

Lecture 1

Introduction and the essence of QM

Purpose of these lectures is to give you an introduction to perhaps the most important theory of modern physics - quantum mechanics. This theory has revolutionized the way we view the microscopic world and is perhaps the best tested scientific theory ever devised - (certain quantities can be calculated and compared to experiment at the level of a part in a million or better).

That said it is a theory that has been contested since its inception - Einstein always considered it a stop gap answer to a fundamental theory

God does not play dice with the Universe

We will (hopefully) have time to discuss why Einstein was wrong in his view later in this course. Philosophers still debate its true meaning and even practising physicists find it counter to intuition

I think I can safely say that nobody understands quantum mechanics

Richard Feynman

It forms the foundation (with general relativity) of all of twentieth century physics and underpins most of modern chemistry. As one of mankind's greatest intellectual achievements it really forms an important part of our culture - although very few people have any understanding of it - hopefully we can address that in these lectures...

OK then, why did QM come to be ? Essentially it was a response to a series of crises in physics at the turn of the century

1. Blackbody radiation
2. Photoelectric effect
3. Stability of atoms and discrete emission spectra

It appeared that classical physics (Newton and Maxwell) was incompatible with some of the new experimental results following from the discovery of the internal structure of atoms. This was the situation in the early years of this century - it took till 1926 before a satisfactory new framework was developed which could encompass and explain these problems. That framework

was QM. Unlike relativity QM owed its birth to a number of physicists - Schroedinger, Heisenberg, Einstein, de Broglie, Bohr, Born, Ehrenfest, Dirac and others.

Although QM may introduce some rather unfamiliar math its basic postulates are not too long or terribly complicated. They can be written easily on the back of a T-shirt. But they do signal a dramatic departure from Newtonian physics. Consider the motion of a particle subject to some force in one dimension. The goal of classical physics is to calculate how the position coordinate $x(t)$ varies with time. The answer can be gotten by solving the differential equation

$$m \frac{d^2 x}{dt^2} = - \frac{dV}{dx} \quad (1)$$

If we specify eg. the initial position and velocity the resulting motion can be predicted from the solution of this equation. In QM the analogous calculation is phrased very differently. Newton's second law 1 is replaced by Schroedinger's equation

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

The wavefunction $\Psi(x, t)$ plays the role of the coordinate in classical physics - once we know it at some time we can use Schroedinger's eqna 2 to find it for all time. Because $\Psi(x, t)$ is a function of both x and t we must use *partial* derivatives in the equation. Notice that a new fundamental constant has appeared - Planck's constant $\hbar = 1.05 \times 10^{-34} Js$. The smallness of this constant is related to the observation that we do not need to invoke quantum methods until we study the realm of the very small. If it were identically zero classical physics would work even at the smallest of scales (although we wouldn't be here to observe it as the very stability of the atoms in our bodies is the result of quantum effects!) It was first measured by Planck in his work on the light emitted by radiant bodies - the theoretical analysis of this problem led to the need for a quantum theory.

But what is Ψ and what does it tell us about the behavior of the original quantum particle ? The answer is simple and yet puzzling $\Psi^* \Psi \Delta x$ gives the *probability* of finding the particle to be between x and $x + \Delta x$! Notice that (because of the square root of minus one in eqn2) the wavefunction Ψ is a *complex* number - hence the need to multiply it by its *complex conjugate* to obtain a positive definite real probability.

In summary, providing we are happy with partial derivatives, complex numbers and probabilities (!) QM gives us a clear prescription for calculating quantum phenomena. But notice it is (according to our intuitive ideas) an apparently *incomplete* picture. QM tells us only the probabilities of measuring certain values for the position of the particle at a specific point in space and time. And that is the most we can expect to be able to know ! It is a radically different picture than that of classical physics - and one that many people including Einstein were/are unhappy with.

For suppose I ask the question - where was the particle just before I measured its position - Einstein and others would have liked to believe that it had a well-defined position described by *hidden variables* and QM, being incomplete, cannot tell me about it. In effect when one averages out the hidden variables a statistical theory results which can only tell me about probabilities - QM. In contrast the orthodox (Copenhagen) position on this is that the particle does not have a position before it is measured - it is a meaningless concept - there are no hidden variables - this is the quantum nature of matter. Recently, a set of experiments, following on theoretical work of John Bell have served to eliminate the possibility of hidden variables theories and strengthened the orthodox position.

It appears that many of of everyday, intuitive concepts about the world fail to describe the behavior of the quantum world. The only area of QM still in debate is the nature of a measurement on a quantum system and how that is affected by a so-called observer. We will discuss this later.

Some math

Lets introduce/review a few things:

Partial derivatives

Suppose we have a function of more than one variable eg. $f(x, t)$. A partial derivative with respect to x is denoted $\frac{\partial f}{\partial x}$ and means 'differentiate with respect to x holding t constant'. An example, if $f(x, t) = x^2 t^3$ then

$$\frac{\partial f}{\partial x} = 2xt^3$$

Similarly,

$$\frac{\partial f}{\partial t} = 3x^2 t^2$$

Complex numbers

Generalize notion of a real number to accommodate square roots of negative numbers. Define $\sqrt{-1} = i$, then

$$\sqrt{-4} = \sqrt{-1 \times 4} = i\sqrt{4} = 2i$$

Complex number is $z = x + iy$. The real part is called x and the imaginary part is y . Its complex conjugate is $z^* = x - iy$. To add complex numbers just add real and imaginary components separately eg.

$$z = (a + ib) + (c + id) = (a + c) + i(b + d)$$

We can multiply complex numbers as follows

$$z = (a + ib) \times (c + id) = a \times c + i^2 b \times d + i(b \times c + a \times d)$$

Since $i^2 = -1$ we then have

$$z = a \times c - b \times d + i(b \times c + a \times d)$$

Probability

Suppose I were to look at snowfall for Syracuse in the month of January during this century. I could imagine constructing a histogram (bar chart) having as x-axis the number of inches and on the y-axis the number of times that number of inches fell during all 100 januaries on record. It is easy to convert this to a picture of the *probability distribution* for snowfall during a Syracuse january - just divide the numbers on the y-axis by 100. The y-axis now runs between 0 and 1 and measures the probability of a certain number of inches of snow falling. Notice now that the area under the histogram is now unity.

Of course snowfall does not really fall in exact inch amounts - one january there might have been 7.4 inches say. Indeed, when we constructed the original histogram we implicitly rounded snowfalls to their nearest integer. We could improve on this by recording the snowfall in 1/2 inch increments. Then 7.4 would be rounded into the bar corresponding to 7.5 rather than 7.0 inches. Providing we have enough data we could imagine continuing this process counting the number of januaries with snowfall in ever decreasing small intervals. Suppose we do this to an increment of 1/100 inch. You

should see that tops of all the vertical bars start to approach a continuous curve - the true probability distribution that describes the possibility of any possible snowfall amount. Call this curve $P(x)$. Notice that the total area under this curve will still be unity.

Suppose you now want to know the probability of having between 6 and 8 inches - this be read off from the area under the curve from $x=6$ and $x=8$. More generally, suppose you want to know the probability of having a snowfall between $x = 6$ and $x = 6 + \Delta x$ where $\Delta x = 0.1$. This will be approximately $0.1 \times P(6)$. In the limit where $\Delta x \rightarrow 0$ this is exact. That is $P(x) * \Delta x$ is the probability that x lies in the range $x \rightarrow x + \Delta x$ as $\Delta x \rightarrow 0$.

Back to QM

Thus in QM the probability distribution $P(x) = \Psi^*(x, t)\Psi(x, t)$ and tells us the probability of a measurement of the particle's position resulting in a value within the range $x \rightarrow x + \Delta x$. Notice that we now have an additional constraint on the wavefunction $\Psi(x, t)$ - we must have

$$\int \Psi^*(x, t)\Psi(x, t)dx = 1$$

We say that the wavefunction must be *normalized*. It is easy to see that it is always possible to modify any solution to the Schroedinger equation to make this true. It is also possible to show that this feature is preserved in time - that is if we normalize a solution at some initial time and then evolve it in time in accord with the Schroedinger equation eqn.2 it will always be so normalized. This is an important consistency check on the framework.

Lecture 2

History

Lest go back and discuss some of the specific problems which forced this radical departure from classical physics:

BlackBody radiation - 1901

Consider a cavity in an oven at uniform temperature. Once everything has come into equilibrium we can sample the radiation emerging from the cavity. We find the distribution of energy $W(f)$ with frequency f initially rises like f^2 but then turns over and falls to zero with large f . The initial rise is easy to understand - the number of modes of the radiation field between f and $f + \Delta f$ is simply $4\pi f^2 \Delta f$. To understand this remember that a em wave is a *vector* quantity - it has both a magnitude and direction. Thus a given mode is described by a wavelength and a direction in space. When we calculate the number of modes of a given wavelength magnitude (or equivalently frequency) we find a result analagous to the surface area of a sphere - but now a sphere in 'frequency space'. In classical thermodynamics) each such mode carries the same energy - basically $k_B T$ so the net energy at frequency f rises like f^2 . Notice not only does this disagree with the measured distribution - it gives the total radiated energy as infinite !

However, Planck was able to fit the distribution with the function

$$\frac{8\pi h f^3}{c^3} \frac{1}{e^{hf/k_B T} - 1}$$

By drawing on ideas in classical statistical physics and recasting the above expression as an infinite sum over different energy states he was led to a very unusual hypothesis: the radiation energy of a single mode of the em field could only come in units of hf . He had no explanation for this ...

Photoelectric effect - 1905

It was noticed that when UV light was incident on a metal plate electrons are ejected. When the energy of the electrons is measured as a function of light frequency it was found that below a certain threshold frequency there were no electrons and above this the energy of the electrons rose linearly with frequency. The light intensity had no effect on the maximum electron

energy - it only affected the number emitted. This was completely at variance with classical ideas which would have yielded electrons whose energy was intensity dependent (essentially the larger amplitude waves would cause the driven electrons to ‘wobble’ more vigorously which would lead to an increase in their kinetic energy). Einstein explained the effect by extending Planck’s idea to suppose that light consists of photons whose energy varies linearly with frequency. A given electron is kicked out from the metal when it collides with a single photon of sufficient frequency. A photon was imagined to be a packet of wavepacket with a certain particular energy (given by the Planck formula)

Rutherford and Bohr - 1911-13

In 1911 Rutherford performed an historic experiment in which he fired a beam of alpha particles (a type of radiation) at a gold foil. He found that most of the alpha particles suffered only small deflections while just a few were scattered through very large angles. He interpreted the results of this scattering experiment as indicating that the gold atoms consisted of a small, dense core of positive charge surrounded by a much larger and more diffuse cloud of negative charge - the electrons. Unfortunately, this planetary model of the atom was in conflict with classical physics - if the electrons were in a circular orbit they would be accelerating and because of Maxwell’s theory they should radiate light energy. But this loss of energy would lead to a spiraling of the electron into the nucleus - atoms would not be stable. Furthermore, the spectrum of light emitted by such an atom would contain light of all frequencies - which was not observed. In fact the light emitted by heated atoms shows a *discrete* structure characteristic of that particular atom – a so-called line emission spectrum.

Bohr tried to fix this in an ad hoc fashion. Specifically he assumed that for hydrogen only certain states were stable - those in which the angular momentum were a multiple of $\frac{h}{2\pi}$. In those states the electron does not radiate. Furthermore, when an electron moves from one such state to another (lower) state it emits the difference in energy as a photon whose frequency is related to its energy via Planck/Einstein’s relation. Using classical physics it is then easy to see that the allowed radii are

$$r_n = \frac{4\pi\epsilon_0 n^2 \hbar^2}{me^2}$$

Similarly the allowed possible energies can be found

$$E_n = -\frac{1}{2} \frac{me^4}{(4\pi\epsilon_0\hbar)^2 n^2}$$

This explanation accounted well for the experimentally observed line spectrum of hydrogen but what justification could there be for the *quantization* of momenta that Bohr had assumed (or the resulting stability at those momenta) ?

de Broglie – 1925

The situation lay fallow for some years before de Broglie started a new line of reasoning. If light could sometimes behave as a particle (a photon) could not matter behave sometimes as a wave ? The energy relation of Planck could be written as

$$E = \hbar\omega \quad \text{and} \quad p = \frac{E}{c} = \hbar k \quad (1)$$

Perhaps a similar relation governed material particles ? Notice that this identification of wavelength with momenta allowed a possible interpretation of the Bohr quantization condition - via $2\pi r_n = n\lambda$ - the condition for standing waves !

