Second Year Quantum Mechanics - Lecture 1
Introduction and Background

Paul Dauncey, 7 Oct 2011

1 Waves and particles

Einstein's work on the photoelectric effect (which applies to photons) showed the photon energy was related to the frequency by

$$E = h\nu = \frac{h}{2\pi} 2\pi\nu = \hbar\omega$$

Following this, de Broglie postulated that this relation applied to all particles, including matter particles like electrons. Critically, he also realised that momentum was related to wavelength

$$p = \frac{h}{\lambda} = \frac{h}{2\pi}\frac{2\pi}{\lambda} = \hbar k$$

where k is the "wave number", which has units of rad/m and so kl gives the phase change over a distance l. Implicit in these equations is the fact that these objects act as both particles and waves. Specifically, the photoelectric effect says an electromagnetic wave of frequency ν will deposit energy in packets of $h\nu$ into a single atomic electron, which seems as if the EM field were a particle at a single point. Hence the EM field behaves as a wave until it interacts (until we "measure" it) at which point it acts as a particle. Similarly, the two-slit experiment with electrons shows the electron behaves as a wave until it impacts the screen, when we "measure" its position; it then behaves as a point-like particle.

Generally, there seems to be a separation between "propagation" with time (which is wave-like and we will see is governed by the Schrödinger wave equation) and "measurement" (which can be particle-like and is instantaneous and does not follow the Schrödinger equation). This bizarre behaviour, that the systems do not obey what we consider to be the fundamental equation when being measured, is a basic concept of quantum mechanics and its meaning has been hotly debated ever since quantum mechanics was first developed, with no sign of agreement yet. Not only does measurement not obey the Schrödinger equation but in general it gives random results. This is not "apparently" random, due to something we cannot measure, but genuinely random and this forms a fundamental physical principle. Quantum mechanics allows us to calculate to high precision the probability of the measurement giving a value, but the outcome of any particular measurement cannot be predicted in general.

The photoelectric effect and two slit experiment are examples of quantum phenomena; these happen because of quantum mechanics but do not give an overall picture. This is similar to the many classical phenomena which were known before Newton, such a Kepler's laws of planetary motion, Galileo's experiment from the Tower of Pisa, ballistics of Roman catapults, etc. However, only when Newton came along with his three laws could all these phenomena be understood as aspects of one theory. In an equivalent way, these lectures will set out the principles and postulates of quantum mechanics, so the phenomena can be understood at a fundamental level; we will develop the basic mechanics laws but for quantum systems rather than classical ones.

2 Classical Hamiltonian equations

Clearly, quantum mechanics has to agree with classical mechanics for large objects and we will want to show this explicitly, so let's very briefly review classical mechanics.

You will have seen classical mechanics done in terms of Newton's laws, from which we use

$$F = ma = m\frac{d^2x}{dt^2}$$

for one-dimensional motion, i.e. $x = x(t)$. This is a second order equation so there are two integrations to be done to solve this equation, which therefore require two constants of integration. These can be taken as the two initial conditions required to completely specify the problem to be solved. The two initial conditions are often given as $x(t = 0)$ and $dx/dt(t = 0)$.

It turns out that in quantum mechanics, we rarely talk in terms of forces and acceleration. Instead, we will work with potentials and energy, and we will restrict ourselves to conservative systems where the energy is constant. There are several alternative formulations of classical mechanics, all of which are based on the same physical principles and all of which give the same answer. One in particular is called the Hamiltonian method. In this we work with two first-order equations rather than one second-order one. This requires us to solve for two variables, not just one. The variables chosen are $x(t)$ and the momentum $p(t)$ where the latter is related to dx/dt by

$$p = m\frac{dx}{dt}$$

so that the above Newton's law becomes

$$\frac{dp}{dt} = F = -\frac{dV}{dx}$$

Note, two first order equations still require two integrations and hence there will still need to be two constants of integration. These are often the initial conditions $x(t = 0)$ and $p(t = 0)$.

In the Hamiltonian method, a function called the Hamiltonian is formed, which is the energy of the system. It is a function of our two variables

$$H(x,p) = T + V = \frac{p^2}{2m} + V(x)$$

and in this formalism, the equations of motion for $x(t)$ and $p(t)$ are given by

$$\frac{dx}{dt} = \frac{\partial H}{\partial p}, \qquad -\frac{dp}{dt} = \frac{\partial H}{\partial x}$$

From the above, these are found to be

$$\frac{dx}{dt} = \frac{p}{m}, \qquad -\frac{dp}{dt} = \frac{dV}{dx}$$

where the second can be written as

$$\frac{dp}{dt} = -\frac{dV}{dx} = F$$

so these are seen to be exactly the same equations we found from Newton's laws. In QM, the equivalent of the Hamiltonian, called the *Hamiltonian operator*, plays a fundamental role.

In principle, the Hamiltonian method is just a different mathematical technique but of course we know the momentum p does have physical significance too so it can give different physics insights. The two methods are equivalent and the choice is really convenience for any given system. The Hamiltonian method can be used for other coordinate systems quite easily and even other pairs of variables than x and p. Most importantly for our purposes, when comparing with quantum mechanics, the Hamiltonian approach gives equations much closer to those in quantum mechanics, hence it is good to know of its existance.

3 Poisson brackets

There are some other useful classical results for comparing with quantum mechanics which we shall show here. Consider any function $Q(x,p)$ of the two variables x and p; how does it change with time? This is straightforward to find using the chain rule

$$\frac{dQ}{dt} = \frac{\partial Q}{\partial x}\frac{dx}{dt} + \frac{\partial Q}{\partial p}\frac{dp}{dt} = \frac{\partial Q}{\partial x}\frac{\partial H}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial H}{\partial x}$$

This combination of derivatives is called the *Poisson bracket* and can be formed using any two functions of x and p, not just something and the Hamiltonian. It occurs so often in the Hamiltonian formalism that it has its own symbol and it is written as

$$\frac{\partial Q}{\partial x}\frac{\partial R}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial R}{\partial x} = \{Q, R\}$$

for any functions $Q(x,p)$ and $R(x,p)$. Some trivial properties of a Poisson bracket are that

$$\{Q, R\} = \frac{\partial Q}{\partial x}\frac{\partial R}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial R}{\partial x} = -\left(\frac{\partial Q}{\partial p}\frac{\partial R}{\partial x} - \frac{\partial Q}{\partial x}\frac{\partial R}{\partial p}\right) = -\{R, Q\}$$

and

$$\{Q, Q\} = \frac{\partial Q}{\partial x}\frac{\partial Q}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial Q}{\partial x} = 0$$

Also

$$\{Q^n, Q\} = \frac{\partial Q^n}{\partial x}\frac{\partial Q}{\partial p} - \frac{\partial Q^n}{\partial p}\frac{\partial Q}{\partial x} = nQ^{n-1}\frac{\partial Q}{\partial x}\frac{\partial Q}{\partial p} - nQ^{n-1}\frac{\partial Q}{\partial p}\frac{\partial Q}{\partial x} = 0$$

In QM, the equivalent is the *commutator* and this also plays a very important role.

In terms of the Poisson bracket, the equation of motion for Q is given by

$$\frac{dQ}{dt} = \{Q, H\}$$

One special case is for the time dependence of the Hamiltonian itself, i.e. for $Q = H$, for which

$$\frac{dH}{dt} = \{H, H\} = 0$$

which demonstrates that the energy is conserved. Two other special cases are found for $\{H, Q\}$ where $Q = x$ or $Q = p$, i.e. treating them as (very simple) functions of x and p. These can trivially be found to give the equations of motion for x and p, so the latter are also seen to just be special cases of the general time dependence.

Finally, another property which we will refer to later is the Poisson bracket of x and p together. For this

$$\{x, p\} = \frac{\partial x}{\partial x}\frac{\partial p}{\partial p} - \frac{\partial x}{\partial p}\frac{\partial p}{\partial x} = 1 \times 1 - 0 \times 0 = 1$$

The relevance of all this to quantum mechanics will become apparent later in the course.

