ and g. We are left with

$$\langle f | \hat{Q} g \rangle + \langle g | \hat{Q} f \rangle = \langle f | \hat{Q} g \rangle^* + \langle g | \hat{Q} f \rangle^*.$$

Then writing down the hermitian property for the case of $\psi = f + ig$

$$\langle (f + ig) | \hat{Q} (f + ig) \rangle = \langle (f + ig) | \hat{Q} (f + ig) \rangle^*$$

which, on remembering being in the $\langle |$ means we take the complex conjugate, expands to

$$\cancel{\langle f | \hat{Q} f \rangle} + i\langle f | \hat{Q} g \rangle - i\langle g | \hat{Q} f \rangle + \cancel{\langle g | \hat{Q} g \rangle} = \cancel{\langle f | \hat{Q} f \rangle^*} - i\langle f | \hat{Q} g \rangle^* + i\langle g | \hat{Q} f \rangle^* + \cancel{\langle g | \hat{Q} g \rangle^*}$$

leaving

$$i\langle f | \hat{Q} g \rangle - i\langle g | \hat{Q} f \rangle = -i\langle f | \hat{Q} g \rangle^* + i\langle g | \hat{Q} f \rangle^*.$$

Dividing through by i and adding to the equation we got from $f + g$ we get

$$2\langle f | \hat{Q} g \rangle = 2\langle g | \hat{Q} f \rangle^*.$$

or

$$\langle f | \hat{Q} g \rangle = \langle g | \hat{Q} f \rangle^*.$$

QED.

Appendix: D   The Radial Wavefunctions of Hydrogen

Starting with the radial equation

$$-\frac{\hbar^2}{2\mu}\frac{d^2U_{nl}}{dr^2} + \frac{l(l+1)\hbar^2}{2\mu r^2}U_{nl} - \frac{Ze^2}{4\pi\epsilon_0 r}U_{nl} = EU_{nl}$$

from section 11.2, where $U_{nl}(r) = rR_{nl}(r)$ and $\psi_{nlm}(r,\theta,\phi) = R_{nl}(r)Y_{lm}(\theta,\phi)$. We can further manipulate this equation to relate it to a standard second order differential equation known as the **associated Laguerre equation**.

First multiply through by $-2\mu/\hbar^2$ and rearrange

$$\frac{d^2U_{nl}}{dr^2} - \frac{l(l+1)}{r^2}U_{nl} + \frac{2\mu Ze^2}{4\pi\epsilon_0 r\hbar^2}U_{nl} = -\frac{2\mu E}{\hbar^2}U_{nl}.$$

For bound states, $E < 0$, so let $k^2 = -2\mu E/\hbar^2$ where k is real.

$$\frac{d^2U_{nl}}{dr^2} = \left(k^2 + \frac{l(l+1)}{r^2} - \frac{2\mu Ze^2}{4\pi\epsilon_0 r\hbar^2}\right)U_{nl}.$$

Dividing by k^2

$$\frac{1}{k^2}\frac{d^2U_{nl}}{dr^2} = \left(1 - \frac{\mu Ze^2}{2\pi\epsilon_0 r\hbar^2 k^2} + \frac{l(l+1)}{r^2 k^2}\right)U_{nl}$$

We can simplify this further by defining $\rho = kr \Rightarrow d\rho/dr = k$ and $d^2U/dr^2 = (d\rho/dr)^2 d^2U/d\rho^2 = k^2 d^2U/d\rho^2$ do that

$$\frac{d^2U_{nl}}{d\rho^2} = \left(1 - \frac{\mu}{\hbar^2}\cdot\frac{Ze^2}{2\pi\epsilon_0 \rho k} + \frac{l(l+1)}{\rho^2}\right)U_{nl} = 0$$

Defining $\rho_0 = \mu Ze^2/(2\pi\hbar^2\epsilon_0 k)$ this becomes

$$\frac{d^2U_{nl}}{d\rho^2} = \left(1 - \frac{\rho_0}{\rho} + \frac{l(l+1)}{\rho^2}\right)U_{nl}.$$

This can be solved by first seeking a solution of the form

$$U_{nl} = \rho^{l+1}\, v(\rho)\, e^{-\rho}$$

where the ρ^{l+1} and $e^{-\rho}$ simplify things as they are the asymptotic solutions when $\rho \to 0$ and $\rho \to \infty$ respectively. Substituting in one finds

$$x\frac{d^2v}{dx^2} + (2(l+1) - x)\frac{dv}{dx} + \left(\frac{\rho_0}{2} - (l+1)\right)v = 0$$

which is of the form of the **associated Laguerre equation**

$$x\frac{d^2v}{dx^2} + (p - x)\frac{dv}{dx} + qv = 0$$

whose solutions, the **associated Laguerre polynomials**[30], can be written as

$$L_q^p = c_0 \sum_{j=0}^{p} \frac{(-1)^j(p+q)!\,x^j}{(p-j)!(q+j)!j!}$$

---

[30]The are also other ways of expressing the associated Laguerre polynomials.