In this case electrons should be able to exhibit phenomena characteristic of waves - such as interference and diffraction ! Such behavior was looked for in a famous experiment of Davisson and Germer (1927) in which a beam of electrons was scattered off a crystal surface in which the interatom separation was comparable to the de Broglie wavelength of the electrons. Lo and behold an interference pattern was observed ! Peaks in intensity were observed whenever the path difference between wave reflected from the first and second atomic layers matched an integral number of wavelengths ($2a \cos \theta = n\lambda$).

We may imagine generalizing this setup to the classic Young's double slit experiment used for light. Electrons are shot at a screen possessing two closely spaced slits. A screen is placed a large distance beyond the slits and is used to record the arrival of electrons which have passed through one of other of the slits. In practice we detect electrons with a detector which flashes when an electron hits it. If we were to do this experiment and record the intensity of electrons recorded by the detector we would find a surprising thing - at certain places on the screen we would *never* see electrons, while at others we would see always a maximum electron intensity! Furthermore these maxima

and minima occur at regular intervals along the screen - we see interference fringes just as we would with light. So the electrons must be associated with a wave as de Broglie had suspected. Clearly the particle character of the electrons emerges as a statistical thing - any individual electron can land anywhere - the wave just gives the probability of finding it at one place or another. It was first thought that the associated wave must somehow describe the aggregate behavior of a bunch of mutually interacting electrons - a given electron will pass through one or other of the slits for certain. But consider the following variation - we can turn down the intensity of the beam until just one electron passes through at a time. If the electron has to go through just one slit then we would predict that the interference pattern would disappear - but it does not we still see an interference pattern ! In some sense the electron passes through both slits ! The associated wave describes the behavior of just a single electron. Equivalently we can say that the electron in passing through the apparatus behaves as a wave but when we come to record it it behaves as a particle ! This is the basis for wave-particle duality. In effect the electron is represented by an abstract state which, depending on what kind of measurement we choose to make, make look alternately particle or wavelike in character !

By 1926, the stage was set - quantum matter should be describable in terms of a wave theory where the momentum of a (free) particle is just $p = h/\lambda$. The interference experiment hints that the intensity of the wave gives the probability for finding the electron. But what is the equation that describes the wave evolution ? Enter Schroedinger ...

Lecture 3.

Arguments for Schroedinger's equation

Free particles

We have seen that free particles are described by the relations

$$E = \hbar\omega \quad (1)$$

$$p = \hbar k \quad (2)$$

It is a reasonable guess that such particles should be associated with the simplest type of wave solution - simple sine or cosine functions.

$$\Psi \sim \sin k(x - vt) \quad \text{or} \quad \cos k(x - vt) \quad (3)$$

Now, $kv = \omega$ the frequency. These solutions describe waves moving in the direction of positive x at speed v . Waves travelling in the opposite direction are obtained by simply switching the sign of v . We will seek a *linear* wave equation that describes the time and space evolution of such waves. This means any linear combination of sine and cosine functions will also be a solution. Specifically we can take the combination $f(kx - \omega t) = \cos(kx - \omega t) + i \sin(kx - \omega t)$. It can be shown that this function $f(kx - \omega t)$ has a very special form

$$f(kx - \omega t) = e^{i(kx - \omega t)} \quad (4)$$

We will take this (complex) exponential function as the free particle solution to the sought-for wave equation.

$$\Psi = A \exp i(kx - \omega t) \quad (5)$$

Additionally, notice the following result

$$\frac{\hbar}{i} \frac{\partial}{\partial x} \Psi = \hbar k \Psi = p \Psi \quad (6)$$

Also,

$$i\hbar \frac{\partial}{\partial t} \Psi = \hbar\omega \Psi = E \Psi \quad (7)$$

Thus, very loosely, the momentum of the particle can be found by differentiating its wavefunction with respect to x and the energy by differentiating

with respect to t . This means in turn that the classical formula for the energy of a free particle $E = p^2/2m$ implies that the wavefunction of such a free particle will satisfy the wave equation

$$\frac{-\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} = i\hbar \frac{\partial \Psi}{\partial t} \quad (8)$$

This is the free particle Schroedinger equation !

Interaction

How can we generalize this to derive a wavefunction for a particle moving in some potential ? Just take the classical energy formula $E = p^2/2m + V$ and do the same replacements !

$$\frac{-\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V\Psi = i\hbar \frac{\partial \Psi}{\partial t} \quad (9)$$

One must be careful. We have not *derived* the Schroedinger equation in the previous pages - rather we have written down the simplest, linear, differential equation that is consistent with conservation of energy and has simple sine or cosine like solutions in the absence of any potential energy. It is an enlightened guess. We must study the consequences of this equation and check them against experiment to be sure of the correctness of this equation. So far this equation has proven fully consistent with all experiments (notice that this equation treats time and space in distinct ways just as did the classical expressions for energy - so it is really only a nonrelativistic equation valid for speeds which are small compared to the speed of light. Dirac was the first person to formulate the analogous relativistic equation which was subsequently named after him).

Averages

We have argued that QM only gives us access to statistical aspects of a particles motion - for example $\Psi^* \Psi \Delta x$ is the probability of finding the particle between x and $x + \Delta x$. How can we find its average position ? Standard probability theory tells us this immediately - multiply the position x by the probability of finding it near x i.e $\Psi^* \Psi \Delta x$ and integrate the result over all positions.

$$\langle x \rangle = \int dx \Psi(x, t)^* x \Psi(x, t) \quad (10)$$

Notice we have assumed that the wavefunction has been normalized.

This is all very well but suppose I want to know not the average position of the particle but say its average momentum. How should I calculate that ? This highlights an aspect of QM that we need to discuss. I have stated that the wavefunction contains all the information that is available about the quantum particle but so far we only know how to calculate information related to its position. In general I should like to be able to calculate the probabilities of measuring specific values for any physical observable, the mean values of those observables etc etc. The general question I want to postpone till later but for now I can give you the correct prescription for computing the average momentum.

Recall that differentiating the wavefunction with respect to position was (up to constant factors) the same as multiplying by the momentum. In general we say that in QM the momentum of a particle is replaced by an **operator** which in this case is just the derivative operator $\frac{\partial}{\partial x}$. Specifically

$$p \xrightarrow{QM} \frac{\hbar}{i} \frac{\partial}{\partial x} \quad (11)$$

Operators are mathematical objects which when applied to functions yield other functions eg. the operator x when applied to the function $f(x) = x^2$ yields another function $f' = xf(x) = x^3$. Similarly the operator $\frac{d}{dx}$ when applied to f yields $f' = \frac{df}{dx} = 2x$. The average value of the momentum in QM is now gotten by sandwiching its associated operator between $\Psi^*(x, t)$ and $\Psi(x, t)$ and integrating over all x .

$$\langle p \rangle = \int dx \Psi(x, t)^* \frac{\hbar}{i} \frac{\partial}{\partial x} \Psi(x, t) \quad (12)$$

This is the same prescription as for the average position if we just replace the (simple!) position operator x by the (more complicated!) momentum operator $p = \frac{\hbar}{i} \frac{\partial}{\partial x}$. We will return to this issue and its generalizations later. At this point you may just consider eqn.12 as another postulate of QM.

What is meant by expectation value ?

In the previous paragraph we introduced a formula for the expectation value of some observable (or operator as it is represented in QM). What is this ? It is **not** the result of measuring that observable for a single particle many times. On the contrary if the first measurement of the particle's position

yields $x = 0.5$ say, then every subsequent measurement of the particle's position will yield $x = 0.5$ (we assume that no other measurements are made in between and the particle is subject to no new forces). Rather, $\langle x \rangle$ measures the average result for measurements of a ensemble of particles all in the same initial state.

Other observables ...

For the simple situation we have discussed so far it is also easy to write down expressions for other mechanical observables such as the kinetic energy or angular momentum eg.

$$\langle T \rangle = \langle p^2 / 2m \rangle = -\frac{\hbar^2}{2m} \int \Psi^* \frac{\partial^2 \Psi}{\partial x^2} dx \quad (13)$$

Lecture 4

Conservation of Probability

We have argued that the wavefunction should be normalized to unity. Physically this corresponds to the simple fact that the probability of finding the particle somewhere should be unity. Whatever happens in the subsequent motion we expect that the total probability to find the particle *somewhere* should still total to unity. It is possible to prove this directly from Schroedinger's equation. Consider

$$\frac{d}{dt} \int \Psi^* \Psi dx = \int \frac{\partial}{\partial t} \Psi^* \Psi dx$$

Using Leibnitz's rule

$$\frac{\partial}{\partial t} \Psi^* \Psi = \Psi^* \frac{\partial \Psi}{\partial t} + \frac{\partial \Psi^*}{\partial t} \Psi$$

Using the Schroedinger equation for the time derivatives allows us to rewrite this as

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

We now notice that this expression can be rewritten

$$\frac{\partial \Psi^* \Psi}{\partial t} = \frac{\partial}{\partial x} \left[\frac{i\hbar}{2m} \left(\Psi^* \frac{\partial \Psi}{\partial x} - \frac{\partial \Psi^*}{\partial x} \Psi \right) \right]$$

We can now integrate this over all x and see since Ψ must go to zero at large x that we obtain the result

$$\frac{d}{dt} \int \Psi^* \Psi dx = 0$$

Thus if it is normalized at $t = 0$ it will remain so - probability is conserved. This is an important and necessary test of the correctness of Schroedinger's approach.

Summary so far

QM tells us that the most information we can obtain about a (microscopic) particle is contained in its wavefunction $\Psi(x, t)$. Once we know the wavefunction at one time Schroedinger's equation allows you to calculate it at

any later time. Physical particles are described by *normalized* wavefunctions $\int \Psi^* \Psi = 1$. In fact QM tells us to interpret $\Psi^* \Psi$ as a *probability density*. This then allows us to write expressions for the expectation value of some observable quantity $Q(x, p)$ as

$$\langle Q \rangle = \int \Psi^* \bar{Q} \Psi dx$$

where the QM operator \bar{Q} is just obtained from its classical expression by replacing x with just x and p by $\frac{\hbar}{i} \frac{\partial}{\partial x}$.

Time independent Schroedinger equation

OK, we have the Schroedinger equation but how do we go about solving it ? It turns out that if the potential V is independent of time this may be accomplished by a method termed *the separation of variables*. What this means is that we seek solutions of the form

$$\Psi(x, t) = \phi(x)f(t)$$

The justification for this is three fold

- More general solutions can be built up from these separable solutions
- They turn out to be states of definite energy
- Expectation values in these states are *independent of time*. They are also termed stationary states.

If we do this we find that the Schroedinger equation reduces to two ordinary differential equations

$$f(t) = \exp -iEt/\hbar \tag{1}$$

$$H\phi(x) = E\phi(x) \tag{2}$$

E is a constant which we will identify shortly as the energy of the state and H is the hamiltonian

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

This second equation involving the Hamiltonian 2 is called the time-independent Schroedinger equation. Notice that, as advertized, the probability density

$\Psi^*\Psi$ is time independent. Furthermore, the operator H is a QM version of the classical energy function for the system and from eqn.2 has an expectation value equal to E . Furthermore, any power of H has expectation value just equal to E raised to that power. Thus the variance of the probability distribution for the energy is zero - the distribution is trivial. Thus any measurement of the energy will return exactly E - stationary states are states of fixed energy.

But what is the energy E - so far we have not specified it. In general we will see that the energy E can take on an infinite number of discrete values dependent on the nature of the potential V . We will call these values E_1, E_2, E_3, \dots and to each allowed value of E there will be an associated solution to the time-independent Schroedinger equation $\phi_1(x), \phi_2(x), \dots$

It is a theorem (we will not attempt to prove it) that the most general solution to Schroedinger's equation is a linear combination of these stationary state solutions (the c_n 's are constants)

$$\Psi(x, t) = \sum_n c_n \phi_n(x) e^{-iE_n t/\hbar} \quad (3)$$

These coefficients c_n can be usually found from a knowledge of the wavefunction at $t = 0$ and the solution of the time-independent problem. The moral of the story is that once we have solved the time-independent equation we have very little left to do to find the most general solution to the time dependent Schroedinger equation! Furthermore, the time-independent equation does not contain i and so we can just look for *real* solutions of this equation.

Examples

The infinite square well

Suppose $V = 0$ for $0 \leq x \leq a$ and is infinite elsewhere. A particle is permanently confined inside this potential well. It could be thought of as a very crude model for a single electron atom. Classically, a particle confined to such a system would just bounce back and forth at constant speed. Its energy could take on any value. We will see that in QM the allowed possible energies are discrete!