Second Year Quantum Mechanics - Lecture 2
The Schrödinger equation

Paul Dauncey, 11 Oct 2011

1 The classical wave equation

We already know a wave equation in classical physics, e.g. for the displacement of a guitar string along its length. The classical wave equation is

$$\frac{\partial^2 \psi}{\partial t^2} = v^2 \frac{\partial^2 \psi}{\partial x^2}$$

which you should have seen several times in the first year. The constant v is the wave velocity. Sines and cosines can be used as solutions; let's try

$$\psi(x,t) = \psi_0 \cos(kx - \omega t)$$

for some constant ψ_0. Substituting in this solution gives

$$\frac{\partial \psi}{\partial t} = \omega \psi_0 \sin(kx - \omega t)$$

so

$$\frac{\partial^2 \psi}{\partial t^2} = -\omega^2 \psi_0 \cos(kx - \omega t) = -\omega^2 \psi$$

and similarly

$$\frac{\partial^2 \psi}{\partial x^2} = -k^2 \psi$$

Hence, for this to be a solution, then we require

$$-\omega^2 = -k^2 v^2$$

which means for any wave satisfying the classical wave equation

$$v = \frac{\omega}{k} = \frac{2\pi\nu}{2\pi/\lambda} = \nu\lambda$$

as expected. Since the velocity of the waves is fixed, then this means we get the usual relation between the frequency and the wavelength, or wavenumber

$$\omega \propto \frac{1}{\lambda} \qquad \text{or} \qquad \omega \propto k$$

2 The matter wave equation

We want to use the de Broglie relations to come up with a wave equation for matter particles. The de Broglie relations are

$$E = \hbar\omega, \qquad p = \hbar k$$

We can use these equations to try to construct a wave equation for matter waves. This will be done through induction and analogy; it is not a proof in any way. Quantum mechanics cannot be deduced from stratch; it needs some postulates.

We know for free particles (i.e. with $V = 0$) that the energy is just the kinetic energy so $E = p^2/2m$. Substituting in the de Broglie relations gives

$$E = \hbar\omega = \frac{p^2}{2m} = \frac{\hbar^2 k^2}{2m}$$

which means

$$\omega = \frac{\hbar k^2}{2m}$$

so now

$$\omega \propto k^2$$

which is different from above. Hence, we cannot simply use the classical wave equation. We need to find the equation for which the above matter relation is a solution.

If we examine where the terms in ω and k came from in the classical wave equation case, then is is clear the power of each comes from the number of derivatives. Hence, if we want k^2, then we need to keep $\partial^2\psi/\partial x^2$. However, to get only a first power of ω, then we need not to have $\partial^2\psi/\partial t^2$, but only $\partial\psi/\partial t$. Therefore, we will try

$$\frac{\partial\psi}{\partial t} = \alpha\frac{\partial^2\psi}{\partial x^2}$$

for some constant α to be determined later. Let's try the cosine solution again; we now find

$$\omega\psi_0 \sin(kx - \omega t) = -\alpha k^2 \psi_0 \cos(kx - \omega t)$$

which cannot hold for all x and t so this cannot be a solution. Instead, we are forced to go to a complex field; let's see what happens if we try

$$\psi(x, t) = \psi_0 e^{i(kx - \omega t)}$$

Plugging this in the equation, we get

$$(-i\omega)\psi = \alpha(ik)(ik)\psi = -\alpha k^2 \psi$$

which gives

$$\omega = -i\alpha k^2$$

This agrees with what we want if we set

$$-i\alpha = \frac{\hbar}{2m}$$

which means

$$\alpha = \frac{i\hbar}{2m}$$

Hence, the wave equation we want is

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

which is normally written as

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

Why is it written like that? For our wave solution, we know

$$\frac{\partial^2\psi}{\partial x^2} = -k^2\psi$$

so the right hand side of the above equation becomes

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} = \frac{\hbar^2 k^2}{2m}\psi = \frac{p^2}{2m}\psi = E\psi$$

and hence written this way it gives the value of the energy. We started by assuming that $E = p^2/2m$, which is the equation for the energy of a free particle, i.e. not in a potential. More generally, then $E = p^2/2m + V$, where $V(x)$ is the potential energy of the particle. Taking a big leap of faith, we assume we can modify the above equation to be more generally

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

We will define a symbol $\hat{H}$ such that

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

This type of object is called an *operator* as the derivatives "operate" on the function it is applied to, here ψ. This particular operator is called the *Hamiltonian operator* and plays a central role in QM. Using $\hat{H}$, we can also write the equation as

$$i\hbar\frac{\partial\psi}{\partial t} = \hat{H}\psi$$

This is the Schrödinger equation, sometimes called the time-dependent Schrödinger equation, TDSE. We have not derived it but simply tried to give a plausibility argument that it might be correct. The only real test of this (or any other equation) is whether it gives results consistent with the real world.

3 The interpretation of the wavefunction

This complex function ψ is called the *wavefunction* and plays a central role in quantum mechanics. We have not discussed the meaning of this much; what is doing the "waving"? The Schrödinger equation contains i and so is complex, hence ψ will also be complex in general. This means we need to consider how to interpret ψ.

Consider the double-slit experiment with light. The light passes through and gives an interference pattern. Wave equations, here the Maxwell equations, apply to the electric field $\boldsymbol{E}$ and the field at each point has contributions from both slits. The intensity of the light at any point is given by $I \propto E^2$. As the intensity is reduced, the light is seen to be quantised, with single photons hitting at random locations on the screen. It is found that the probability of a photon being found at a given location is $\propto E^2$, which clearly becomes the intensity for large numbers of photons.

How do we generalise this for matter waves? The analogy (known as the Born rule) is that a wave equation, which for matter is the Schrödinger equation, applies to the "wavefunction" ψ and that the probability of finding the particle at a given point is proportional to $|\psi|^2$. Note, because ψ is complex in general for matter waves, then we need to take the square of the modulus of ψ, i.e. $|\psi|^2 = \psi^*\psi$, to get the probability, as this is real and positive, by definition. To be exact, the quantity $|\psi|^2$ is the *probability density*. Formally, the probability of finding the particle within a narrow range from x to $x + dx$ is $|\psi|^2\,dx$. This means that the probability of finding the particle between $x = a$ and $x = b$ is

$$P(a \leq x \leq b) = \int_a^b |\psi|^2\,dx$$

The total probability of finding the particle anywhere must be one, so this means

$$\int_{-\infty}^{\infty} |\psi|^2 \, dx = 1$$

We usually fix an overall multiplicative constant to make this true; this is called "normalisation" of the wavefunction. Note, this condition means the wavefunction must go to zero as x goes to infinity, or else this integral would not be finite.

Although the wavefunction is complex, the physical interpretation of all measurable quantities, such as the probability, comes through quantities which are definitely real. Hence, we do not end up with complex physical variables. In fact, the complex phase of ψ drops out in all physical variables; this is a basic property of quantum mechanics. Hence, you will often hear it said that the phase of the wavefunction is unobservable.

4 Use of complex variables

Even classically, a wave is often written as

$$\psi(x,t) = \psi_0 \cos(kx - \omega t) = \mathcal{R}e \left[\psi_0 e^{i(kx-\omega t)} \right]$$

but here the complex form is used for mathematical convenience and to get the actual result we take the real part. This is not true in quantum mechanics; in this case we don't just use the real part but must keep both parts as they both play a role.

We found the equation had to be complex to work. It is common to feel uncomfortable about this as the real world is indeed real. However, we should understand the complex wavefunction is nothing more than a convenient combination of two real fields. For the Schrödinger equation, it is convenient purely because of the way the equations of quantum mechanics turn out.

Any complex equation requires both the real parts and the imaginary parts to be equal and so it equivalent to two real equations. Let's see what happens if we actually write the equations in terms of the real and imaginary parts of ψ. Using the suggestive notation

$$\psi(x,t) = X(x,t) + iP(x,t)$$

where $X(x,t)$ and $P(x,t)$ are real functions, then the left hand side of the Schrödinger equation can be written as

$$i\hbar \frac{\partial \psi}{\partial t} = i\hbar \frac{\partial X}{\partial t} - \hbar \frac{\partial P}{\partial t}$$

For the right hand side of the Schrödinger equation, then note all the terms in $\hat{H}$ are real so it separates into real and imaginary parts as

$$\hat{H}\psi = \hat{H}X + i\hat{H}P$$

This means the Schrödinger equation becomes the two equations; the imaginary and real parts give

$$\frac{\partial X}{\partial t} = \frac{1}{\hbar}\hat{H}P, \qquad -\frac{\partial P}{\partial t} = \frac{1}{\hbar}\hat{H}X$$

respectively. These should be compared with Hamilton's classical equations we saw in the last lecture

$$\frac{dx}{dt} = \frac{\partial H}{\partial p}, \qquad -\frac{dp}{dt} = \frac{\partial H}{\partial x}$$

Hence, we understand the Schrödinger equation to be of a similar form to Hamilton's equations and so it is not too suprising that our comparisons with classical mechanics later will be more easily done in the Hamiltonian formalism too.