up to a overall normalization constant c_0. By identification this requires $\rho_0/2 = n$, an integer, so that q is an integer. Hence substituting back in we find the radial wavefunctions are

$$R_{nl} \propto \rho^l e^{-\rho} L_{n-l-1}^{2l+1}(2\rho)$$

with the condition that $\rho_0 = 2n$ with $n = 1, 2, 3....$ but this condition means that k is quantised and hence E is quantised too

$$E = -\frac{\hbar^2 k^2}{2\mu}, \quad \text{where } k = \frac{\mu Z e^2}{2\pi\epsilon_0 \hbar^2 \rho_0} = \frac{\mu Z e^2}{4\pi\epsilon_0 \hbar^2 n}$$

$$\Rightarrow E = -\frac{\mu}{2\hbar^2}\left(Z\frac{e^2}{4\pi\epsilon_0}\right)^2 \frac{1}{n^2}$$

$$E = -13.6 \frac{Z^2}{n^2} \frac{\mu}{\mu_H} \text{ eV}$$

This only depends on n, not on l and we can see from the Laguerre polynomials that $n \geq l+1$ so $l \leq n-1$. These are only features of our $1/r$ potential - other central potentials do not do this though *all* have energies not dependent on m.

Appendix: E   The Virial Theorem

First let us show that

$$[H, p] = i\hbar \frac{dV}{dx}.$$

Consider for an arbitrary function $f(x)$ acted on by the commutator

$$
\begin{aligned}
[H, p]f &= \left[\frac{p^2}{2m} + V(x), p\right] f(x) \\
&= \left(\left[\frac{p^2}{2m}, p\right] + [V(x), p]\right) f(x) \\
&= V(x)\left(-i\hbar\frac{d}{dx}\right) f(x) + i\hbar\frac{d}{dx}\left(V(x)f(x)\right) \\
&= -i\hbar V(x)\frac{df}{dx} + i\hbar V(x)\frac{df}{dx} + f(x)i\hbar\frac{dV}{dx} \\
&= f(x)i\hbar\frac{dV}{dx} \quad \text{QED.}
\end{aligned}
$$

It now follows that

$$[H, xp] = [H, x]p + x[H, p] = [(-i\hbar/m)p]p + x[i\hbar dV/dx].$$

We can use this in

$$\frac{d\langle xp\rangle}{dt} = \frac{i}{\hbar}\langle[H, xp]\rangle + \left\langle\frac{\partial(xp)}{\partial t}\right\rangle$$

which is the result from section sec:tevol with $Q = xp$, to get

$$\frac{d\langle xp\rangle}{dt} = \frac{i}{\hbar}\left\langle -\frac{i\hbar}{m}p^2 + i\hbar x\frac{dV}{dx}\right\rangle = \left\langle\frac{p^2}{m} - x\frac{dV}{dx}\right\rangle.$$

Hence in equilibrium $d\langle xp\rangle/dt = 0$ implying

$$\left\langle\frac{p^2}{m}\right\rangle = \left\langle x\frac{dV}{dx}\right\rangle \quad \text{or} \quad \langle T\rangle = \frac{1}{2}\left\langle x\frac{dV}{dx}\right\rangle \qquad \text{(Virial Theorem)}$$

where $T = p^2/(2m)$ is the kinetic energy.

Revision Lecture April 2018

This is just a list of results that I will briefly discuss in the revision lecture. The exam questions won't ask you to reproduce the proofs of these results from the main notes, but understanding the proofs in detail will be helpful as you may need to apply similar techniques to solve parts of the exam questions.