First, notice that $\phi = 0$ for $x < 0$ and $x > a$ since there is no probability of finding the particle outside the well. Inside the well, where $V = 0$, the

time-independent equation reduces to

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

Assuming that $E > 0$ we may introduce the variable $k = \sqrt{2mE/\hbar^2}$ and write this equation as

$$\frac{d^2 \phi}{dx^2} = -k^2 \phi$$

The general solution to this is

$$\phi = A \sin kx + B \cos kx$$

The constants A and B are fixed by applying the boundary conditions $\phi(0) = 0$ and $\phi(a) = 0$. This yields $B = 0$ and the *quantization condition*

$$\sin ka = 0$$

The latter means that $ka = n\pi$. Thus not all wavelengths are allowed - only those which correspond to standing waves in the well. We have seen that the Bohr quantization condition for the hydrogen atom could be understood on a similar basis - here we see for the first time that the formal theory of QM is able to explain many of the quantum phenomena which had been observed and which had proven so difficult for classical physics to account for. What remains - well we still have to normalize the solution - that is the origin of the remaining freedom in the constant A ! Thus we find that the stationary states of this potential are of the form

$$\phi_n(x) = \sqrt{\frac{2}{a}} \sin \frac{n\pi}{a} x$$

The energy of this state is $E_n = \frac{n^2 \pi^2 \hbar^2}{2ma^2}$.

We speak of the ground state as the state of lowest energy which here corresponds to $n = 1$ with $E_1 = \frac{\pi^2 \hbar^2}{2ma^2}$. Classically the state of lowest energy corresponds to the particle at rest with $E = 0$. We see in QM that such a state is impossible. The minimum energy the particle can have is E_1 which increases as we confine the particle to smaller and smaller regions ($a \rightarrow 0$). This is a corollary of a very general theorem in QM called Heisenberg's uncertainty principle. This roughly states that the more accurately ones

knows the position of a quantum particle the less certain we are about its momentum (and hence its energy). In fact the product of the uncertainty in its position times the uncertainty in its momentum is *always* greater than some minimum which is equal to $\hbar/2$. We will prove this theorem later in the course but one immediate corollary is that no particle can ever be at rest at a point since then it would have a well-defined position and momentum (zero!). Thus even at zero temperature particles always suffer fluctuations in their positions and momenta - they are somewhat smeared out. This of course is required if they sometimes behave like waves ...

Other points to notice: as n increases the number of zero crossings of the wavefunction increases. By symmetry the expectation value for x is at $x = a/2$ for all states.

Lecture 5

Orthogonality - example for infinite well

Imagine taking the ground state and first excited state, multiplying them and integrating between $x = 0$ and $x = a$.

$$\int_0^a dx \sin(\pi x/a) \sin(2\pi x/a)$$

Using the identity

$$\sin(\pi x/a) \sin(2\pi x/a) = \frac{1}{2} (\cos(\pi x/a) - \cos(2\pi x/a))$$

We can see that the integral is zero - the two wavefunctions are said to be *orthogonal*. Actually, it is not too hard to see that this is true for the product of any two wavefunctions $\phi_n(x)$ and $\phi_m(x)$ - their product integrates to zero. This allows us to determine the coefficients c_n easily knowing the initial wavefunction. Taking the general expansion result $\Psi = \sum c_n \phi_n$, multiplying both sides by $\phi_m(x)$ and integrating (at $t = 0$) leads to the result

$$c_m = \int \Psi(x, 0) \phi_m(x) dx$$

Thus a particle starting in a stationary state will always remain in that state (all the c_n 's are zero bar one) and all average properties of that particle will be time independent.

It will turn out to be a general feature of solutions to the time-independent Schroedinger equation for arbitrary potential - different stationary states (with differing energies) will be orthogonal (in this sense) to each other. Furthermore, any state of the particle can be expanded out on this set of special stationary states and the expansion coefficients determined using the orthogonality condition.

Time dependence

Suppose we manage to create an initial wavefunction that is of the form

$$\Psi(x, 0) = \sqrt{\frac{1}{a}} (\sin(\pi x/a) + \sin(2\pi x/a))$$

Thus we find that $c_1 = c_2 = \sqrt{\frac{1}{a}}$ and all other c 's are zero. The energy of this state is no longer fixed but can fluctuate - it is *not* a stationary state. We can now compute the time dependence of the probability density of the particle. This yields

$$\begin{aligned} a\Psi^*\Psi &= \sin^2(\pi x/a) + \sin^2(2\pi x/a) \\ &+ 2\sin(\pi x/a)\sin(2\pi x/a)\cos(E_1 - E_2)t/\hbar \end{aligned}$$

In obtaining this we have used the result

$$2\cos\theta = e^{i\theta} + e^{-i\theta}$$

Thus the probability density (and hence the mean position of the particle) oscillates with time. Let's calculate $\langle x(t) \rangle$

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

Thus

$$a\langle x(t) \rangle = \int x\sin^2(\pi x/a)dx + \int x\sin^2(2\pi x/a)dx \quad (1)$$

$$+ \int 2x\sin(\pi x/a)\sin(2\pi x/a)\cos(\Delta Et/\hbar)dx \quad (2)$$

where $\Delta E = E_2 - E_1$ and using the result

$$\int x\sin^2(px) = \frac{x^2}{4} - x\frac{\sin(2px)}{4p} - \frac{\cos(2px)}{8p^2}$$

$$\int_0^a x\sin(\pi x/a)\sin(2\pi x/a)dx = -\frac{8a}{9\pi^2}$$

we can show that

$$\langle x(t) \rangle = \frac{a}{2} \left(1 - \frac{32}{9\pi^2} \cos(\Delta Et/\hbar) \right)$$

The angular frequency of this motion is thus

$$\omega_{\text{quantum}} = \frac{3}{2} \frac{\pi^2 \hbar}{ma^2}$$

Thus the mean position of the particle oscillates with time in a way reminiscent of the classical oscillations of a particle in the same well. In fact if we

calculate the mean energy of this state $\bar{E} = (E_1 + E_2)/2$ and use it to derive a speed v we find

$$v_{\text{classical}} = \frac{5\pi^2\hbar^2}{4ma^2}$$

the frequency of such a motion is then $\omega_{\text{classical}} = \frac{2\pi v}{2a}$ and yields

$$\omega_{\text{classical}} = \sqrt{\frac{5}{2}} \frac{\pi^2\hbar}{ma^2}$$

Thus the quantum oscillation frequency is pretty close to the classical oscillation frequency of a particle whose energy is close to the mean quantum energy.

$$\frac{\omega_{\text{classical}}^2}{\omega_{\text{quantum}}^2} = \frac{10}{9}$$

Thus certain linear combinations of stationary states can produce probability 'lumps' whose averaged motion resembles the classical motion of a particle. We will see another example of this when we return to the discussion of free particles. Another example of such a connection to classical physics is given the *Bohr Correspondence Principle* which states that in the limit of large quantum numbers n the results of quantum physics agree with the classical motion. Actually the correctness of this principle relies on the fact that the separation between energy levels becomes small for large n - which is not in fact true for potentials which go to infinity for large distance such as the infinite square well but it is true for the potentials within atoms and molecules - the situation which Bohr was trying to describe.

Ehrenfest's theorem

The connection to classical physics can be made more explicit in a celebrated theorem due to Paul Ehrenfest which shows that quantum expectation values evolve according to Newton's 2nd law. Consider the time rate of change of the expectation value of the momentum $\langle p \rangle$

$$\frac{d\langle p \rangle}{dt} = \int dx \frac{\partial}{\partial t} \left(\Psi^* \frac{\hbar}{i} \frac{\partial}{\partial x} \Psi \right)$$

Using Leibnitz we find

$$\frac{d\langle p \rangle}{dt} = \frac{\hbar}{i} \int dx \left(\frac{\partial \Psi^*}{\partial t} \frac{\partial \Psi}{\partial x} + \Psi^* \frac{\partial^2 \Psi}{\partial x \partial t} \right)$$

Substituting in from the Schroedinger equation and using integration by parts we can see

$$\frac{d \langle p \rangle}{dt} = - \int dx \Psi^* \frac{\partial V}{\partial x} \Psi$$

The right hand side is nothing but the expectation value of the force on the particle and we have the result. Similarly it is easy to see that

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

Wave packets and free particles

We have seen that the Schroedinger equation in free space admits simple plane wave solutions $e^{ikx - \omega t}$ where $\omega = \hbar k^2 / 2m$. However, although we initially hypothesised that these represented free particles it is easy to see that this cannot be true.

- They are not *normalizable*
- If we compute their wave velocity $v_{\text{quantum}} = \frac{\hbar k}{2m}$ it is *half* the speed of the classical particle they are supposed to represent! ($v_{\text{classical}} = p/m = \hbar k/m$)
- Such states do not satisfy the Uncertainty Principle - they have a definite momentum.

These problems are all evaded once we realise that physical particles are necessarily restricted to a finite region of space. They should hence be represented not by infinitely long plane wave states but superpositions of such states whose wavelengths/momenta are restricted to lie in a small band. Such a solution is termed a wavepacket - $\Psi(x, t)_P$.

$$\Psi(x, t)_P = \int dk c(k) e^{ikx - \omega t}$$

If we assume that $c(k)$ is strongly peaked about some wavenumber k_0 we can expand $\omega(k)$ around k_0

$$\omega(k) \sim \omega(k_0) + (k - k_0) \left. \frac{d\omega}{dk} \right|_{k_0}$$

Changing variables to $s = k - k_0$ we can write this as

$$\Psi_P \sim e^{-\omega_0 t} \int ds c(k_0 + s) e^{i(k_0 + s)x - \omega'_0 s t}$$

This can be rewritten

$$e^{i(-\omega_0 t + k_0 \omega'_0 t)} \int ds c(k_0 + s) e^{i(k_0 + s)(x - \omega'_0 t)}$$

Thus the wavepacket at time t is just of the same form as at $t = 0$ but translated to $x - \omega'_0 t$. Thus the velocity of the packet as a whole (the so-called group velocity) is just ω'_0 which is none other than the classical velocity! Furthermore, by choosing a suitably rapidly decaying function $c(s)$ we can make the resulting wavepacket go to zero sufficiently fast as $x \rightarrow \infty$ as to render it normalizable. Finally, such wavepackets are compatible with the Uncertainty Principle - if you compute the typical spread in momenta k it is possible to show that it varies inversely with the typical region in position over which the wavepacket is non zero - the coefficient of proportionality being at least $\hbar/2$ (the latter being realised by a gaussian function $c(k)$).

In summary (free) physical particles do not possess a definite energy or momentum but possess a spread compatible the uncertainty principle given their spatial localisation. Mathematically they are represented by a wavepacket which is just a linear superposition of plane wave states. Such a wavepacket propagates approximately as a single probability lump with a velocity equal to the classical velocity of the free particle of that energy.

Lecture 6

Solving the Schroedinger equation on the computer

Types of solutions

We have seen that finding solutions to the full Schroedinger equation reduces to solving the time independent Schroedinger equation

$$H\Psi = E\Psi$$

where H is the hamiltonian (or energy) operator. The solutions to this equation are of two types

- Bound state solutions
- Scattering solutions

The former are exemplified by solutions of the infinite square well potential - we see a discrete spectrum of *normalizable* allowed states. A particle in such a state has zero chance of being found at infinity - it is essentially confined to a certain region of the x-axis. The reason for this is simply that its energy E is less than the potential at large distance $E < V(\infty)$ - even quantum effects cannot allow it to tunnel over infinite distances. The scattering solutions on the other hand allow the particle to escape to large distances and are characterized by $E \geq V(\infty)$. One example of such a state is the free particle. We have seen that such states have a continuous range of possible energies E and can be represented by non-normalizable wavefunctions - like e^{ikx} . To achieve physical scattering wavefunctions we must then superpose many of these *mathematical* solutions to Schroedinger's equation leading to the creation of wavepackets. Such wavepackets contain a set of momenta and energies centred around some central (classical) value and hence satisfy the uncertainty principle. They are then normalizable.

For the moment we will concentrate on the bound states (for potentials like the infinite square well these are the only possibility). For most realistic potentials we will find that it is impossible to solve the (time independent) Schroedinger equation to find the allowed states and energies and so we must turn to numerical methods to find approximate solutions.