Second Year Quantum Mechanics - Lecture 3
Solving the Schrödinger equation

Paul Dauncey, 14 Oct 2011

1 Solving the Schrödinger equation

In the last lecture, we tried to justify the form of the time-dependent Schrödinger equation (TDSE)

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

This is a linear partial differential equation (PDE) and these can be tricky to solve. The standard mathematical trick is to try "separation of variables" (SoV). We assume we can write

$$\psi(x,t) = u(x)T(t)$$

i.e. we separate the two variables into two functions, each only depending on one variable. Of course, it is not clear that all solutions will be of this form and in fact they are not. However, the mathematical trick is that we can solve for the SoV solutions and then make a general solution by adding SoV solutions together. Hence, we will continue by substituting this SoV form for ψ into the Schrödinger equation. This gives

$$i\hbar u\frac{dT}{dt} = -\frac{\hbar^2}{2m}T\frac{d^2u}{dx^2} + VuT$$

where the partial derivatives have become total derivatives as u and T only depend on one variable. Dividing throughout by $\psi = uT$, then this gives

$$i\hbar\frac{1}{T}\frac{dT}{dt} = -\frac{\hbar^2}{2m}\frac{1}{u}\frac{d^2u}{dx^2} + V$$

The left hand side is now a function of t only and the right hand side of x only. However, this means for a given time, no matter what the value of x, the right hand side must always give the same value, i.e. it is not a function of x after all. Similarly, for a fixed position, the left hand side cannot vary as t changes, so it must be equal to a value which does not depend on time. Hence, both sides must be equal to a constant, E, called the "constant of separation" and we have

$$i\hbar\frac{1}{T}\frac{dT}{dt} = E, \qquad -\frac{\hbar^2}{2m}\frac{1}{u}\frac{d^2u}{dx^2} + V = E$$

2 Time dependence

Consider the time-dependent part, which can be solved generally. For this

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

so

$$\frac{dT}{T} = -\frac{iE}{\hbar}dt$$

which then can be integrated to give

$$\ln(T) = -\frac{iEt}{\hbar} + C$$

where C is a constant of integration. This then gives

$$T = T_0 e^{-iEt/\hbar}$$

where $T_0 = e^C$ is a constant. Note for a wave, which goes as $e^{-i\omega t}$, the de Broglie relation for frequency says the energy is $\hbar\omega$ which implies the constant of separation E could be the energy.

The total solution includes the spatial part, which we haven't solved for yet, but it can be generally written as

$$\psi = u(x)e^{-iEt/\hbar}$$

What does this give for a probability density?

$$|\psi|^2 = \psi^*\psi = u(x)^* T_0^* e^{iEt/\hbar} u(x) T_0 e^{-iEt/\hbar} = u(x)^* u(x) T_0^* T_0 = |u(x)T_0|^2$$

i.e. it does not change with time. Hence, the SoV solutions are called "stationary states" (for obvious reasons) or "energy eigenstates" (for reasons which we will get on to in a later lecture).

3 Position dependence

The other side of the equation gives

$$-\frac{\hbar^2}{2m}\frac{1}{u}\frac{d^2u}{dx^2} + V = E$$

or

$$-\frac{\hbar^2}{2m}\frac{d^2u}{dx^2} + Vu = Eu \qquad \text{i.e.} \qquad \hat{H}u = Eu$$

This is called the time-independent Schrödinger equation (TISE) as it is clearly for a function of position only, not time. We cannot solve this generally, as we did for T, but given a particular physical situation, we know $V(x)$ and then can (in principle) solve this equation for u. We will generalise this to three dimensions later. Generally, there will be constraints on the solutions due to the intepretation of the wavefunction in terms of probability; see the lecture slides.

Note, all the terms in $\hat{H}$ are real so it is possible, and not uncommon, for the solutions $u(x)$ to also be able to be purely real.

4 Free particle solution

Let's check this all makes sense with where we started. The original equation was motivated by considering a free particle so let's solve the time-independent Schrödinger equation for this case, for which $V = 0$. Hence, we have

$$\frac{d^2u}{dx^2} + \left(\frac{2mE}{\hbar^2}\right)u = 0$$

which you should recognise as having exactly the same structure as an equation you have seen before

$$\frac{d^2l}{dt^2} + \alpha l = 0$$

The TISE in terms of x rather than t but mathematically the solutions don't care what the variable is, of course. For positive α, then we can write $\alpha = \omega^2$ and you will then recognise it as the simple harmonic oscillator equation which you know has solutions which are $\sin(\omega t)$ and $\cos(\omega t)$. However, for negative α, then we write $\alpha = -\gamma^2$ and the solutions are $e^{\gamma t}$ and $e^{-\gamma t}$.

The same is true for the time-independent Schrödinger equation above; mathematically, the solutions for $E < 0$ will give exponentials, now in x, but these will become infinite for $x \to \pm\infty$ and so are not physically allowed. Hence, we are required to have $E \geq 0$. For this case, we would expect oscillating solutions so let's try a sine

$$u = u_0 \sin(kx)$$

This gives

$$\frac{d^2u}{dx^2} = \frac{d^2}{dx^2}[u_0 \sin(kx)] = \frac{d}{dx}[ku_0 \cos(kx)] = -k^2 u_0 \sin(kx) = -k^2 u$$

Hence, this is a solution if k satisfies

$$-k^2 + \frac{2mE}{\hbar^2} = 0$$

so

$$E = \frac{\hbar^2 k^2}{2m} = \frac{p^2}{2m}$$

Hence, it is clear the constant of separation E is indeed the particle energy. It turns out this is a general result and is not just true for the free particle case.

We arbitrarily chose a sine above; a cosine would also have worked. The general solution is a combination of both

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

where A and B are constants which we have to allow to be complex. Since they do not have to be real, then the general solution can also be expressed alternatively as

$$u = Ce^{ikx} + De^{-ikx}$$

again for (generally complex) constants C and D. It is easy to see these are equivalent using

$$e^{\pm ikx} = \cos(kx) \pm i\sin(kx)$$

Note, we found in the last lecture than sines and cosines didn't seem to work; there we were trying a real solution for both x *and* t. We must not forget the complex $e^{-iEt/\hbar}$ time dependence here and this means the total solution is still complex, even if $u(x)$ is real.

The other thing to note is that any value of k will give a solution and gives a value of E which is always positive. Hence, a free particle can have any positive energy, as we would expect. However, we are restricted not to have negative energies for free particles. It is usual for the time-independent equation to give constraints on E although this one is quite a weak constraint. Boundary conditions or the shape of the potential mean the allowed energy values can be quite restricted in some cases.

The total solution for ψ for the free particle is then, picking the form we used previously

$$\psi = Tu = \psi_0 e^{-iEt/\hbar} e^{ikx} = \psi_0 e^{-i(Et-px)/\hbar} = \psi_0 e^{-i(\omega t - kx)}$$

where we have used the de Broglie relations $E = \hbar\omega$ and $p = \hbar k$. The probability is then

$$|\psi|^2 = \psi^*\psi = \psi_0^*\psi_0 e^{i(Et-px)/\hbar} e^{-i(Et-px)/\hbar} = \psi_0^*\psi_0 = |\psi_0|^2$$

and so is constant both in space and time. This seems a bit wierd; we think of a free particle as moving from somewhere to somewhere else but here the probability of finding the particle is the same everywhere and doesn't change with time. To get something which changes with time, we have to add the SoV solutions to get the more general solutions.

5 Superposition of states

Consider two solutions, ψ_1 and ψ_2, of the TDSE with some non-zero $V(x)$. By definition, ψ_1 and ψ_2 satisfy

$$i\hbar\frac{\partial\psi_1}{\partial t} = \hat{H}\psi_1, \qquad i\hbar\frac{\partial\psi_2}{\partial t} = \hat{H}\psi_2$$

Consider a sum of these two with some coefficients

$$\psi_s = \alpha\psi_1 + \beta\psi_2$$

for any values of α and β. For this, then

$$i\hbar\frac{\partial\psi_s}{\partial t} = \alpha i\hbar\frac{\partial\psi_1}{\partial t} + \beta i\hbar\frac{\partial\psi_2}{\partial t} = \alpha\hat{H}\psi_1 + \beta\hat{H}\psi_2 = \hat{H}(\alpha\psi_1 + \beta\psi_2) = \hat{H}\psi_s$$

so the *superposition* of the two wavefunctions also satisfies the Schrödinger equation. This is called the "principle of superposition". Note, this is true for any α and β, even complex values. This works because the Schrödinger equation is linear, i.e. it contains only first powers of the wavefunction ψ. The fact that superposition works has far-reaching consequences as we shall see later. This also allows us to make general solutions of the TDSE by adding SoV solutions.