Wave-particle duality and the wavefunction

Wave-particle duality implies particles are described by a wavefunction $\Psi(x,t)$ with

$$P(x,t)dx = |\Psi(x,t)|^2 dx = \Psi^*(x,t)\Psi(x,t)\,dx.$$

This implies wavefunctions must be normalized such that $\int \Psi^*(x,t)\Psi(x,t)dx = \langle\Psi|\Psi\rangle = 1$

Operators and expectation values

All measurable quantities have their associated operator

$$
\begin{aligned}
\hat{x} &= x \\
\hat{p}_x &= -i\hbar\frac{\partial}{\partial x} \\
\hat{H} &= \frac{\hat{p}^2}{2m} + V(x) = \frac{-\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x) \\
\hat{E} &= i\hbar\frac{\partial}{\partial t}.
\end{aligned}
$$

For any such quantity the associated expectation value (ensemble average value) is given by

$$\langle Q\rangle = \int \Psi^*\hat{Q}\Psi\,dx = \langle\Psi|\hat{Q}\Psi\rangle.$$

As $\langle Q\rangle$ is real such operators must be **Hermitian**, i.e.

$$\int \Psi^*\hat{Q}\Psi\,dx = \int \Psi\hat{Q}^*\Psi^*\,dx, \qquad \langle\Psi|\hat{Q}\Psi\rangle = \langle\hat{Q}\Psi|\Psi\rangle = \langle\Psi|\hat{Q}\Psi\rangle^*$$

for any Ψ. Since this is true for any Ψ being Hermitian also implies

$$\int f^*\hat{Q}g\,dx = \int g\,\hat{Q}^*f^*\,dx, \qquad \langle f|\hat{Q}g\rangle = \langle\hat{Q}f|g\rangle = \langle g|\hat{Q}f\rangle^*.$$

The Schrödinger equation

Heuristically the energy conservation equation

$$H = E, \qquad p^2/2m + V = E$$

implies the **time dependent Schrödinger equation**

$$\hat{H}\Psi = \hat{E}\Psi, \qquad \frac{-\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} + V(x)\Psi = i\hbar\frac{\partial\Psi}{\partial t}.$$

If $V(x)$ does not depend on time we can find separable solutions $\Psi(x,t) = \psi(x)e^{-iEt/\hbar}$, where $\psi(x)$ satisfies the **time independent Schrödinger equation**

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

Eigenfunctions and superpositions

This is an eigenfunction-eigenvalue equation. It will have multiple solutions, $\psi_n(x)$ labelled by their energy eigenvalue E_n, i.e.

$$\frac{-\hbar^2}{2m} \frac{d^2\psi_n}{dx^2} + V(x)\psi_n = E_n\psi_n.$$

E.g. for the infinite square well $\qquad$ ($V(x) = 0$ for $0 < x < L$ and $V(x) = \infty$ elsewhere)

$$\psi_n = \sqrt{\frac{2}{L}} \, \sin\left(\frac{n\pi x}{L}\right).$$

Eigenfunctions are orthonormal (orthogonal and normalized)

$$\int \psi_m^* \psi_n \, dx = \langle \psi_m | \psi_n \rangle = \delta_{mn}.$$

Such eigenfunctions form a complete set, meaning any function $\phi(x)$ can be expressed as a linear superposition with weights c_n, i.e.

$$\phi(x) = \sum_n c_n \psi_n(x)$$

where orthonormality implies

$$c_n = \langle \psi_n | \phi \rangle = \int \psi_n^* \phi \, dx.$$

$$\sum_n |c_n|^2 = 1 \qquad \left(\text{proved using } \int \phi^* \phi \, dx = 1\right)$$

and the $|c_n|^2$ can be identified as the probability of being in state ψ_n. E.g. if a system is prepared in state $\phi(x) = \sum_n c_n \psi_n(x)$ and the energy is measured, the probability of measuring E_n is $|c_n|^2$.

Square wells and other potentials

For other potentials, e.g. the finite square or harmonic oscillator $P(x) = \psi^*(x)\psi(x)$ will penetrate into the classically forbidden region.

If $V(x)$ is continuous $\psi(x)$, $d\psi/dx$ and $d^2\psi/dx^2$ must all be continuous. (So that all terms in the Schrödinger equation are continuous). If $V(x)$ has finite discontinuities this still requires $\psi(x)$ and $d\psi/dx$ to be continuous. These properties can be used to sketch wavefunctions, e.g. for the finite square well.

Ladder operators

E.g. for the harmonic oscillator if $H\psi_n = E_n\psi_n$ then $Ha_+\psi_n = (E_n + \hbar\omega)a_+\psi_n$ and similarly $Ha_-\psi_n = (E_n - \hbar\omega)a_-\psi_n$ where the $a_\pm$ are *ladder* operators.

Just need one solution to get started and then we can find all the rest by moving up and down the ladder in energies! I.E. $\psi_{n+1} \propto a_+\psi_n$ and for ground state $a_-\psi_0 = 0$

Also useful for angular momentum states, if $L_z\psi = m\hbar\psi$ then $L_z L_+\psi = (m+1)\hbar L_+\psi$.