Discretization

The simplest method to solve the one-dimensional problem for bound states is called the *shooting method*. First, we replace derivatives with *finite differences*

$$\frac{d\phi}{dx} \rightarrow \frac{\phi_{n+1} - \phi_n}{\Delta x}$$

Here, ϕ_n represents the value of $\phi(x)$ at $x = n\Delta x$ and Δx is a (small) interval on the x-axis. The idea is that for Δx small enough the approximation for the derivative will be accurate enough and the final solution ϕ_n will approximate closely the solution to the continuum equations. To proceed we first write the Schroedinger equation as two first order equations

$$\begin{aligned}\frac{d\phi}{dx} &= p \\ \frac{dp}{dx} &= \frac{2m}{\hbar^2} (V - E) \phi\end{aligned}$$

Next we apply our simple *discretization* recipe and rearrange the two equations into the simple form

$$\begin{aligned}\phi_{n+1} &= \phi_n + \Delta x p_n \\ p_{n+1} &= p_n + \Delta x \frac{2m}{\hbar^2} (V_n - E) \phi_n\end{aligned}$$

This way of writing the Schroedinger equation makes it clear what we have to do - if we specify an initial value (say at $x = 0$) for $\phi(0)$ and $\frac{d\phi(0)}{dx}$ we can use these equations to predict their values at $x = \Delta x$. Having their values at $x = \Delta x$ we can use the equations again to get their values at $x = 2\Delta x$ etc etc. In this way we can generate the wavefunction for all x . What value should we choose at $x = 0$? If the potential is symmetric about $x = 0$ i.e $V(x) = V(-x)$ it is easy to see that the final wavefunction obeys one of two conditions $\phi(x) = \phi(-x)$ the so-called *even parity* solutions or $\phi(x) = -\phi(-x)$ the *odd parity* solutions. The former are even functions and hence have gradient $p = 0$ at $x = 0$. The latter have $\phi(0) = 0$. Furthermore, we can adjust the scale of $\phi(x)$ arbitrarily at this stage so that a suitably general set of initial conditions are

- $\phi(0) = 1, p(0) = 0$ (even parity)
- $\phi(0) = 0, p(0) = 1$ (odd parity)

Allowed energies - shooting and bisection

OK, so now we have initial conditions and a recipe eqn. 1 to generate the wavefunctions. At this point you may be wondering - what determines the allowed values of E ? We will see that only for rather special values of the energy E will iteration of these equations yield a normalizable wavefunction - in general the numerical solution will yield wavefunctions which diverge very rapidly at large x . Thus the shooting technique consists of guessing a value for E , iterating the equations determining ϕ_n and seeing whether in some region far from the center of the potential the wavefunction is 'small' - if not we use some criterion for improving on our guess for E . If we do this carefully, gradually increasing the energy E , we can find all the allowed energies and associated wavefunctions.

One good way to locate the energies precisely is called *bisection*. By experimenting with the local Java applet you will notice that if you have found two values of E say E_1 and E_2 which straddle an allowed energy they will diverge with opposite signs at large x (say some fixed point $x = x_L$). Thus if you overshoot the allowed energy the wavefunction gets large in say a positive direction ($\phi(x_L) > 0$) while undershooting yields a wavefunction which diverges in a negative direction ($\phi(x_L) < 0$). Now consider the mean energy $(E_1 + E_2)/2$. If its wavefunction at $x = x_L$ has the same sign as say the wavefunction computed at E_1 you can use this midpoint energy as a new guess for E_1 . If not it must have the same sign as that corresponding to E_2 and you may use the midpoint energy as the new E_2 . In either case you have now halved the region in which the true allowed energy is located. By iterating this procedure many times you can locate the energy to arbitrary precision.

General Strategy

- Decide a region of E in which you want to search for allowed energies. Decide also the minimum possible energy.
- Guesstimate a typical energy level separation (say using dimensional analysis)
- Decide on a value for x_L - the potential will contain a length scale - use some multiple of this.

- Set a lower energy E_1 to the minimum possible energy. Set the upper energy E_2 to the same. Compute $\phi(x_L)$ for this energy.
- Scan upward in the energy E_2 (using the average energy level separation as a guide) until you find $\phi(x_L)$ changing sign - an allowed energy lies now between E_1 and E_2 .
- Bisect to find exact energy.
- Reset lowest energy E_1 to (just above) this allowed energy
- Repeat last three steps until you have exhausted the initial energy region of interest.

The C code you will use employs this strategy. To change the potential you are examining you just edit the function `potential()`

Harmonic Oscillator

We will first use this technique to find the allowed energies and wavefunctions of the harmonic oscillator potential $V = \frac{1}{2}m\omega^2x^2$. First we simplify the discrete equations by working in terms of rescaled energy $\epsilon = \frac{2m}{\hbar^2}E$. Also, the potential is now of the form $v(x) = \frac{2m}{\hbar^2}\frac{m\omega^2x^2}{2} = \alpha^2x^2$. Initially in the code we set $\alpha = 1/2$. First we set `PARITY=0` (even parity) and compile the code. When we run it the energies are (approx) 0.50, 2.50, 4.50, 6.50, 8.50. Similarly when we set `PARITY=1` (odd parity) we find the energies 1.50, 3.50, 5.50, 7.50, 9.50. Thus we see that the energies ϵ of the harmonic oscillator are equally spaced (and for $\alpha = 1/2$) separated by unity. The ground state has non-zero energy as required by the uncertainty principle. The plotted wavefunctions oscillate and then decay rapidly to zero at large x . As the energy increases the number of oscillations increases just as for the infinite square well. Furthermore, we can show that the product of any two such states integrated over the interval is zero as required by orthogonality!

If we vary α we will find that the energy level splitting varies as 2α . In fact we can demonstrate numerically that

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

We will see later that we can derive this result analytically. However, by repeating this calculation for the anharmonic oscillator $V = \alpha x^4$ we can use

these numerical methods to solve a system which is not tractable by analytic methods. Similarly we can use these techniques to study the bound states of the finite square well - for a deep well the low lying states will look like those of the infinite well - although the wavefunction will now extend outside the range of the well - decaying exponentially out to large distance. Also, a finite well possesses only a finite number of bound states - for large energies the particle can move off to infinite distance with finite probability - it exhibits a continuum of so-called 'scattering' states. These are not accessible with the current code.

Lecture 7.

Vectors, Operators and the Hamiltonian evolution

We now turn to a more abstract discussion of QM. We have seen that in wave mechanics a quantum system is described mathematically by a function of space and time called the *wavefunction*. This wavefunction forms a convenient representation for the more abstract notion of quantum state. Remember that the description of a quantum state is quite different from its classical counterpart - it must be consistent with the Heisenberg uncertainty principle and the principle of superposition. Thus a quantum state describing a single particle may admit the possibility that a measurement of the particle's position or momentum may return more than one value. To explain double slit interference we must also assume it may be split and recombined to yield other new quantum states.

Thus the mathematical quantity which is used to represent this quantum state should contain a large number of different *component* pieces of information and two such objects may be added together with different weights to produce another such quantum state object. If we look around in mathematics for objects which behave in this way we will see that quantum states have the same properties as *vectors*. Unlike the familiar vectors of three dimensional space the quantum state vectors inhabit a so-called *complex vector space* whose dimension may be infinite! To see how this all works out lets summarize the important properties of ordinary vectors - once this is done it will be obvious which of these properties carries over to more general vector spaces. These properties are as follows

1. The sum of two vectors \mathbf{a} and \mathbf{b} is a vector $\mathbf{c} = \mathbf{a} + \mathbf{b}$.
2. We can multiply any vector by a *scalar* to yield another vector eg. $\mathbf{b} = \lambda\mathbf{a}$.
3. Vectors can be expanded in components; that is it is possible to choose a suitable *basis* set of vectors (like $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ of Cartesian coordinates) and express every vector as a sum over basis vectors weighted by (real number) coefficients.

$$\mathbf{a} = a_i\mathbf{i} + a_j\mathbf{j} + a_k\mathbf{k}$$

The set of 3 numbers (a_i, a_j, a_k) are then the components of the vector.

4. There exists a dot product between two vectors \mathbf{a} and \mathbf{b} denoted $\mathbf{a}\cdot\mathbf{b}$ which is just a scalar. If \mathbf{a} and \mathbf{b} are expanded on a Cartesian basis this is just

$$\mathbf{a}\cdot\mathbf{b} = a_i b_i + a_j b_j + a_k b_k$$

The components of a vector are then nothing else than the dot product of the vector with the (unit) basis vectors eg. $a_i = \mathbf{i}\cdot\mathbf{a}$.

The generalization that is needed to discuss vectors for QM is

- Introduce a new notation $|a\rangle$ for quantum state vector.
- Replace real scalars by complex scalars. (vector components are also in general complex numbers)
- Let the dimension of the space in which QM takes place be as big as you like. Thus the expansion of a state vector in components can be written

$$|a\rangle = \sum_i a_i |e_i\rangle$$

where the set of vectors $\{|e_i\rangle\}$ forms a basis in the space.

- There exists a dot product which combines two vectors to yield a *complex* number. This is denoted

$$\langle a|b\rangle$$

- The analog of a Cartesian basis is one in which the basis vectors are *orthogonal* - $\langle e_i|e_j\rangle = \delta_{ij}$.
- The components a_i can just be interpreted as the dot products of the basis vectors with the vector eg. $a_i = \langle e_i|a\rangle$.
- The dot product between two vectors $|a\rangle$ and $|b\rangle$ in such a basis is now

$$\mathbf{a}\cdot\mathbf{b} = a_i^* b_i + a_j^* b_j + a_k^* b_k$$

The latter result can be understood if we allow ourselves two types of vector - the original (sometimes called *ket*-vector) $|a\rangle$ and a dual *bra* vector $\langle a|$ whose components a_i^{bra} with respect to the dual bra basis $\langle e_i|$ are just the complex conjugates of the ket components $a_i^{\text{bra}} = a_i^*$. The vector $\langle a|$ dual to $|a\rangle$ is sometimes also called the *adjoint* vector. Notice that this

definition of the dot product ensures that the dotproduct of a vector with itself is a positive real number (this will be a necessary ingredient in order to allow for the probability interpretation of the theory - it will allow us to normalize the quantum state vector to be of unit length $\langle a|a \rangle = 1$ completely analogously to the normalization of the Schroedinger wavefunction).

With this technology we can expand the state vector $|\Psi \rangle$ on some n -dimensional basis $\{|e_i \rangle\}$ - we will assume from now on the basis is *orthonormal*;

$$|\Psi \rangle = \sum_{i=1}^n \psi_i |e_i \rangle$$

The coefficients ψ_i are just the generalized dot products of the state vector with the basis vectors $\phi_i = \langle e_i | \Psi \rangle$. Let us postpone discussion of what appropriate set of basis vectors to choose. We will see that this is intimately connected to the choice of observable to measure. Just notice that this mathematical expansion of the state vector embodies the physical principle of superposition - that is the quantum probability wave can be the sum of many contributions each of which may correspond to classically distinct possibilities (for example each basis vector might represent a possible location of the particle on the x-axis - the quantum state is a sum over these - allowing the particle to be simultaneously at many positions and forcing a probabilistic interpretation of the theory).

The evolution of the state vector follows from the Schroedinger equation of wave mechanics:

$$i\hbar \frac{\partial |\Psi \rangle}{\partial t} = H |\Psi \rangle \tag{1}$$

The quantity H is called the *Hamiltonian operator*. Its purpose is to transform one vector into another (neighboring) vector. It has the dimensions of energy - a primary observable.

Normalization of the wavefunction translates into the statement $\langle \Psi | \Psi \rangle = 1$ - i.e the state vector is a kind of unit vector. This normalization condition was necessary for a probabilistic interpretation of the theory - and it is necessary that it remain true for all time - thus as the state vector evolves in time it remains always of unit 'length'. What properties must H have in order that this be true ?

Imagine solving the equation 1 over a small time period Δt .

$$|\Psi(t + \Delta t) \rangle = |\Psi \rangle - \frac{i\Delta t}{\hbar} H |\Psi(t) \rangle$$

To check the normalization condition we need to introduce the concept of *adjoint* operator. Suppose $|\phi\rangle = A|\psi\rangle$ then the adjoint vector $\langle\phi|$ is given by

$$\langle\phi| = \langle\psi|B^\dagger$$

This is often written as

$$\langle B^\dagger\psi|$$

which emphasises that B^\dagger operates on the bra vector - here $\langle\psi|$. Thus we find that

$$\langle\Psi(t+\Delta t)|\Psi(t+\Delta t)\rangle = \langle\Psi(t)|\Psi(t)\rangle + \frac{i\Delta t}{\hbar} (\langle\Psi(t)|(H^\dagger - H)\Psi(t)\rangle)$$

Thus we require the Hamiltonian to be a *self-adjoint* or *Hermitian* operator $H = H^\dagger$. We will see that hermitian operators play a central role in quantum mechanics.