This means we can take two stationary states and add them to give another solution of the TDSE. However, unless the energies of the two states happen to be equal, then the superposition state will *not* be a stationary state and will not be separable into functions of time and position. However, as shown, it will still be a solution of the TDSE. Let's take a simple case with

$$\psi_1 = u_1 e^{-iE_1 t/\hbar}, \qquad \psi_2 = u_2 e^{-iE_2 t/\hbar}$$

where the u_1 and u_2 are real functions and form a superposition state with $\alpha = \beta = 1$, i.e.

$$\psi_s = u_1 e^{-iE_1 t/\hbar} + u_2 e^{-iE_2 t/\hbar}$$

The probability density of the superposition state is then

$$|\psi_s|^2 = u_1^2 + u_2^2 + u_1 u_2\left[e^{i(E_1-E_2)t} + e^{i(E_2-E_1)t}\right] = u_1^2 + u_2^2 + 2u_1 u_2 \cos(E_1 - E_2)t$$

The third term is now time-dependent; the t terms do not cancel. Hence, we now have a probability density which changes with time. This is because we have combined two solutions with different energy, which therefore have different time dependences. Hence, although ψ_s is a solution of the TDSE, it is not a stationary state because it contains more than one energy value. Note, in contrast, if we chose solutions for which $E_1 = E_2$, then the cosine term above would be constant and we would be back to a constant probability. Hence, to get motion, i.e. a change to the probability density, we need to superimpose stationary states which results in a non-stationary total wavefunction. Classical motion, where we see particles moving around, corresponds to such states, normally a large superposition of many states.

To finish, an interesting question is; what is the energy of a superposition state? It turns out even this simple question is effectively impossible to answer and we will see why later in the course.

Second Year Quantum Mechanics - Lecture 4
The infinite square well

Paul Dauncey, 18 Oct 2011

1 Introduction

One of the most crucial phenomena explained by quantum mechanics was the existence of atomic spectra, e.g. the observed spectral lines of the hydrogen atom. These arise because the electron orbit only exists with particular energy values (due to quantum mechanics) and so movements between those values give out specific values of energy which are seen as specific frequencies (also due to quantum mechanics). We would like to demonstrate the first of these two quantum effects, i.e. that only particular energies are allowed. The hydrogen atom is mathematically quite complicated so we will start with a much simpler system.

2 The infinite square well

One of the simplest potentials for which the time-independent Schrödinger equation can be solved is the "infinite square well". By this, we mean a potential which is infinitely large everywhere except for a restricted region of space, where the potential is (defined to be) zero. This means the particle can only be found within this region. Classically, this corresponds to a particle in a box, bouncing back and forth. This is clearly an idealised system, but has many of the same features as more realistic potentials which bind particles to a restricted region of space.

Hence, we take

$$V = 0 \text{ for } |x| \leq a, \qquad V = \infty \text{ for } |x| > a$$

Within the square well, then the time-independent Schrödinger equation has the same form as for a free particle, i.e.

$$-\frac{\hbar^2}{2m}\frac{d^2u}{dx^2} = Eu$$

and so we know the solutions are sines and cosines (or equivalently complex exponentials). The general solution is then

$$u(x) = A\cos(kx) + B\sin(kx)$$

where the energy E is related to k through

$$E = \frac{\hbar^2 k^2}{2m}$$

What about outside the square well? The potential is infinite, so the Vu term in the Schrödinger equation will be infinite unless $u = 0$ throughout this region. Another way to see this is that the particle cannot be found in a region of infinite potential without giving it infinite energy. Hence, it must have zero probability of being outside the well, so $|u|^2 = 0$ everywhere outside, for which it is clear $u = 0$.

We now need to think about the boundaries. We have seen that the wavefunction must be continuous at the boundary. Hence, we need to pick the solutions within the well which go to zero at the boundaries, which are at $x = \pm a$. For $x = a$, then

$$u(a) = A\cos(ka) + B\sin(ka) = 0$$

and for $x = -a$, then

$$u(-a) = A\cos(ka) - B\sin(ka) = 0$$

Normally, we would also have to make sure the first derivative of the wavefunction is also continuous; otherwise this would make the second derivative, which appears in the Schrödinger equation, infinite. However, as here we have an infinite potential, this can "cancel" the infinite second derivative and so we are allows discontinuous derivatives in this case. We will justify this further in the next lecture.

For the above equations, obviously there is a trivial solution with $A = B = 0$ but this gives $u = 0$ even within the square well, meaning the probability of finding the particle is zero everywhere, i.e. there is no particle. We want a different solution so let's try adding the two equations to get

$$2A\cos(ka) = 0$$

which is satisfied if $A = 0$ or if

$$ka = \frac{(2m+1)}{2}\pi$$

for some integer $m \geq 0$, so that

$$k = \frac{(2m+1)\pi}{2a}$$

Now, for these values of ka, the sine cannot also be zero at the boundary, so to get $u(a)$ and $u(-a)$ to be zero, we have to set $B = 0$. Hence the solutions for these k values are purely cosines. Similarly, subtracting the above equations gives

$$2B\sin(ka) = 0$$

which is satisfied if $B = 0$ (which is the pure cosine solution above) or if

$$ka = m\pi$$

for integer $m \geq 1$, so that k is required to take only the values

$$k = \frac{m\pi}{a} = \frac{(2m)\pi}{2a}$$

For these k values, the cosine is not zero at the boundary and so we must set $A = 0$; these solutions are pure sines. These two sets of conditions on k can be summarised as

$$k_n = \frac{n\pi}{2a}$$

for any integer n, where we have

$$\begin{array}{lll} n = 1, 3, 5, \ldots & u_n(x) & = A\cos k_n x \\ n = 2, 4, 6, \ldots & u_n(x) & = B\sin k_n x \end{array}$$

Clearly, any values of A for the cosine solution (or B for the sine solution) will give a wavefunction which satisfies the Schrödinger equation. However, there is another constraint which we should take into account. If there is one particle in the well, then the probability of finding the particle anywhere in the well has to be one. The quantity $|\psi|^2$ is actually the probability density, meaning the probability of finding the particle between x and $x + dx$ is $|\psi(x)|^2 dx$. This means the probability of finding it within the square well must give one, i.e.

$$\int_{-a}^{a} |u(x)|^2\, dx = \int_{-\infty}^{\infty} |u(x)|^2\, dx = 1$$

since $u = 0$ outside the square well. This allows us to fix the magnitude of A or B.

3 Energy values

We have seen that requiring boundary conditions has restricted the allowed values of k. However, this also then means

$$E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{\hbar^2 \pi^2 n^2}{8ma^2}$$

which means the allowed energies of the particle are also restricted. This is very general; in quantum mechanics, a particle in a potential well cannot have an arbitrary energy in a stationary state, but only one of a restricted set of values. This is called "quantisation" of the energy and is of course where quantum mechanics gets its name.

The states above can be labelled by the integer n; we already did so for both k_n and E_n. The value of n defines both the wavenumber and the energy. Such labels are often called the "quantum numbers" of the states. They do not always have to be integers, although the infinite square well is one of several cases where they happen to be so.

The quantisation of wavenumber (or equivalently wavelength) due to boundary conditions is not new to quantum mechanics but occurs in classical physics also. A guitar string gives a "fundamental" note of a well-defined frequency because it is constrained not to move at both ends so the allowed wavelength has to fit into the distance available. Higher harmonics are also possible, as long as they also have no displacement of the string at the ends. In fact, the spectrum of vibrations of a guitar string are effectively identical to those of the quantum infinite square well, as the boundary conditions are the same.

4 Parity

It is often the case that a potential can have a symmetry, due to some property of the system being studied. An example is the gravitational field around the Earth which is the same in all directions (isotropic), at least in the idealised case of an exactly spherical Earth. Another form of symmetry is parity, which is the technical term for whether a function is reflection symmetric (or antisymmetric) i.e. comparing $f(x)$ to $f(-x)$. The infinite square well is a case of a reflection symmetric potential as $V(-x) = V(x)$ for all x.

Such potentials occur both in classical and quantum mechanics. In both cases, any given solution does not have to reflect the symmetry of the potential, but the family of all possible solutions will show the same symmetry. In the case of a reflection symmetry, the solutions can be symmetric or antisymmetric. For the infinite square well, the $n = 1, 3, 5, \ldots$ solutions are cosines, which have the property that $u(-x) = u(x)$ and so are symmetric. These are said to have a positive parity, or parity $= +1$. The $n = 2, 4, 6, \ldots$ solutions are sines, which have the property $u(-x) = -u(x)$ and so are antisymmetric. They are correspondingly said to have negative parity, or parity $= -1$. This can be summarised into parity $= (-1)^{n+1}$.