Commutators and uncertainty

The commutator of two operators $\hat{A}$ and $\hat{B}$ is $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$.

If $\hat{A}$ and $\hat{B}$ commute $[\hat{A}, \hat{B}] = 0$, then they have a common set of eigenfunctions

$$\hat{A}f_n = a_n f_n \qquad \text{and} \qquad \hat{B}f_n = b_n f_n.$$

This means the values of A and B can both be known simultaneously. If for A you measure a_n this implies the system has wavefunction f_n and so if you measure B you will get b_n.

If $\hat{A}$ and $\hat{B}$ do not commute $[\hat{A}, \hat{B}] \neq 0$, then they do not have a common set of eigenfunctions. We instead have

$$\hat{A}f_n = a_n f_n \qquad \text{and } \hat{B}g_n = b_n g_n,$$

where because the g_n are a complete set we can expand

$$f_n = \sum c_n g_n$$

and the probability that for B we measure b_n is $|c_n|^2$. I.E. knowing A does not determine the value of B (only probabilites) and measuring B (e.g. finding b_m collapses the wavefunction from being the superposition $\sum c_n g_n$ to just g_m and so then the system is no longer in an eigenfunction of $\hat{A}$. If A is measured again its value may change. I.E. non-zero commutators imply an uncertainty principle.

E.g.

$$[x, \hat{p}] = i\hbar \qquad \Rightarrow \qquad \sigma_x \sigma_p \geq \hbar/2.$$

Note in 3D this extends to

$$[x, \hat{p}_x] = [y, \hat{p}_y] = [z, \hat{p}_z] = i\hbar,$$

but along differing directions the operators commute

$$[x, \hat{p}_y] = [y, \hat{p}_z] = [z, \hat{p}_x] = \text{etc} = 0$$

Ehrenfest and virial theorems

We showed that for any operator Q,

$$\frac{d\langle Q \rangle}{dt} = \frac{i}{\hbar}\langle [H, Q] \rangle + \left\langle \frac{\partial Q}{\partial t} \right\rangle.$$

One useful example of this is the **virial theorem**

$$\left\langle \frac{p^2}{m} \right\rangle = \left\langle x\frac{dV}{dx} \right\rangle \quad \text{or} \quad \langle T \rangle = \frac{1}{2}\left\langle x\frac{dV}{dx} \right\rangle$$

This is useful as often the potential V is known and this then gives an easy way of computing the kinetic energy T.

3D wavefunctions, angular momentum and Spherical Harmonics

In 3D, solutions of Schrödinger's equation can be found by seeking solutions that are separable. If the potential is separable in Cartesian coordinates $V(x, y, z) = V_x(x) + V_y(y) + V_z(z)$ then $\psi(x, y, z) = \psi_x(x)\psi_y(y)\psi_z(z)$. Only the infinite square well and harmonic oscillator $V = (k_x x^2 + k_y y^2 + k_z z^2)/2$ separate in this way.

3D wavefunctions have 3 independent quantum numbers. Symmetric potentials will have energy levels that are degenerate, i.e. distinct wavefunctions with different sets of quantum numbers that have the same energy.

If a potential is spherically symmetric $V(r)$, then Schrödinger's equation is separable in spherical polar coordinates, r, θ, ϕ.and has the form

$$\frac{-\hbar^2}{2m}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial \psi}{\partial r}\right) + V(r)\psi + \frac{\hat{L}^2\psi}{2mr^2} = E\psi, \quad \text{where } \psi = R(r)Y(\theta, \phi).$$

Here all the angular dependence is in the total angular momentum operator $\hat{L}^2$. The eigenfunction of $\hat{L}^2$ are the spherical harmonics $Y_{lm}(\theta, \phi)$ which satisfy

$$\hat{L}^2 Y_{lm} = l(l+1)\hbar^2 Y_{lm} \quad \text{and} \quad \hat{L}_z Y_{lm} = m\hbar Y_{lm} \quad \text{with} \quad -l < m < l.$$

We choose to normalise them such that

$$\int_{\text{sphere}} Y_{lm}Y_{lm}^* \, d\Omega = \int_{\theta=-\pi/2}^{\pi/2}\int_{\phi=0}^{2\pi} Y_{lm}Y_{lm}^* \sin\theta \, d\theta d\phi = 1.$$