Notice that our argument implies that the operator $(1 - \frac{i\Delta t}{\hbar}H)$ applied to *any* vector preserves its length to $O(\Delta t^2)$. In the limit of vanishing Δt it is an example of a *unitary operator* U . Such operators U have the property $U^\dagger = U^{-1}$ where U^{-1} is the inverse operator - the operator which undoes the effect of U . To see this is norm preserving consider

$$|a\rangle = U|b\rangle$$

$$\langle aU^\dagger U|a\rangle = \langle a|a\rangle$$

In general, such an operator may be written as $\exp iH$ where the operator H is hermitian (not necessarily the Hamiltonian). This is the case for the Schroedinger evolution eqn. 1 which has a formal solution

$$|\Psi(t)\rangle = \exp -iHt/\hbar|\Psi(0)\rangle$$

Notice that general unitary operators take one orthonormal frame into another - they correspond to a change of basis. For example if we have some orthonormal basis $\{|e_i\rangle\}$ then for any unitary operator S we may construct another basis $\{|e'_i\rangle\}$ which is also orthonormal.

$$|e'_i\rangle = S|e_i\rangle$$

then

$$\langle e'_j|e'_i\rangle = \langle e_j|S^\dagger S|e_i\rangle = \delta_{ij}$$

The length of any vector $|\Psi\rangle$ is then invariant under such a change of basis (although its components will change). Unitary changes of basis are just analogous to using a rotated frame of reference in a discussion of two or three dimensional vectors.

Eigenvalues and Eigenvectors

Consider again vectors in ordinary three dimensional space. And consider the rotation operator - most vectors will change under rotation - the exception are vectors which lie along the axis of rotation. They don't change at all. Also, vectors lying in the plane at 90 degrees to the axis just flip sign if the angle of rotation is 180 degrees. In a complex vector space such as used by QM *every* linear transformation has 'special' vectors such as these - they are called *eigenvectors*. They transform into multiples of themselves under the operator/transformation. The multiplying constant is called the *eigenvalue*.

$$T|\alpha\rangle = \lambda|\alpha\rangle$$

Hermitian operators are special because

- They have real eigenvalues
- Their eigenvectors are orthogonal (and can be made orthonormal)
- They span the space - that is any vector can be expanded as a linear combination of the eigenvectors - they can hence be used as a basis set.

The last statement is strictly only always true for finite dimensional vector spaces. The first of these is easily proved:

$$\langle \alpha | T | \alpha \rangle = \lambda \langle \alpha | \alpha \rangle$$

because of hermiticity we may rewrite this as

$$\langle T \alpha | \alpha \rangle = \lambda^* \langle \alpha | \alpha \rangle$$

(we have used the fact that the Hermitian adjoint of a scalar is just its complex conjugate) Hence $\lambda = \lambda^*$ QED. To prove the second statement suppose

$$T|\alpha\rangle = \lambda|\alpha\rangle$$

and

$$T|\beta\rangle = \mu|\beta\rangle$$

Thus

$$\langle \alpha | T\beta \rangle = \mu \langle \alpha | \beta \rangle$$

Using hermiticity we see that this can be rewritten

$$\langle T\alpha | \beta \rangle = \mu \langle \alpha | \beta \rangle$$

But the LHS is just $\langle \lambda\alpha | \beta \rangle$ and since λ is real and $\mu \neq \lambda$ we see

$$\langle \alpha | \beta \rangle = 0$$

Generalized Statistical Interpretation

Suppose now that we have a quantum system described by a state vector $|\Psi\rangle$ evolving according to eqn. 1. It is a postulate of QM that every observable will be represented in the theory by an hermitian operator. Suppose we choose to make a measurement of some physical observable corresponding to an hermitian operator Q . The possible results of that measurement are the eigenvalues measuring q_i is simply

$$| \langle \phi_i^Q | \Psi \rangle |^2$$

After the measurement the state ‘collapses’ to the state $|\phi_i^Q\rangle$ and will then continue to evolve according to eqn. 1 once more.

Lecture 8.

The Uncertainty Principle

As we have described, for any observable A we will associate an hermitian operator A . The expectation value of the operator in the state $|\Psi\rangle$ is then naturally given by the expression

$$\langle A \rangle = \langle \Psi | A \Psi \rangle$$

Consider two such observables A and B . The product of the (squared) uncertainty in A will be

$$\sigma_A^2 = \langle (A - \langle A \rangle) \Psi | (A - \langle A \rangle) \Psi \rangle = \langle f | f \rangle$$

Similarly for the observable B (with f replaced by g) Therefore

$$\sigma_A^2 \sigma_B^2 = \langle f | f \rangle \langle g | g \rangle \geq | \langle g | f \rangle |^2$$

This is called the Schwarz inequality. It is a theorem for all vector spaces equipped with a notion of dot product. Now the RHS of this expression is always bigger than the square of the imaginary part of $\langle f | g \rangle$ so we can also write this as

$$\sigma_A^2 \sigma_B^2 \geq \left(\frac{1}{2i} [\langle f | g \rangle - \langle g | f \rangle] \right)^2$$

But

$$\langle f | g \rangle = \langle \Psi | (A - \langle A \rangle) (B - \langle B \rangle) | \Psi \rangle$$

This is simply

$$\langle AB \rangle - \langle A \rangle \langle B \rangle$$

Similarly

$$\langle g | f \rangle = \langle BA \rangle - \langle B \rangle \langle A \rangle$$

So we find

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

where the square brackets are defined by

$$[A, B] = AB - BA$$

This is the Uncertainty Principle in its most general form. It says that for any pair of observables whose quantum operators do not commute there will be an associated uncertainty relation for the product of the fluctuations in their expectation values. This will be true for any quantum state vector $|\Psi\rangle$. If we choose $A = x$ and $B = \frac{\hbar}{i} \frac{d}{dx}$ we find

$$[x, p] = i\hbar$$

and hence

$$\sigma_x \sigma_p \geq \frac{\hbar}{2}$$

This is the famous Heisenberg uncertainty principle which we have encountered several times already.

Operators which do not commute do not share a complete set of common eigenvectors - thus if we make a measurement of one of them we will (by the generalized statistical interpretation described above) collapse the state vector to a particular eigenvector of that observable, which will not then be an eigenvector of the other observable. Indeed, if we choose to expand this collapsed state vector on the basis of eigenvectors of the second observable we will generate a range of values - i.e a measurement of that second observable would be *uncertain*. The magnitude of that uncertainty would be given by the uncertainty relation.

Finite dim representations - matrix mechanics

Let us now derive a concrete realization of these ideas by thinking of a finite dimensional system endowed with some orthonormal basis set $\{|e_i\rangle\}$, $i = 1, \dots, n$. An operator T takes every vector into some other vector. Thus acting on a particular basis vector $|e_1\rangle$ we would find

$$T|e_1\rangle = T_{11}|e_1\rangle + T_{21}|e_2\rangle + \dots + T_{n1}|e_n\rangle$$

and similarly for all the other basis vectors. Thus the effect of the transformation can be encoded in n^2 coefficients T_{ij} where we can see that

$$T_{ij} = \langle e_i | T | e_j \rangle$$

Furthermore if $|\alpha\rangle$ is an arbitrary vector

$$|\alpha\rangle = a_1|e_1\rangle + a_2|e_2\rangle + \dots + a_n|e_n\rangle$$

then

$$T|\alpha\rangle = \sum_j a_j (T|e_j\rangle) = \sum_i \left(\sum_j T_{ij} a_j \right) |e_i\rangle$$

Evidently T takes a vector with components a_1, \dots, a_n into a vector with components $a'_i = T_{ij} a_j$. It is easy to see that a compound operator $C = ST$ is just represented by a set of n^2 numbers C_{ij} given by

$$C_{ij} = \sum_k S_{ik} T_{kj}$$

Thus if we *represent* the state vector by its components in a particular basis, then operators may be represented by *matrices*. Furthermore, if we think of the components of a ket vector $|a\rangle$ as just a column matrix, then the dot product $\langle b|a\rangle$ is just the matrix product

$$b^\dagger a$$

where the dagger operation transposes the column vector $|b\rangle$ into the row vector $\langle b|$ and takes its complex conjugate.

Now consider the scalar

$$\langle \phi|A|\psi\rangle = \sum_{i,j} \phi_i^* \psi_j \langle e_i|A|e_j\rangle \quad (1)$$

$$= \sum_{i,j} \phi_i^* A_{ij} \psi_j \quad (2)$$

If A is Hermitian we must have that the LHS equal

$$\langle H\phi|\psi\rangle$$

This will only be true when

$$A_{ij} = A_{ji}^*$$

This then is the definition of a Hermitian matrix. It is a result in matrix theory that the eigenvalues of a hermitian matrix are purely real and that the eigenvectors are orthonormal and span the original space - just as for the abstract operators. Unitary matrices are then simply those matrices whose (matrix) inverse is just equal to its (matrix) hermitian conjugate as you would expect.

Thus any observable can be represented by either an operator or a (possibly infinite dimensional) hermitian matrix. The possible (real) values which

can result from measurement of that observable are just the eigenvalues of that matrix and the probability of measuring any such value is just the (mod) square of the component of its state vector along the eigenvector corresponding to that eigenvalue. Thus most of day to day business of doing a QM calculation, written in this language, corresponds to finding the eigenvalues and eigenvectors of some Hermitian matrix A . For small systems this can be done by setting the determinant of $A - \lambda I$ equal to zero. This generates a polynomial equation with n roots - the eigenvalues $\lambda_i, i = 0 \dots n$.

Notice that a matrix M will take on a diagonal form in a basis of its own eigenvectors

$$M_{ij} = \langle e_i | M | e_j \rangle = \lambda_j \langle e_i | e_j \rangle = \lambda_j \delta_{ij}$$

The transformation to this basis from the original basis must be a unitary matrix transformation (since it must preserve the length of the state vector). Thus

$$|e_i^D \rangle = S_{ij} |e_j \rangle$$

The form of the matrix eigenvalue equation may then be preserved if the matrix M_{ij} undergoes a so-called similarity transformation.

$$A^{\text{diagonal}} = SAS^\dagger$$

The only remaining question is how is the unitary matrix S determined? The unitary matrix can be built from the eigenvectors of the matrix by assembling them into successive columns of S . Thus if we know the eigenvectors of an given matrix (in a given basis) we can construct the unitary matrix that effects the change of basis which renders the matrix diagonal.

It is also clear then that if two operators have a common set of eigenvectors they can be *simultaneously* diagonalized (put in diagonal form) by a common unitary transformation. But diagonal matrices commute - hence there will be no uncertainty relation holding between the two corresponding observables. So such observables are often said to be *compatible*. The opposite reasoning is also true - incompatible operators are those which do not commute and have a non-trivial mutual uncertainty relation. They cannot be simultaneously diagonalized and have different eigenvectors.

Numerical methods for solving matrix mechanics problems often turn this around - they typically focus on finding iteratively a transformation S which is capable of rendering A *diagonal*. The diagonal elements are then the eigenvalues and the matrix which effects the diagonalization yields the eigenvectors.

The Schroedinger equation just becomes a matrix equation now. This can be seen by taking the vector form and expanding $|\Psi\rangle$ on a (time-independent) set of basis states $\{|e_i\rangle\}$

$$|\Psi\rangle = \sum_i \langle e_i|\Psi\rangle |e_i\rangle$$

Inserting this into the Schroedinger equation we find

$$i\hbar \sum_i \frac{d \langle e_i|\Psi\rangle}{dt} = \sum_i H|e_i\rangle \langle e_i|\Psi\rangle$$

Take dot product of this with $\langle e_j|$. Orthogonality of the basis says

$$i\hbar \frac{d \langle e_j|\Psi\rangle}{dt} = \sum_i \langle e_j|H|e_i\rangle \langle e_i|\Psi\rangle$$

or equivalently ($\langle e_i|\Psi\rangle = c_i$ etc)

$$i\hbar \frac{dc_i}{dt} = H_{ij}c_j$$

where H_{ij} is the Hamiltonian matrix. If we look for solutions of this equation in which all the c_i vary with time in the simple way $c_i \sim e^{-iEt/\hbar}$ (the stationary solutions) we find that the *allowed energies E are just the eigenvalues of the Hamiltonian matrix and the eigenvectors the allowed stationary states*

Lecture 9.