Note, however, that the probability distribution $|u|^2$ changes as $|u(-x)|^2 = |u(x)|^2$ for even parity states and $|u(-x)|^2 = |-u(x)|^2 = |u(x)|^2$ for odd parity states and so the probability is *always* even. This makes physical sense; why would the probability of the particle being on the left be different from on the right?

5 Time dependence

We need to remember that the solution of the TISE is not the full solution of the Schrödinger equation as we need to add the time-dependent part. The total wavefunction is

$$\psi_n(x,t) = u_n(x) e^{-iE_n t\hbar}$$

Any complex number z can be expressed as real and imaginary parts as $z = x + iy$ or can be expressed as a modulus $|z| = r$ and a phase θ as $z = re^{i\theta}$. Hence, we see that for real $u_n(x)$ as we have here, then $|\psi| = |u_n(x)| = u_n(x)$ and does not depend on t, while the phase is $-E_n t/\hbar$. Hence, the solution rotates clockwise round the Argand diagram. If $u_n(x)$ is not zero at some particular x, then the value of ψ also never goes to zero at that x. Note again; if the wavefunction has a unique energy, then there is no motion of the probability distribution. Only by superimposing solutions with more than one energy do we get a change with time.

6 General features of TISE solutions

We will move on to solving more complicated potentials than the infinite square well. However, we should think about the form of the solutions we might expect. We are solving the time-independent Schrödinger equation

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

which we can write as

$$\frac{d^2u}{dx^2} = -\frac{2m}{\hbar^2}[E - V(x)]u$$

Consider this for constant V to start with. As mentioned in the previous lecture, if the factor on the RHS is positive, i.e. if $E > V$, then we have sine and cosine solutions. For these

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

so the wavenumber is given by

$$k = \frac{\sqrt{2m(E-V)}}{\hbar}$$

and so for $E \gg V$, then we get large k and hence short wavelengths, while for E only just above V, we get small k and long wavelengths. The limit of $E \to V$ is a constant u.

In general, V is not constant but changes with position. If there were a step in V at some x, then the value of k would be different on each side of the step and we would have the wavefunction changing from long to short wavelengths. (A similar thing happens in optics when a light wave goes from one medium to another with different refractive indices.) Note, we need to make the wavefunction value and its derivative value the same on both sides of the step, because of our continuity restrictions. If there were many such steps, then we can see that the value of k to use would be different at each different position in x, so a generally changing V means we effectively get waves but with a wavelength that is position dependent; roughly speaking, $k = k(x)$. Note, unfortunately, things are *not* so simple as to do this generally using

$$k(x) \neq \frac{\sqrt{2m[E - V(x)]}}{\hbar}$$

so don't fall into that trap.

What if $E < V$? Here we get exponentials, but the same principle holds. With

$$u = Ce^{\pm\gamma x}$$

then

$$\frac{du}{dx} = \pm\gamma Ce^{\pm\gamma x} = \pm\gamma u$$

so

$$\frac{d^2u}{dx^2} = \gamma^2 u$$

Hence, for constant V, this is a solution if

$$\gamma = \frac{\sqrt{2m[V-E)]}}{\hbar} = ik$$

Again, for $V \gg E$, the exponential falls or rises very fast, while for V just above E, u changes very slowly. Again, in the limit, then as $V \to E$, then u is constant and so matches to the previous solution. Clearly, with $V(x)$, the γ will change with position, in a similar way to k. We can even mix such solutions; in some parts of space, $E > V$ and we get oscillations while in other parts of space, $E < V$ and we get exponential falloffs. We will see such a case as this in the next lecture.

Another way to see this is that the value of $-[E - V(x)]u$ at each point in space determines the "curvature" (i.e. the second derivative) of the solution. If E is greater than $V(x)$, then the curvature has the opposite sign to u, so a positive u will have the solutions bending negative and a negative u will have it bending positive. This is why these solutions oscillate. However, if E is less than $V(x)$, then the curvature has the same sign as u, so the solution will always tend to bend away from $u = 0$, giving exponential-like behaviour. Classically, the particle can only be found in the former regions, so when we see particles corresponding to the classical positions, they have associated waves, as postulated by de Broglie.

This leads to a very general argument that for any system, $E > V_{\mathrm{min}}$, where V_{min} is the lowest value of the potential anywhere in x. If this was not true, then for every value of x, $V > E$ and so there would be only exponential-like solutions everywhere. These would go to infinity for either (or both) of $x \to \pm\infty$. In terms of curvature, the curvature would always be away from the $u = 0$ axis and so would be always increasing for either (or both) of positive and negative x and so cannot go to zero at infinity. Such solutions are not allowed for energy eigenstates, so we conclude the energy must always be greater than the potential minimum.

Second Year Quantum Mechanics - Lecture 5
The finite square well

Paul Dauncey, 20 Oct 2011

1 Introduction

The infinite square well is clearly an idealisation, as a realistic potential cannot change from zero to infinity. The finite square well is therefore much closer to a realistic situation which could be measured in the laboratory and so we will look at this next. While it is mathematically harder than the previous case, there is nothing new in terms of physics and we will follow a very similar procedure.

2 The finite square well

For this, we will take the potential to be

$$V(x) = -V_0 \text{ for } |x| \leq a, \qquad V(x) = 0 \text{ for } |x| > a$$

for a (positive) constant V_0. Within the square well, then the time-independent Schrödinger equation has the same form as before, i.e.

$$-\frac{\hbar^2}{2m}\frac{d^2u}{dx^2} = (E - V)u = (E + V_0)u$$

while outside the square well, the equation is very similar but with a different constant on the right

$$-\frac{\hbar^2}{2m}\frac{d^2u}{dx^2} = Eu$$

Because the well is now finite, it is possible to have solutions with energy above the potential at $x \to \pm\infty$, i.e. $E > 0$ as well as $-V_0 < E < 0$ solutions. No physical solutions are possible for an energy below the potential minimum, i.e. $E < -V_0$.

3 Unbound states

The case with solutions for $E > 0$ has oscillatory solutions in both regions, i.e. inside and outside the well. Here, the particle acts similarly to a free particle, but the well causes a change to the wavelength when the particle is inside it. This is analogous to the classical case of a free particle moving over a potential well which does not trap it; it will speed up and slow down as it moves past, depending on the slope of the potential, but it will not be stopped. There is no constraint on energy for these solutions (as long as $E > 0$ of course). The particle is not kept close to the well and so is not constrained to that region of x. Such solutions are said to be "unbound states".

4 Bound states

The other case of solutions have $-V_0 < E < 0$, meaning that we would expect oscillations within the well but exponential falloff outside the well. This means the particle is kept within, or close to, the well and so these states are called "bound states". This is analogous to a classical particle trapped in a potential well.

Inside the well, we can write

$$\frac{d^2u}{dx^2} + k^2 u = 0$$

while outside, since $E < 0$, we have

$$\frac{d^2u}{dx^2} - \gamma^2 u = 0$$

where

$$k = \frac{\sqrt{2m(E+V_0)}}{\hbar}, \qquad \gamma = \frac{\sqrt{-2mE}}{\hbar},$$

The full derivation of the solutions is given in Handout 2. Here, I will sketch out how they come about. Let's assume the lowest energy state is again even parity so it will be a cosine within the well and exponentials outside the well with equal rates of decay on both sides. Hence the solution will go something like

$$\begin{aligned} u_L(x) &= Be^{\gamma x} \\ u_C(x) &= A\cos(kx) \\ u_R(x) &= Be^{-\gamma x} \end{aligned}$$

The derivatives are

$$\begin{aligned} \frac{du_L}{dx} &= \gamma Be^{\gamma x} \\ \frac{du_C}{dx} &= -kA\sin(kx) \\ \frac{du_R}{dx} &= -\gamma Be^{-\gamma x} \end{aligned}$$

Matching the solutions at $x = a$ (or $x = -a$) gives

$$Be^{-\gamma a} = A\cos(ka)$$

Matching the derivatives is now required as there is no longer an infinite potential in the system. Hence matching at $x = a$ (or $x = -a$) gives

$$-\gamma Be^{-\gamma a} = -kA\sin(ka)$$

so combining these two gives

$$\begin{aligned} -\gamma A\cos(ka) &= -kA\sin(ka) \\ \gamma &= k\tan(ka) \end{aligned}$$

A similar calculation, using $\sin(kx)$ to get odd parity states, yields the equation

$$\gamma = -k\cot(ka)$$

Note, $k = k(E)$ and $\gamma = \gamma(E)$ so some arbitrary energy value is highly unlikely to give values of k and γ which satisfy the above tangent or cotangent relation. Hence, these are equations which will only be satisfied by E for particular values. Hence, these equations give the energies of the solutions and the energy is again quantised.