The angular equation does not depend on the form of $V(r)$ and the solutions have the form

$$\psi(r, \theta, \phi) = R_{nl}(r)Y_{lm}(\theta, \phi).$$

Remembering that the volume element in spherical polars $dV = r^2 \sin\theta d\theta d\phi dr$ the radial probability distribution

$$P(r)dr = r^2 R_{nl}^* R_{nl} \, dr.$$

For the Coulomb potential the energy level depend only on the principal quantum number, n, e.g. for hydrogen

$$E = -\frac{13.6 \, \text{eV}}{n^2}.$$

In decays between atomic electronic states the change in energy is taken away by the emitted photon

$$hf = hc/\lambda = \Delta E.$$

Non-degenerate perturbation theory

If we perturb a system whose unperturbed Hamiltonian, eigenfunctions and energy levels satisfy

$$H_0 \psi_n^0 = E_n^0 \psi_n^0,$$

such that the perturbed Hamiltonian is $H = H_0 + H'$ then corresponding perturbations to the energy levels are

$$E_n^1 = \left\langle \psi_n^0 | H' | \psi_n^0 \right\rangle.$$

Degenerate perturbation theory

If we have two degenerate states that satisfy

$$H_0\,\psi_a^0 = E^0\psi_a^0 \quad \text{and} \quad H_0\,\psi_b^0 = E^0\psi_b^0$$

then the eigenfunctions of the perturbed Hamiltonian $H = H_0 + H'$ are

$$\psi = \alpha\psi_a + \beta\psi_b,$$

where α, β and the perturbation to the energy level, E^1, are given by the solution of

$$\begin{pmatrix} W_{aa} - E^1 & W_{ab} \\ W_{ba} & W_{bb} - E^1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

where $W_{ab} = \langle \psi_a | H' | \psi_b \rangle$ etc.

If $W_{ab} = W_{ba} = 0$, which is true if ψ_a and ψ_b are eigenfunctions of H', then this reduces to the non-degenerate perturbation theory result.

If the original eigenfunctions are eigenfunctions of the perturbation use non-degenerate perturbation theory.

Hydrogen quantum numbers and fine structure

The solutions to the basic Schrödinger equation for hydrogen are highly degenerate.

To fully specify the wave function, ψ_{nlmm_s}, one must define the values of

- n principal quantum number. $(n = 1, 2, 3...)$

- l total orbital angular momentum quantum number $L^2 = l(l+1)\hbar^2$. $(l = 0, ..., n-1)$

- m magnetic quantum number, component of orbital angular momentum along z-axis $m\hbar$. $(m = -l, ..., l)$

- m_s component of spin angular momentum along z axis $m_s\hbar$. $(m_s = \pm 1/2)$

The energy levels only depend on n and so are $2n^2$ fold degenerate.

e.g. $n = 3$ can have $l = 0\ m = 0\ m_s = \pm 1/2$,   2 states
             plus $l = 1\ m = -1, 0, 1\ m_s = \pm 1/2$,   6 states
             plus $l = 2\ m = -2, -1, 0, 1, 2,\ m_s = \pm 1/2$ 10 states,
which equals $18 = 2 \times 3^2$

We can use an alternative decomposition of these degenerate states in which we label the wave function ψ_{nljm_j} by

- n principal quantum number. $(n = 1, 2, 3...)$

- l total orbital angular momentum quantum number $L^2 = l(l+1)\hbar^2$. $(l = 0, ..., n-1)$

- j total angular momentum quantum number $J^2 = j(j+1)\hbar^2$. $(j = 1/2$ for $l = 0$ or otherwise $j = l - 1/2$ or $l + 1/2$ as j must be positive)

- m_j component of the total angular momentum along z axis $m_j\hbar$. $(m_j = -j, ..., j)$

e.g. $n = 3$ can have $l = 0$ $j = 1/2$ $m_j = \pm 1/2$,   2 states

plus $l = 1$ $j = 1/2$, $m = \pm 1/2$ or

$j = 3/2$, $m = -3/2, -1/2, 1/2, 3/2$ $2 + 4$ states

plus $l = 2$ $j = 3/2$, $m = -3/2, -1/2, 1/2, 3/2$ or

$j = 5/2$, $m = -5/2, -3/2, -1/2, 1/2, 3/2, 5/2$ $4+6$ states,

which equals $18 = 2 \times 3^2$

Relativistic corrections to the energy levels can be modelled by treating them as perturbations to the Hamiltonian of the non-relativistic Schrödinger equation. For this we use the second labelling as the perturbations commute with J enabling us to use non-degenerate perturbation theory. These give rise to the **fine structure** of hydrogen and we showed

$$E_{nj} = E_n^0 \left(1 + \frac{\alpha^2}{n^2} \left[\frac{n}{(j + 1/2)} - \frac{3}{4} \right] \right),$$

but I don't expect you to remember this or be able to prove it!