Chemical Bonding revisited

We can use this matrix formulation of QM to return to the problem of understanding the QM origin of the chemical bond. Suppose we are looking at say the H_2^+ ion which consists of just one electron shared between two protons. If these two protons are a long way apart we can envisage the electron as being attached to one or the other - giving two physical states $|1\rangle$ and $|2\rangle$. These will be our base states. This truncation of the space of all states to just two will be sufficient to discuss simple properties of the ground state of the system. Imagine taking the two protons infinitely far apart - then we expect no overlap $\langle 1|H|2\rangle = 0$ - the (local) Hamiltonian cannot cause an electron to jump from one state to another. In this limit we can trivially solve the Schroedinger equation for the two state system - we find $|1\rangle \sim \exp(-iH_{11}t/\hbar)$. Thus $H_{11} = E_0$ is to be interpreted as the energy an electron would have in a single hydrogen ion. By symmetry it is equal to H_{22} - the energy if the electron lived on the other proton. As we put the two protons closer and closer together we find a non-zero chance for the electron to jump from one proton to another - this is represented by a nonzero matrix element between the two states - $\langle 1|H|2\rangle \neq 0 = -A$ say. Thus the two state system we must solve takes the form

$$i\hbar \begin{pmatrix} \frac{dc_1}{dt} \\ \frac{dc_2}{dt} \end{pmatrix} = \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

where $c_1 = \langle 1|\Psi\rangle$ etc. Lets look for solutions where both amplitudes c_1, c_2 have the same time dependence $\exp -iEt/\hbar$. Thus we need to solve the matrix eigenvalue problem $Hc = Ec$. The possible eigenvalues turn out to be

$$E_{+/-} = E_0 + / - A$$

with corresponding eigenvectors $|e_{+/-}\rangle = \frac{1}{\sqrt{2}}(|1\rangle + / - |2\rangle)$. Thus the ground state is lowered in energy with respect to a single hydrogen atom and its amplitude is an even function of the two separate amplitudes. In order to get the true ground state we must use a state in which the electron is equally split between the two protons! This energy will *decrease* as the distance between the two protons is decreased (i.e we assume that the magnitude of A increases as the distance decreases) - hence chemical bonding.

Notice also, that if we start out with the electron located on atom $|1\rangle$ it will not remain so - the presence of the two time dependent factors for the eigenstates, which oscillate at different frequencies, ensures that at some later time it will contain an admixture of the second state - i.e there will be some probability that the electron would be found on the second atom!

Suppose we had two distinct atoms trying to share an electron in this picture. It is easy to modify the calculation to handle the situation when $H_{11} \neq H_{22}$. For small A the new energies are

$$E_I = H_{11} + \frac{A^2}{H_{11} - H_{22}}$$

$$E_{II} = H_{22} - \frac{A^2}{H_{11} - H_{22}}$$

This is typically much smaller than the splitting for equal energies and explains why single electron bonding in non-symmetric molecules is not very common.

A more typical situation is the two electron chemical bond in which two electrons are shared by the two atoms. The hydrogen molecule furnishes a nice example. The situation can still be modelled approximately as a two state system corresponding to the physical situation in which electron **a** is around the first proton and electron **b** is around the second and vice versa. As before the base states for these two situations have the same energy by symmetry but as before there is a possibility of hopping or exchange of the two electrons between the two protons. The mathematics is identical to the single electron problem and so we expect that the allowed energies of the combined system are split by this quantum mechanical hopping and we predict that the energy of the ground state is *lowered* as a result. This decrease in ground state energy is accentuated for small inter-proton separations and results in a chemical bond - the covalent bond.

If we admit two more base states into the picture - those corresponding to the two electrons being on one or the other of the two atoms (if they are dissimilar) we can allow for *ionic bonding* in this picture. Thus by extensions of these ideas we can start to understand the quantum mechanical basis for chemistry.

Emergence of wave mechanics

Consider a one dimensional crystal with lattice sites $i = 1 \dots N$ spaced a apart. A quantum particle can move on this lattice and we can adopt a set of basis states to describe this motion as follows $|e_i\rangle$ corresponds to the particle being located at site i .

$$|\Psi\rangle = \sum_i \langle e_i | \Psi \rangle |e_i\rangle$$

Writing down the Schroedinger equation for this system we find

$$i\hbar \frac{dc_i}{dt} = \sum_j H_{ij} c_j$$

where $\psi_i = \langle e_i | \Psi \rangle$ is the amplitude for the particle to be located at site i (the component of the state vector on the basis vector $|e_i\rangle$). The matrix elements H_{ij} are just

$$H_{ij} = \langle e_i | H | e_j \rangle$$

and measure the component of $H|e_j\rangle$ on the basis vector $|e_i\rangle$. Remember that the hamiltonian measures the change in a vector under an infinitesimal amount of time. We expect that for very short times the only state a particle can hop to from lattice site i are its neighbours $i - 1$ and $i + 1$. Furthermore we expect that the probability for going in either direction is the same. Thus we might guess

$$H|e_i\rangle = (V + 2A)|e_i\rangle - A|e_{i-1}\rangle - A|e_{i+1}\rangle$$

where A and V are some constants. Furthermore, we expect that for a free particle the vector corresponding to a uniform probability distribution for the particle is time independent. This identifies V as a simple function of the potential energy of the particle. We thus have found

$$H_{ij} = -Aa^2\Delta_{ij} + V\delta_{ij}$$

where Δ_{ij} is a discrete form of the operator $\frac{d^2}{dx^2}$ restricted to the crystal lattice. Furthermore, we can rewrite this matrix equation in the suggestive form

$$\sum_j a \left(-Aa\Delta_{ij} + V\frac{\delta_{ij}}{a} \right) \psi_j = E\psi_i$$

In the limit in which $a \rightarrow 0$ with $N \rightarrow \infty$ the sums become integrals and $\delta_{ij}/a \rightarrow \delta(i - j)$ where i and j can now be thought of as continuous valued positions. Thus this equation will be the same as the fundamental equation of wave mechanics if we require

$$Aa = \frac{\hbar^2}{2m}$$

In this limit can clearly adopt a normalization for the eigenvectors in which

$$\sum_i a |\psi_i|^2 = 1$$

which just goes over into the usual form

$$\int dx \psi^*(x) \psi(x) = 1$$

when it is realized that $\psi_i = \langle e_i | \Psi \rangle = \psi(x_i)$ is just the probability amplitude for finding the particle at lattice site x_i . Notice that the eigenvectors then take the form of Dirac delta functions $\psi^i(x_j) \rightarrow \frac{\delta_{ij}}{\sqrt{a}}$.

Thus we can see that wave mechanics is just one representation of QM - in which we focus on measurements of position and expand all vectors on a basis corresponding to the eigenvectors of the position operator. On a finite lattice this is a finite dimensional matrix problem and as $a \rightarrow 0$ it goes over to a infinite dimensional matrix problem. Differential operators - like $\frac{d^2}{dx^2}$ being just convenient representations of infinite dimensional matrices !

Indeed you can see that the space of all functions satisfies all the requirements of a vector space - the sum of two functions is a function. One can define operators and scalars (just complex numbers) and most importantly one can define the notion of a dot product where summation over components is replaced by integration over a continuous valued index - the position x . The most delicate remaining issue consists of *completeness* - that any function can be expanded out as a sum over a set of basis functions. The boundary conditions play a crucial role here - if the wavefunction is subject to suitable boundary conditions this postulate may also be satisfied - eg. Fourier series and the infinite square well. Such functions are also square integrable - that is their dot product is finite always. The latter is necessary for a probability interpretation of the theory. Functions satisfying these requirements are said to live in a *Hilbert space* after the famous mathematician.

Band structure

Suppose we use this lattice Schroedinger equation to model the situation in a crystal in which an additional constant potential energy is acquired when an electron is at a lattice site - $V_i = V_0$ say. It is easy to see that e^{ikna} is a solution of the time independent Schroedinger equation with energy

$$E = V_0 - \frac{\hbar^2}{m} (1 - \cos ka)$$

for any k from π/a to zero (for an infinitely long crystal). Thus the possible energies form a *band* with energies ranging from V_0 to $V_0 - \frac{2\hbar^2}{m}$. Clearly V_0 is the energy an electron would have on a isolated lattice site - its original atomic energy level - this is split into a band of allowed energies in the crystal. Furthermore, a similar process will be true for all the original atomic levels - leading to a sequence of bands of allowed energies separated by finite energy gaps. This structure is a crucial component to the arguments which are used to explain all metallic, insulator and semiconductor behavior. Electrons fill these stationary states up to some maximum energy - the so-called Fermi energy (there is a principle called the Pauli exclusion principle that prevents electrons from being in the same state). A filled band is inert for conduction purposes - it has an equal number of left and right traveling electron waves and no free states available to take more electrons such as those responsible for carrying an electric current. A partially filled band however has available electron states for carrying current and such a substance is a metal. An insulator has a filled band and hence cannot use electrons to carry heat or electric current. A semiconductor has a filled band but a small bandgap to the next (empty) band - hence at room temperature a small number of electrons can become thermally excited to this new band and carry current. Furthermore, the *holes* left behind in the nearly filled band can also carry current - this substance will be a semiconductor.

Lecture 10.

Angular Momentum

Let us now turn to QM in more than one dimension. In one dimension we know how to represent the position and momentum operators (when referred to the basis of eigenvectors of the position operator). In three dimensions we will immediately need to know how to represent the *angular momentum* of a particle. This is a vector quantity in classical mechanics with components L_x, L_y, L_z where

$$\begin{aligned}L_x &= yp_z - zp_y \\L_y &= zp_x - xp_z \\L_z &= xp_y - yp_x\end{aligned}$$

In QM we expect it to be represented by a set of three Hermitian operators corresponding to these components. By plugging in the 1D correspondence $p_x = \frac{\hbar}{i} \frac{\partial}{\partial x}$ etc we find the operators

$$\begin{aligned}L_x &= \frac{\hbar}{i} \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \\L_y &= \frac{\hbar}{i} \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) \\L_z &= \frac{\hbar}{i} \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right)\end{aligned}$$

The first question we should ask is: what are the eigenvalues and eigenfunctions of these operators? We know that the former are the possible values of a measurement of the corresponding component of angular momentum L and the latter are important for extracting the probabilities of measuring any one such eigenvalue. We can also ask the question - what is the maximal amount of information we can have concerning the angular momentum of a particle in QM - that is, can we know all these components precisely and *simultaneously*? This latter question can be answered by computing the commutator of these operators with each other. We know if this is zero the operators are simultaneously diagonal in the basis of their common eigenvectors. Otherwise there will be a generalized uncertainty principle governing the minimum size of the fluctuations in these two operators. We can easily

show that

$$\begin{aligned} [L_x, L_y] &= i\hbar L_z \\ [L_z, L_x] &= i\hbar L_y \\ [L_y, L_z] &= i\hbar L_x \end{aligned}$$

From these fundamental commutation relations the entire theory of angular momentum can be deduced. Evidently, the different components are incompatible observables

$$\sigma_{L_x} \sigma_{L_y} \geq \frac{\hbar}{2} | \langle L_z \rangle |$$

On the other hand the square of the total angular momentum $L^2 = L_x^2 + L_y^2 + L_z^2$ does commute with any of these components eg.

$$[L^2, L_z] = 0$$

Thus we can choose eigenstates which are simultaneously eigenstates of both L^2 and say L_z .

$$L^2 f = \lambda \hbar^2 f$$

and

$$L_z f = \mu \hbar f$$

Define the two operators

$$L_{\pm} = L_x \pm iL_y$$

The commutators with L_z are

$$[L_z, L_{\pm}] = [L_z, L_x] \pm i[L_z, L_y] = \pm i\hbar L_{\pm}$$

and, of course, L_{\pm} commutes with L^2 . Imagine some function f which is an eigenfunction of L^2 and L_z then we can show that $L_{\pm}f$ is also an eigenfunction of these two operators with eigenvalues $\lambda\hbar^2$ and $(\mu \pm 1)\hbar$. L_+ is termed a raising operator and L_- a lowering operator. Thus, for a given value of λ the application of these two operators generates a ladder of states, each one separated from its neighbors by a L_z eigenvalue different by \hbar . We can keep applying say L_+ and move up the ladder. Eventually, however we must find a state with the highest possible value of the z-component of angular momentum. For such a state $\mu = l$ say and we have

$$L_+ f_l = 0$$

Now, we can show that

$$L_{\pm}L_{\mp} = L_x^2 + L_y^2 \mp i(L_xL_y - L_yL_x)$$

or

$$L^2 = L_{\pm}L_{\mp} + L_z^2 \mp \hbar L_z$$

Thus

$$L^2 f_l = (\hbar^2 l^2 + \hbar^2 l) f_l$$

Thus $\lambda = l(l + 1)$. In the same way there is a bottom value of $\mu = \bar{l}\hbar$ such that

$$L_- f_{\bar{l}} = 0$$

we find similarly that $\bar{l}(\bar{l} - 1) = l(l + 1)$. Thus $\bar{l} = -l$. Thus the eigenvalues of L_z are $m\hbar$ where m runs from $-l$ to $+l$ in $2l + 1$ integer steps. Thus l must be integer *or half-integer*. Notice that the maximum component of the angular momentum l can never equal its total value $l(l + 1)$ since this would invalidate the generalized uncertainty principle (since then we would precisely the values of L_x and L_y - zero!). The computation of the eigenfunctions takes a little more effort!