5 Energy values

Unfortunately, the above equations cannot be solved analytically but can be done graphically or numerically.

It is good to do this in a way that gives an intuitive grasp on what is going on; e.g. how do the number of bound states depend on V_0, is there always at least one bound state, etc? One way of plotting the graphical solution is to note that since

$$k = \frac{\sqrt{2m(E+V_0)}}{\hbar}, \qquad \gamma = \frac{\sqrt{-2mE}}{\hbar},$$

then

$$\begin{aligned} \gamma^2 + k^2 &= \frac{2mV_0}{\hbar^2} \\ \gamma &= \sqrt{\frac{2mV_0}{\hbar^2} - k^2} \end{aligned}$$

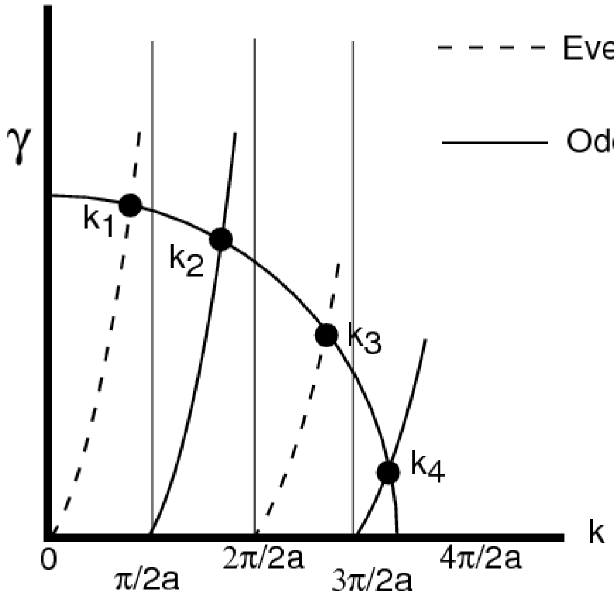
Hence, we can consider this, together with $\gamma = k\tan(ka)$ (or $\gamma = -k\cot(ka)$), as two equations for γ in terms of k. As usual, if you have $y = f(x)$ and $y = g(x)$, you can find the solutions by plotting $f(x)$ and $g(x)$ and seeing where they cross. Here we have $\gamma = f(k)$ and $\gamma = g(k)$ but it is exactly the same principle. The useful thing about doing it this way is that

$$\gamma = k\tan(ka), \qquad \gamma = -k\cot(ka),$$

are independent of V_0 and

$$\gamma = \sqrt{\frac{2mV_0}{\hbar^2} - k^2}$$

is independent of a. The crossing points give the value of k (and γ) for the solutions. The latter is the function of a circle and it is now easier to see that there is a finite number of solutions, as the circle will overlap only a certain number of tangent or cotangent lines, see figure below.



It is now clear there is always at least one solution as no matter how small the circle, i.e. no matter how small V_0 is, there will still be at least one crossing point. Also, as a goes smaller, the spacing along the k axis of the tangent and cotangent lines gets bigger, so the solutions are more widely spaced out.

The most striking thing about the finite square well solutions are that the wavefunction is not zero outside the potential well. This region is forbidden for a classical particle with energy $E < V_0$, as it never has enough energy to climb the potential wall and get outside. However, even though its energy is less than V_0, in the quantum case, there is still some probability that it will be found outside. It seems it must have "negative kinetic energy" is such regions, although how we interpret this is not at all clear. Clearly, if we make a measurement of the position and find the particle is outside the well, then it must at that point have at least an energy equal to V_0, i.e. larger than previously. Hence its energy must be increased by the act of measuring its presence. The only possible source of this "extra" energy is the measuring device itself. This effect that a measurement even in principle disturbs the particle is very basic to quantum mechanics and will come up many times.

6 Comparison with the infinite square well

The equations fixing the value of k and hence E for the even parity states were

$$\gamma = k\tan(ka) \qquad \text{and} \qquad \gamma^2 + k^2 = \frac{2mV_0}{\hbar^2}$$

This means

$$\frac{2mV_0}{\hbar^2} = k^2\tan^2(ka) + k^2 = k^2\left[\frac{\sin^2(ka)}{\cos^2(ka)} + 1\right] = \frac{k^2}{\cos^2(ka)}$$

so

$$\cos(ka) = \pm\frac{\hbar k}{\sqrt{2mV_0}}$$

This has to be used with care as we introduced extra (spurious) solutions in taking the square, so only values of k for which $\tan(ka)$ is positive are valid. However, in this form, it is clear that the infinite square well case corresponds to the right hand side being zero, for which $\cos(ka) = 0$ as we found previously.

Note, for a very shallow well, there will always be at least one bound state. For a deep well, the lowest energy states are close to the energies of the lowest states of the infinite square well. For a very deep well, the wavefunctions decay very rapidly in the classically forbidden regions. In the limit of the infinite well, these become discontinuities in du/dx, which justifies our previous treatment of this case.

Second Year Quantum Mechanics - Lecture 6
The potential step and barrier

Paul Dauncey, 21 Oct 2011

1 Introduction

So far we have only looked in detail at bound states. We will now look at potentials which do not result in bound states but have some interesting properties.

2 The potential step for low energy

The first we will consider is a potential step, where the potential is constant in each of the two regions $x < 0$ and $x > 0$ but has a different value in each. Take $V = 0$ for $x < 0$ and $V = V_0 > 0$ for $x > 0$. Consider a beam of particles incident from the left. There are two cases; $E < V_0$ and $E > V_0$.

For the case $E < V_0$, then classically the particles would all bounce off the barrier and go back along the x axis. In QM, we expect oscillatory solutions for $x < 0$ and exponential solutions for $x > 0$. The relevant equations are

$$\frac{d^2u_L}{dx^2} + \frac{2mE}{\hbar^2}u = 0, \qquad \frac{d^2u_R}{dx^2} - \frac{2m(V_0 - E)}{\hbar^2}u = 0$$

for which the solutions are most conveniently written as

$$u_L = Ae^{ikx} + Be^{-ikx}, \qquad u_R = Ce^{-\gamma x} + De^{\gamma x}$$

where

$$k = \frac{\sqrt{2mE}}{\hbar}, \qquad \gamma = \frac{\sqrt{2m(V_0 - E)}}{\hbar}$$

When coupled to the usual time-dependent energy (phase) factors in the full TDSE solution, then the e^{ikx} terms give right-going waves and the e^{-ikx} terms give left-going waves. Physically, the term Ae^{ikx} is the incoming wave and Be^{-ikx} is a reflected wave. The terms for $x > 0$ are exponential falloff and rise but to keep the wavefunction from becoming infinite as $x \to \infty$, we must set $D = 0$.

The derivatives of the wavefunctions are

$$\frac{du_L}{dx} = ikAe^{ikx} - ikBe^{-ikx}, \qquad \frac{du_R}{dx} = -\gamma Ce^{-\gamma x}$$

Hence, matching the wavefunction and its derivative at the boundary $x = 0$ gives

$$A + B = C, \qquad ikA - ikB = -\gamma C = -\gamma A - \gamma B$$

Hence

$$(k - i\gamma)A = (k + i\gamma)B$$

so

$$B = \left(\frac{k - i\gamma}{k + i\gamma}\right)A$$

Hence

$$C = A + B = \left[1 + \left(\frac{k - i\gamma}{k + i\gamma}\right)\right] A = \left(\frac{k + i\gamma + k - i\gamma}{k + i\gamma}\right) A = \left(\frac{2k}{k + i\gamma}\right) A$$

The probability density in the incoming wave is

$$|Ae^{i(kx-\omega t)}|^2 = |A|^2$$

while the probability density in the reflected wave is

$$|Be^{-i(kx+\omega t)}|^2 = |B|^2 = \left|\frac{k - i\gamma}{k + i\gamma}\right|^2 |A|^2 = \frac{k - i\gamma}{k + i\gamma}\frac{k + i\gamma}{k - i\gamma}|A|^2 = |A|^2$$

Hence, the amount of the reflected wave $|B|^2$ is the same as the incident wave $|A|^2$, i.e. there is total reflection. The wavefunction for $x > 0$ does not represent a travelling wave so there is no transmission of particles out to $x \rightarrow \infty$. This is the same result as we expect classically. However, note that although the particles do not move out to large x, there is some probability of finding the particle just to the right of the barrier.