Eigenvectors

We employ *spherical coordinates*

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

We find

$$\begin{aligned} L_+ &= \hbar e^{i\phi} \left(i \cot \theta \frac{\partial}{\partial \phi} + \frac{\partial}{\partial \theta} \right) \\ L_- &= \hbar e^{-i\phi} \left(i \cot \theta \frac{\partial}{\partial \phi} - \frac{\partial}{\partial \theta} \right) \end{aligned}$$

Notice that they are *independent* of the coordinate r - thus so will be the eigenfunctions. If we can solve the equation $L_+ f_{l,l} = 0$ for the highest L_z eigenstate with L^2 eigenvalue $l(l + 1)$ we can always get the other $2l$ states by

applying L_- . Furthermore, let's assume that the eigenstates can be written as a product state

$$f_{l,l} = F(\theta)G(\phi)$$

We find that $G(\phi) = e^{im\phi}$ where m must be an integer in order for G to be single valued. Furthermore, since $L_z = -i\hbar \frac{\partial}{\partial \phi}$ we see $m = l$. The remaining equation for θ reads

$$\frac{dF}{d\theta} = l \cot \theta F$$

Writing $l \cot \theta = \frac{d}{d\theta} \ln \sin^l \theta$

$$\frac{d \ln F}{d\theta} = \frac{d}{d\theta} \ln \sin^l \theta$$

Thus

$$F(\theta) = A \sin^l \theta$$

and the total eigenfunction looks like

$$f_{l,l} e^{il\phi} \sin^l \theta$$

By applying L_- we find

$$f_{l,l-1} \sim e^{(l-1)\phi} \sin^{l-1} \theta \cos \theta$$

These eigenfunctions are called *spherical harmonics* and play a crucial role in the states of 3D systems which are rotationally invariant.

Spin

Usual orbital angular momentum corresponds to integral l . But it is observed that particles may also possess *intrinsic* angular momentum analogous to rotation about an axis - called *spin*. This may be half-integral. For example, the electron has spin $1/2$ meaning that $L^2 = 1/2(1/2+1)\hbar^2$ and its projection on the z -axis is $\pm \frac{1}{2}\hbar$. It is an example of a 'exact' two state system, similar to the ones introduced earlier to describe (approximately) chemical bonding. In this case the representation we have derived for integral l is not valid. Instead the spin operators can be represented by 2×2 matrices - the σ (or Pauli) matrices introduced in the homework. A general spin state may be represented by a 2 component column vector or *spinor*.

$$|\chi\rangle = a \left| \frac{1}{2} \right\rangle + b \left| -\frac{1}{2} \right\rangle$$

The components of $|\chi\rangle$ are the vector $\begin{pmatrix} a \\ b \end{pmatrix}$. Also,

$$\begin{aligned} S_x &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ S_y &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ S_z &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned}$$

As we have seen these obey the fundamental commutation relations of angular momentum and yield the correct eigenvalues for S^2 and S_z . Suppose the system is known to be in the state (a, b) and we ask the question: what is the probability of finding the particle 'spin-up' along the x-axis? First, we must confirm that S_x has the same two eigenvalues $\pm\frac{1}{2}$ and then express the state in terms of the corresponding eigenvectors of S_x . This yields

$$|\chi\rangle = \left(\frac{a+b}{\sqrt{2}}\right) \left|\frac{1}{2}\right\rangle^x + \left(\frac{a-b}{\sqrt{2}}\right) \left|-\frac{1}{2}\right\rangle^x$$

Thus $\frac{1}{2}|a+b|^2$ is the probability of finding the particle spin up along the x-axis!

Spin in a magnetic field

It is known that a particle with spin interacts with a magnetic field B with Hamiltonian

$$H = -\gamma B \cdot S$$

For an electron with magnetic field in the z-direction Schroedinger's equation takes the form

$$i\hbar \frac{d}{dt} \begin{pmatrix} a \\ b \end{pmatrix} = -\gamma B \hbar / 2 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$$

The general state is then a linear combination

$$\chi = \begin{pmatrix} \cos(\alpha/2)e^{i\gamma Bt/2} \\ \sin(\alpha/2)e^{-i\gamma Bt/2} \end{pmatrix}$$

where we have automatically imposed a normalization condition on the state. Thus

$$\langle S_z \rangle = \cos^2(\alpha/2) - \sin^2(\alpha/2)\hbar = \cos\alpha\hbar/2$$

Similarly

$$\langle S_x \rangle = \chi^\dagger S_x \chi$$

with χ as above. We find

$$\langle S_x \rangle = \frac{\hbar}{2} \sin\alpha \cos(\gamma B t)$$

and a similar result for the y-component. We see that the spin vector classically precesses around the field direction with angular frequency (the Larmor frequency) $\omega = \gamma B$.

Addition of angular momentum

Consider two spin 1/2 particles. What is the total angular momentum for the system? More generally, consider two particles with angular momentum eigenvalues l_1, m_1 and l_2, m_2 . Think of them initially as very far apart or *uncoupled*. It is easy to see that there are 4 commuting angular momentum operators for this system: $L^2(1), L_z(1), L^2(2)$ and $L_z(2)$. Any linear combinations of them will also be commuting. The obvious ones we are interested in are $L_z = L_z(1) + L_z(2)$ and $L^2 = \mathbf{L}(1) + \mathbf{L}(2)$ together with the squares of the individual angular momenta. We can equally well choose these latter four as the maximal commuting set. The eigenstates of these will correspond to the total angular momentum of the combined system and its component along some axis. The eigenstates of these operators will be built out of *products* of eigenstates of the original angular momentum operators

$$|l_1 m_1 l_2 \rangle = \sum c_{m_1, m_2} |l_1 m_1 \rangle |l_2 m_2 \rangle$$

If, initially there are $(2l_1 + 1)(2l_2 + 1)$ states for the two particles this will be preserved in this new representation.

Clearly, the maximal value for m will be $l_1 + l_2$. This means the two particles can be found in a state with $l_{\max} = l_1 + l_2$. The minimal value of m will be $|l_1 - l_2|$. In fact we can show that this yields l_{\min} the possible minimal value of the total angular momentum quantum number. Since angular momentum is quantized we expect any angular momentum in between will

also be seen. One way to see this is to see what value of l_{\min} will yield the correct total number of states

$$\sum_{l_{\min}}^{l_1+l_2} (2l+1) = (2l_1+1)(2l_2+1)$$

To make things more concrete reconsider our original example of the spin of two electrons. The possible combined state with largest z-component of angular momentum $m = 1$ is are

$$|\uparrow\rangle_1 |\uparrow\rangle_2$$

Lets operate with $L_- = L_-(1) + L_-(2)$ on this. We get

$$|\downarrow\rangle_1 |\uparrow\rangle_2 + |\uparrow\rangle_1 |\downarrow\rangle_2$$

This has $m = 0$. Operate again with L_- and we find

$$|\downarrow\rangle_1 |\downarrow\rangle_2$$

which has $m = -1$ as expected for an $l = 1$ state. But wait, this is only 3 states and I started with four base states. Also, my previous prescription implies there is also the possibility of a spin zero state $l = 0$ which must have $m = 0$. Of course, there is indeed another $m = 0$ state orthogonal to the one we have already written down. It is

$$|\uparrow\rangle_1 |\downarrow\rangle_2 - |\downarrow\rangle_1 |\uparrow\rangle_2$$

Thus the combined system can exist in a triplet $l = 1$ or singlet $l = 0$ state.

1 Lecture 11.

QM in three dimensions

The Schroedinger equation is

$$i\hbar \frac{\partial \Psi}{\partial t} = H \Psi$$

where the hamiltonian takes the form

$$H = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) + V(x, y, z)$$

and $p_x \rightarrow \frac{\hbar}{i} \frac{\partial}{\partial x}$ etc. Normalization of the wavefunction is simply

$$\int dx dy dz |\Psi|^2 = 1$$

The time-independent Schroedinger equation is just the same with E replacing the LHS. For systems which are rotationally invariant (i.e where the potential depends only on r) we may guess that solutions will look simpler in spherical coordinates (r, θ, ϕ) . To solve the Schroedinger equation in such a coordinate system we again try to find *separable* solutions - that is we write $\Psi(x, y, z) = R(r)Y(\theta, \phi)$. This means we must transform terms involving derivatives in x eg. $\frac{\partial^2}{\partial x^2}$ into derivatives with respect to r, θ and ϕ . This is a straightforward but tedious calculus exercise. I won't do it here but simply quote the result We find

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

This leads to two equations

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

and

$$\frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) - \frac{2mr^2}{\hbar^2} (V(r) - E) = C$$

Now, since we have shown that the Hamiltonian commutes with the angular momentum generators, the states of the system will be eigenstates of the

angular momentum operators - thus, the angular equation here must simply generate the *spherical harmonics* we have already met. In which case the constant $C = l(l + 1)$ with l integral (it must be L^2 that comes in since this is the only operator which does not single out a direction in space - unlike L_z). Thus the angular shape of the allowed wavefunction does not depend on the details of the potential if it only depends on radial distance r . We can simplify the radial equation still further by making the change of variables $R = u/r$. Hence $\frac{dR}{dr} = \left(r \frac{du}{dr} - u\right) / r^2$ and $\frac{d}{dr} \left(r^2 \frac{dR}{dr}\right) = r \frac{d^2u}{dr^2}$. We find

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

This is identical in form to the 1D Schroedinger equation except that the potential is replaced by the effective potential

$$V_{\text{eff}} = V + \frac{\hbar^2}{2m} \frac{l(l+1)}{r^2}$$

The normalization condition becomes just

$$\int dr u^2 = 1$$

If you don't like all this math it is also possible to guess this equation for $u(r)$. The argument goes like this; the Hamiltonian is rotationally invariant - hence it can only depend on the radial coordinate r plus constants. We might guess

$$H \sim \frac{1}{2m} p_r^2 + V(r)$$

where p_r denotes the momentum along the radial direction. This cannot be the full story since such a Hamiltonian does not give the correct expression for the energy of classical circular motion. The latter can be written

$$E_{\text{circular motion}} = \frac{L^2}{2mr^2}$$

where L is the angular momentum of the particle. Now we know that \mathbf{L}^2 is an operator in QM which commutes with H for a rotationally invariant system. Hence we can add such a term to our previous quantum Hamiltonian without changing the rotationally invariant nature of the Hamiltonian. Such a Hamiltonian will give the correct expression for purely radial motion

and also for simple circular motion. In addition we know that eigenstates of the Hamiltonian will be simultaneously eigenstates of L^2 - their angular dependence will hence just be given by the spherical harmonics we introduced previously! Furthermore, for such states we may replace the operator L^2 by its eigenvalue $l(l+1)\hbar^2$ and we arrive at the same equation as before.

Lets solve this first in the case where $V = 0$ for $R < a$ and is infinite for $r > a$. The radial equation inside the well reads

$$\frac{d^2u}{dr^2} = \left[\frac{l(l+1)}{r^2} - k^2 \right] u$$

where $k = \sqrt{2mE}/\hbar$. We need to solve this equation subject to the boundary condition $u(a) = 0$. The case $l = 0$ is easy

$$u = A \sin(kr) + B \cos(kr)$$

Since the true wavefunction is u/r so $B = 0$. The other boundary condition then requires $\sin(ka) = 0$ which yields $ka = n\pi$. The allowed energies are simply

$$E_n = \frac{n^2 \pi^2 \hbar^2}{2ma^2}$$

and the allowed wavefunctions are just

$$\psi = \sin(n\pi r/a)/r$$

The general solution for nonzero l is $r j_l(kr)$ where the function $j_l(kr)$ is called a spherical Bessel function. The allowed energies are given by the solution of the equation

$$j_l(ka) = 0$$

which yields an infinite discrete set of solutions. If β_{nl} is the n^{th} solution to this equation we find the allowed energies

$$E_{nl} = \frac{\hbar^2}{2ma^2} \beta_{nl}^2$$

and the wavefunctions are just

$$\psi(R, \theta, \phi) = j_{nl}(\beta_{nl}r/a) Y_{lm}(\theta, \phi)$$

Notice that each energy level is $(2l+1)$ fold *degenerate* which reflects the fact that the energy of the system does not depend on choice of the z -axis.

The hydrogen atom

To find the energy levels and wavefunctions we merely substitute in the form for the Coulomb potential $V = \alpha/r$ and solve the radial equation again. The solutions are again a set of special functions - the *Laguerre functions*. Its easy to modify the C code I gave you to determine the energies and wavefunctions. You will see that all the allowed wavefunctions are oscillatory but with exponentially decaying amplitudes. The allowed energy levels turn out to go like

$$E_n = - \left[\frac{m}{2\hbar^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \right] \frac{1}{n^2}$$

and l is restricted to lie in the range $0 \dots n - 1$. Notice that each of these states is $2l + 1$ degenerate also. For the ground state the energy in physical units is $E_0 = -13.6\text{eV}$.

Spectrum

If you put an atom in some stationary state then it will stay there forever but if that state is not one of the lowest energy it will be prone to make transitions to lower energy states (this can be caused by thermal collisions between atoms for example). To conserve energy a photon whose energy matches the atomic energy differences is emitted.

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

Transitions to the ground state lie in the UV - Lyman series. The Balmer series are transitions to $n = 2$ etc and occur in the visible part of the spectrum. This spectra can be produced by first passing an electric spark through the gas and then watching for the photons produced as the atoms relax back to lower energies.

Indistinguishability, atoms and the periodic table

Suppose I want to study a system with more than one particle. One might want to start with systems in which the two particles do not interact strongly. In this case the allowed energies/wavefunctions can be computed from 2 independent Schroedinger equations. Just as in the case with two independent

angular momenta we can write down a wavefunction for the combined system by taking products of the *single particle* wavefunctions.

$$\Psi(r_1, r_2) = \psi_a(r_1)\psi_b(r_2)$$

Of course, we are implicitly assuming that we can tell which particle is which - this is absolutely impossible for two identical particles such as electrons. We have already seen a precursor to this when we discussed simple 2-state models for covalent chemical bonding - we found that the allowed states of the system were linear combinations of states in which the electrons initially labelled were at one or the other of the two parent atoms. Thus we found that

$$\Psi(r_1, r_2) = \psi_a(r_1)\psi_b(r_2) \pm \psi_a(r_2)\psi_b(r_1)$$

Indeed, quantum indistinguishability *forces* us to use such a basis of states when discussing systems with more than one identical particle. If we introduce the exchange operator P such that

$$Pf(r_1, r_2) = f(r_2, r_1)$$

then clearly $P^2 = 1$ and it follows that the eigenvalues of P are ± 1 . This operator will commute with the Hamiltonian and hence we can find a set of functions which are simultaneously eigenstates of P and H . Such eigenstates will either be even or odd under exchange of particle label. These are just the previous wavefunctions. Particles for which $P = 1$ are termed bosons and those with $P = -1$ are called fermions. In relativistic QM we can also show that fermions have half-integral spin while bosons have integer spin (this is called the CPT-theorem). In our case we must just postulate it as an additional assumption.

Notice immediately that two identical fermions *cannot* occupy the same state - for then $\Psi = 0$. This is termed the Pauli exclusion principle. Its use in combination with the existence of discrete allowed energies gives an explanation for the atomic structure of atoms and the periodic table. Imagine that as a first approximation we treat each electron around the atom as moving independently of all the others (that is we neglect their mutual electrical interaction and use only the Coulomb force of attraction of each electron to the positively charged nucleus). Just as in the hydrogen atom if I solve the Schrodinger equation for this system I will find a set of *single particle* allowed energies. Each of these will be labelled by some radial quantum number and an angular momentum. Typically the energy of the system increases

with both radial quantum number (shell number) and angular momentum. To assemble the atom we put electrons in one by one in order of increasing energy - the Pauli principle stops us putting 2 electrons in the same state (thus we can put 2 electrons in a state with $l = 0$ since they can have distinct spin states, 6 electrons in a state with $l = 1$ since there two spin states for every value of l_z). Once we have filled a given l -state the resultant electrons have no net spin and angular momentum and such a shell is chemically inert. Thus the chemistry of an atom is determined only by its outermost partially filled single electron states. This leads to a *periodic* variation in the properties as the atomic number (number of electrons) increases. If we include electron repulsion and spin properly into this model we can account quantitatively for all the known chemistry of the atoms.

Lecture 12.

Symmetry and conservation laws in QM

Imagine computing the time variation of the expectation value of some observable corresponding to the operator A . We have

$$\frac{d \langle A \rangle}{dt} = \frac{\partial \langle \Psi |}{\partial t} A | \Psi \rangle + \langle \Psi | A \frac{\partial | \Psi \rangle}{\partial t}$$

From the Schroedinger equation we have

$$\frac{d \langle A \rangle}{dt} = i/\hbar \langle \Psi | [HA - AH] | \Psi \rangle$$

Thus any operator which commutes with the Hamiltonian will have an expectation value in *any* state which does not vary with time - it is *conserved*. Furthermore, we know from our discussion of angular momentum that eigenstates of the hamiltonian can also be chosen to be eigenstates of the operator A if it commutes with H . Thus they can be labelled with the corresponding eigenvalue of A . This *quantum number* will not then change with time.

There are three basic operators we know that satisfy classical conservation laws - the energy, momentum and angular momentum. In QM the associated operators will hence commute with the Hamiltonian (in the first case this is of course trivial!). We might ask the question: is there any fundamental principle of nature which might guarantee that these classical conservation laws also hold true in QM? The answer is yes and it has to do with symmetry.

Consider the operator $O = \exp(iap/\hbar)$. By expanding this operator in powers of p - the momentum operator it can be seen that this operator effects a translation of the wavefunction by the distance a . If the system is translation invariant this does not change the energy of the state. This will be true if the Hamiltonian of the system is invariant under translation and it follows that O must commute with H . This in turn means that p will commute with H and hence will be conserved. This invariance of the hamiltonian is called a symmetry - and we see the intimate connection between symmetry and conservation laws. A similar result is true for angular momentum and symmetry under rotations.

Perturbation Theory

Suppose we have solved the Schrodinger equation for some potential V_0 and now want to find the eigenvectors/values for another potential which differs from V_0 by a small amount. Perturbation theory is a tool for finding approximations to the latter knowing the eigenvectors/values for V_0 . Writing

$$H = H_0 + \lambda V$$

We will take λ to be a small number. We write

$$|\psi_n\rangle = |\psi_n^0\rangle + \lambda|\psi_n^1\rangle + \dots$$

$$E_n = E_n^0 + \lambda E_n^1 + \dots$$

If we substitute this ansatz into the Schrodinger equation and equate powers of λ we find

$$V|\psi_n^0\rangle + H_0|\psi_n^1\rangle = E_n^1|\psi_n^0\rangle + E_n^0|\psi_n^1\rangle$$

Taking the dot product with $\langle\psi_n^0|$ we find

$$E_n^1 = \langle\psi_n^0|V|\psi_n^0\rangle$$

This is the first order shift in the energy. To find the eigenvector to first order in λ we expand it as

$$|\psi_n^1\rangle = \sum_m c_m |\psi_m^0\rangle$$

Notice that we do not need to include any term c_n in this expression as such a term is not $O(\lambda)$. Thus we find

$$\sum_m (E_m^0 - E_n^0) c_m^{(n)} |\psi_m^0\rangle = -(V - E_n^1) |\psi_n^0\rangle$$

Taking the dot product with $|\psi_l^0\rangle$ we find

$$(E_l^0 - E_n^0) c_l^{(n)} = -\langle\psi_l^0|V|\psi_n^0\rangle$$

Notice that the denominator is never zero unless there is *degeneracy*.

Fine Structure of Hydrogen

The spectrum of hydrogen is in leading approximation given by that of the non-relativistic Schrödinger equation with Coulomb potential. However, the true Hamiltonian of the system should treat the electron relativistically. Since terms are suppressed by $O(v^2/c^2)$ relative to the leading term their effects can be computed using perturbation theory. Using the (relativistic) formula

$$T = p^2/2m - p^4/8m^3c^2 + \dots$$

the lowest order correction to the kinetic energy is

$$E^1 = -\frac{1}{8m^3c^2} \langle p^2\psi | p^2\psi \rangle$$

Now $p^2|\psi\rangle = 2m(E - V)|\psi\rangle$. Hence

$$E^1 = -\frac{1}{2mc^2} \left[E_n^2 + 2E_n\alpha \left\langle \frac{1}{r} \right\rangle + \alpha^2 \left\langle \frac{1}{r^2} \right\rangle \right]$$

This yields

$$E^1 = -\frac{E_n^2}{2mc^2} \left[\frac{4n}{l + 1/2} - 3 \right]$$

Notice that this breaks the original degeneracy of the levels and makes them l -dependent ($\alpha = e/4\pi\epsilon_0$).

There is a further contribution to the energy of the electron coming from relativity. Imagine the electron 'at rest' with the proton circling around it. Such a charge will generate a magnetic field which interacts with the electron's spin in the way we described earlier

$$V = -\gamma S \cdot B$$

Notice that the magnetic field will be proportional to the orbital angular momentum vector of the electron.

$$B = \alpha \frac{1}{mc^2} \frac{1}{r^3} L$$

The factor γ is given by

$$\gamma = -\frac{e}{2m} S$$

Thus we expect a lowest order contribution to the energy of the form from this *spin-orbit* term to be

$$E_n^1 = \left\langle \frac{e^2}{8\pi\epsilon_0} \frac{1}{m^2 c^2 r^3} S \cdot L \right\rangle$$

Now, this term does not commute with L and S separately so these two are not separately conserved. However, V does commute with $J = L + S$. Thus states of the hydrogen atom should be labelled by the conserved quantum numbers n and j . Furthermore, $L \cdot S = \frac{1}{2}(J^2 - S^2 - L^2)$. Thus the eigenvalues of this new term are proportional to

$$j(j+1) - s(s+1) - l(l+1)$$

This allows us to write the electron energy shift as

$$E^1 = \frac{E_n^2}{mc^2} \left[\frac{n[j(j+1) - l(l+1) - 3/4]}{l(l+1/2)(l+1)} \right]$$

Combining this result with the correction from the kinetic energy we find the final *fine structure* formula

$$E_{nj} = -\frac{13.6}{n^2} \left[1 + \frac{\alpha^2}{n^2} \left(\frac{n}{j+1/2} - \frac{3}{4} \right) \right]$$

Degenerate Perturbation Theory

If the unperturbed states are degenerate in energy then ordinary perturbation theory fails (unless the two degenerate states have $\langle \psi_a | V | \psi_b \rangle = 0$ which occurs if the perturbation commutes with the original Hamiltonian - this is the case for the fine structure calculation above). Thus we must find some other way to handle the problem. Suppose that $|\psi_a\rangle$ and $|\psi_b\rangle$ are two orthogonal, degenerate states with energy E_0 . Notice that *any* linear combination of states will also be an eigenstate of the unperturbed Hamiltonian. Typically, V will break this degeneracy. The clue as to how to proceed can be seen from the formula for the coefficients $c_m^{(n)}$. The two (say) degenerate states will start to dominate in this sum and we may analyze the system as a simple 2-state system by including only those two states in the equation for the new state vector. We will find

$$H^0 \psi^1 + V \psi^0 = E^0 \psi^1 + E^1 \psi^0$$

with

$$\psi^0 = \alpha\psi_a^0 + \beta\psi_b^0$$

we find by taking appropriate dot products that

$$\alpha \langle \psi_a^0 | V | \psi_a^0 \rangle + \beta \langle \psi_b^0 | V | \psi_b^0 \rangle = \alpha E^1$$

This is just a simple matrix problem where the matrix elements are just taken with respect to the original, unperturbed wavefunctions.