3 The potential step for high energy

The second case has an energy $E > V_0$, where we would expect oscillations on both sides of the barrier. Classically, we would expect all the particles to pass over the barrier. We have wavefunctions in the two regions

$$u_L = Ae^{ik_L x} + Be^{-ik_L x}, \qquad u_R = Ce^{ik_R x} + De^{-ik_R x}$$

where

$$k_L = \frac{\sqrt{2mE}}{\hbar}, \qquad k_R = \frac{\sqrt{2m(E - V_0)}}{\hbar}$$

Particles can in principle be reflected or transmitted at the barrier, but if we just consider the initial particles as coming from the left, then we must have $D = 0$. We now have

$$\frac{du_L}{dx} = ik_L\left(Ae^{ik_L x} - Be^{-ik_L x}\right), \qquad \frac{du_R}{dx} = ik_R Ce^{ik_R x}$$

Applying the boundary conditions at the boundary $x = 0$, then

$$A + B = C, \qquad k_L(A - B) = k_R C = k_R A + k_R B$$

Hence

$$(k_L - k_R)A = (k_L + k_R)B$$

so

$$B = \left(\frac{k_L - k_R}{k_L + k_R}\right) A$$

Hence

$$C = A + B = \left[1 + \left(\frac{k_L - k_R}{k_L + k_R}\right)\right] A = \left(\frac{k_L + k_R + k_L - k_R}{k_L + k_R}\right) A = \left(\frac{2k_L}{k_L + k_R}\right) A$$

Note, this solution is identical to the previous case if we substitute $k_R = i\gamma$, so we could have done both cases in one go allowing $\sqrt{V_0 - E}$ to be imaginary.

The above equations define B and C in terms of A. We have a (semi) infinitely spread out free particle beam to the left (and also to the right for $E > V_0$) so this wavefunction is not normalised. However, if we only consider ratios, then we can get sensible answers. We want to know how many of the particles are reflected at the barrier, so this is the ratio of incident to reflected, as before. Here, the "reflection coefficient" is

$$R = \frac{|B|^2}{|A|^2} = \frac{(k_L - k_R)^2}{(k_L + k_R)^2} < 1$$

so not all particles are reflected now. Some must make it past the barrier and go out to $x \to \infty$; they have enough energy to have positive kinetic energy on the right. We can calculate the flux of particles for $x > 0$ using C, but care is required as the particle speed is lower.

$$v_L = \frac{p_L}{m} = \frac{\hbar k_L}{m}, \qquad v_R = \frac{p_R}{m} = \frac{\hbar k_R}{m} < v_L$$

If there is a given flux of particles, i.e. number of particles per second, then there must be a higher density of particles where they go slower. In other words, a given density needs to be multiplied by the velocity to get the flux. Hence the fluxes of incoming and transmitted particles are

$$F_L = v_L|A|^2, \qquad F_R = v_R|C|^2$$

The ratio of the incident to outgoing fluxes is the "transmission coefficient" and hence is given by

$$T = \frac{F_R}{F_L} = \frac{|C|^2\hbar k_R/m}{|A|^2\hbar k_L/m} = \frac{|C|^2}{|A|^2}\frac{k_R}{k_L} = \frac{4k_L^2}{(k_L + k_R)^2}\frac{k_R}{k_L} = \frac{4k_L k_R}{(k_L + k_R)^2}$$

We did not have to worry about the k factors for the reflection coefficient because they were the same, as the magnitude of the velocity of the incident and reflected particles is the same. As a check

$$R + T = \frac{k_L^2 + k_R^2 - 2k_L k_R}{(k_L + k_R)^2} + \frac{4k_L k_R}{(k_L + k_R)^2} = \frac{k_L^2 + k_R^2 + 2k_L k_R}{(k_L + k_R)^2} = 1$$

as required.

4 The potential barrier

It is striking that there is some probability for the particle to be to the left of the step for the $E < V_0$ case above. This brings up an interesting idea. What if we put the potential back down to zero at some point later, so that some of the exponential tail reaches a region where it will have oscillatory solutions again? This is the potential barrier. This is a more complicated version of the step and is the main illustration used in almost every quantum mechanics book of the physical implications of the wavefunction being non-zero in classically forbidden regions.

The barrier potential is therefore V_0 for some limited region of space of width w and is zero elsewhere. If $E > V_0$, then we get behaviour similar to the step case; some waves are reflected and some transmitted. In fact, this is mathematically identical to the finite square well for the unbound states; as long as E is greater than the bottom of the potential, the calculation is identical whether V_0 is positive or negative.

The interesting new behaviour happens when $0 < E < V_0$.

$$u_L = Ae^{ikx} + Be^{-ikx}, \quad u_C = Ce^{-\gamma x} + De^{\gamma x}, \quad u_R = Fe^{ikx} + Ge^{-ikx}$$

where

$$k = \frac{\sqrt{2mE}}{\hbar}, \qquad \gamma = \frac{\sqrt{2m(V_0 - E)}}{\hbar}$$

For particles incident from the left, then there is no left-going wave so $G = 0$. Since the D term only holds within the barrier we do not need to worry about it blowing up for large x so cannot put $D = 0$ this time.

The solution of this is mathematically tedious and the details are given in Handout 2. The critical result is that for large γ, i.e. for $E \ll V_0$, then the transmission coefficient goes as

$$T \sim e^{-2\gamma w}$$

The exponential dependence means it is extremely sensitive to the width of the barrier w and so this can be used to measure distances down to the atomic scale, e.g. in a scanning tunnelling microscope.

Second Year Quantum Mechanics - Lecture 7
The simple harmonic oscillator

Paul Dauncey, 25 Oct 2011

1 The simple harmonic oscillator

Unfortunately, this isn't very simple mathematically... However, it is a critical result with huge implications in physics so we need to consider it in some detail.

A simple harmonic potential goes as x^2; explicitly it is usually written as

$$V(x) = \frac{1}{2}m\omega_0^2 x^2$$

There are many situation in which this is at least an approximation. Basically, any system with a potential minimum at x_m, for which $dV/dx|_{x_m} = 0$, can be approximated using a Taylor expansion around the minimum as

$$V(x_m + \delta x) = V(x_m) + \left.\frac{d^2V}{dx^2}\right|_{x_m} \frac{\delta x^2}{2}$$

and so the change of the potential at such a location always goes quadratically; $\Delta V \propto \delta x^2$. Specific examples are the Lennard-Jones potentials between molecules, the motion of atoms in ion traps, and the quantisation of electric and magnetic fields, where it results in photons.

The time-independent Schrödinger equation is

$$-\frac{\hbar^2}{2m}\frac{d^2u}{dx^2} + \frac{1}{2}m\omega_0^2 x^2 u = Eu$$

It is convenient to define scaled variables

$$y = \sqrt{\frac{m\omega_0}{\hbar}}x, \qquad \alpha = \frac{2E}{\hbar\omega_0}$$

Hence

$$\frac{d}{dx} = \sqrt{\frac{m\omega_0}{\hbar}}\frac{d}{dy}$$

so $2/\hbar\omega_0$ times the TISE equation can be written as

$$\frac{d^2u}{dy^2} + (\alpha - y^2)u = 0$$

This is not at all trivial to solve explicitly. I could just write down the answer but let's try to motivate this a little more. We need a solution which produces y^2 when it is differentiated twice. This suggests something like an exponential in $y^2/2$ as this brings down a y each time. It has to go to zero for large y so trying

$$u = He^{-y^2/2}$$

with some constant H, then

$$\frac{du}{dy} = -yHe^{-y^2/2} = -yu$$

$$\frac{d^2u}{dy^2} = -u - y\frac{du}{dy} = -u + (-y)^2u = (y^2 - 1)u$$

Hence we have

$$\frac{d^2u}{dy^2} + (\alpha - y^2)u = y^2u - u + \alpha u - y^2u = (\alpha - 1)u = 0$$

This is satisfied if we choose the energy such that $\alpha = 1$. It turns out this is indeed the ground state solution. However, this is only valid for one particular value of α and so we need to be more general. Instead, let's try

$$u = H(y)e^{-y^2/2}$$

for some more general function of y. For this

$$\frac{du}{dy} = \frac{dH}{dy}e^{-y^2/2} - Hye^{-y^2/2}$$

and

$$\begin{aligned}\frac{d^2u}{dy^2} &= \frac{d^2H}{dy^2}e^{-y^2/2} - \frac{dH}{dy}ye^{-y^2/2} - \frac{dH}{dy}ye^{-y^2/2} - He^{-y^2/2} + Hy^2e^{-y^2/2} \\ &= \frac{d^2H}{dy^2}e^{-y^2/2} - 2\frac{dH}{dy}ye^{-y^2/2} + (y^2 - 1)He^{-y^2/2}\end{aligned}$$

Hence we need to solve

$$\frac{d^2u}{dy^2} + (\alpha - y^2)u = \frac{d^2H}{dy^2}e^{-y^2/2} - 2\frac{dH}{dy}ye^{-y^2/2} + (y^2 - 1)He^{-y^2/2} + (\alpha - y^2)He^{-y^2/2} = 0$$

which means $H(y)$ satisfies

$$\frac{d^2H}{dy^2} - 2y\frac{dH}{dy} + (\alpha - 1)H = 0$$

This is the equivalent of the TISE in terms of H rather than u.

2 Energy values

It is often useful to express solutions as polynomials so we will write

$$H_n(y) = \sum_j^n a_j y^j$$

We must not allow $n \to \infty$ or else the polynomial will beat the exponential and the total $u(y)$ will become infinite as $y \to \infty$. Hence we need n finite, so let's look for a solution which works for large y. As $H_n(y)$ has a leading term y^n, then

$$H_n \approx y^n, \qquad \frac{dH_n}{dy} \approx ny^{n-1}, \qquad \frac{d^2H_n}{dy^2} \approx n(n-1)y^{n-2}$$

Hence for large y

$$n(n-1)y^{n-2} - 2ny^n + (\alpha - 1)y^n = 0$$

For large y, then $y^n \gg y^{n-2}$ so the leading term is

$$-2ny^n + (\alpha - 1)y^n = 0$$

which means

$$-2n + \alpha - 1 = 0$$

or

$$\alpha = 2n + 1$$

This is clearly in agreement with the ground state solution we already found, where $H(y) \propto y^0$ so $n = 0$ and hence from the above, we expect $\alpha = 1$ as we indeed found.

The energies of the states are given by the α values as α was defined as

$$\alpha = \frac{2E}{\hbar\omega_0}$$

Hence, since we found $\alpha = 2n + 1$, then we have energies

$$E = \frac{\alpha}{2}\hbar\omega_0 = \left(n + \frac{1}{2}\right)\hbar\omega_0$$

for integer n. This is a critical result; the energies are equally spaced with an energy difference $\Delta E = \hbar\omega_0$. Note the convention that the quantum number n starts at $n = 0$ here, whereas for both the infinite and finite square well, the ground state has $n = 1$.

3 Hermite polynomials

To see how to get exact solutions, then it is useful to look back at the approximation we made above. We found for large y we did get a solution, so why can't we take our exact solution to be only this largest term, i.e. $H_n = a_n y^n$? The problem is that the second derivative gives an y^{n-2} term, which we could neglect above in the approximation. However, generally, we see that for an exact solution, we will need to include lower powers of y in the solution to cancel off the neglected term in y^{n-2}. Specifically, we would need

$$u = \left(y^n + a y^{n-2}\right) e^{-y^2/2}$$

for some a. However, the new term itself would require lower powers of y to cancel *its* remaining part, etc. This would continue with lower and lower powers and hence in fact generate the polynomial we assumed is the solution. It stops at y^0 or y^1 as the second derivative then gives zero and no further cancelling terms are needed.

Note, each term is two powers of y less and so we will need only even powers if n is even and only odd powers if n is odd. Hence the sequence continues until we get to y^0 (for even n) or y^1 (for odd n). There is a recursion relation between the terms which allows us to work down from y^n to solve for all the a_j in the sequence.

We are not the first people to write down this and the solutions of H have been worked out; they are called "Hermite polynomials". They are labelled $H_n(y)$, where the highest power term is y^n. Each different Hermite polynomial function has a corresponding allowed value of α which is $\alpha = 2n + 1$ as we saw. The first function is H_0 which only has one term, y^0, i.e. a constant. Conventionally, it is

$$H_0 = 1$$

with $\alpha = 1$ and this is the ground state solution we already found. The next function H_1 is also quite simple; it also only has one term, y^1, and conventionally

$$H_1 = 2y$$

with $\alpha = 3$. The next is a little more complicated, it has terms of y^2 and y^0 and the solution is

$$H_2 = 4y^2 - 2$$

with $\alpha = 5$. This can clearly be continued although it is tedious. There are details on the recursion relations and the standard form of the polynomials on Handout 3. Note, we have seen that we either get Hermite polynomials of even powers of y (and hence x) or of odd powers, but not mixtures. Hence, the solutions separate into even parity and odd parity solutions, just as for the bound states of the square wells.

4 Energy values and quantisation

The quantisation has appeared above but it is not clear what caused it. Let's look at the recursion relation in a bit more detail. We found above that the term y^n in the equation for $H(y)$ will give

$$(\alpha - 1)H \propto (\alpha - 1)y^n, \qquad 2y\frac{dH}{dy} \propto 2ny^n, \qquad \frac{d^2H}{dy^2} \propto n(n-1)y^{n-2}$$

If we collect terms with the same power of y, y^j, then the first two terms need $j = n$ while the third needs $j = n-2$ which means $n = j+2$. (Alternatively we could have taken two derivatives of the y^{n+2} term to start with instead.) Hence the equation can be written as

$$\sum_j [a_{j+2}(j+2)(j+1) - 2a_j j + a_j(\alpha - 1)]y^j = 0$$

This can only hold for all y if each factor multiplying y^j is zero. Hence

$$a_{j+2}(j+2)(j+1) - a_j(2j + 1 - \alpha) = 0$$

or

$$a_{j+2} = \frac{2j + 1 - \alpha}{(j+1)(j+2)} a_j$$

There is nothing here preventing the series from continuing for ever, in which case it would contain arbitrarily large powers of y. At that point, it would outweigh the $\exp(-y^2/2)$ term and make the wavefunction grow as $x \to \infty$. Hence, we need to truncate the series after a finite number of terms. For H_n, we stop after y^n, so the only way this will happen is by having this recursion relation give zero for a_{n+2}, which will occur if

$$2n + 1 - \alpha_n = 0$$

or

$$\alpha_n = 2n + 1$$

for which the series stops at y^n, i.e. exactly the values we saw previously. Hence, it is the boundary condition, specifically the requirement that we have the wavefunction go to zero for large x, which forces particular values of α_n and hence quantises the energy. This is just as for the square wells, although significantly more complicated mathematically.

Second Year Quantum Mechanics - Lecture 8
The postulates of quantum mechanics

Paul Dauncey, 27 Oct 2011

1 Introduction

For far, we have used QM to predict the energies of stationary states and the probability distributions for finding particles at some position.

There are many other things we would like a theory of mechanics to be able to do. How do transitions occur between energy levels? What can we say/know/measure about momentum? What about angular momentum? Are there real-world problems for which we can calculate exact solutions?

We will now review what we have seen so far and identify "structural" features. Some of these can be derived from others but some cannot and have to be regarded as the basic postulates of QM. Hence we will try to put the theory on a more formal setting. Comparing this to relativity, where there are only two postulates then we will find QM has (at least) five. There is also some freedom about which ones we chose and which we derive from these. I will follow Rae and label the postulates in the same way as in that textbook.

Postulate 1: For every dynamical system, there exists a wavefunction that is a continuous, square-integrable, single valued function of the parameters of the system and of time and from which all possible predictions about the physical properties of the system can be obtained.

2 Implications

We have seen that $|\psi(x,t)|^2$ is the probability density, meaning that there is some uncertainty about where we would find the particle if we tried to measure its position.

For a classical situation such as a particle in a box (the infinite square well case), then I could put the particle in the box at some time with some initial conditions, i.e. an initial position and momentum. The particle would then bounce backwards and forwards in a regular manner and (in principle) I could predict exactly where it would be at any future time. If I did this but then did not tell you the initial conditions, then what would you be able to predict? Clearly, as far as you were concerned, the particle could have started anywhere. As you don't know its momentum, you don't know how rapidly it will be bouncing backwards and forwards either. Hence, you have no precise predictive power about where the particle would be later. All you could conclude it that it is equally likely to be anywhere within the box and so would say it has a probability density for being measured at any given point, where the probability density would be flat. (Note, this probability density is not the same as the one we found for the quantum energy eigenstates of the infinite square well.)

Classically, this inability to predict the position is because you don't have the full information required; if you did then you could do an exact prediction. Is the same true in quantum mechanics? Postulate 1 says the answer is no; the wavefunction contains *all* the possible information about the state of a system. It allows us to calculate the probability density but there are no other pieces of information which are missing; there are no "hidden variables" in the jargon. Hence, Postulate 1 says that the fact that we can only make probabilistic statements about the position measurement is an intrinsic property of quantum mechanics, not just due to the fact that we didn't determine enough information to start with.

3 Dynamical variables

So far, our handling of position and energy have been completely different. Energy values are given by the solutions of the TISE, while positions are given by a probability distribution, as discussed above. Despite this, it turns out that these two variables are actually formally handled in the same way in quantum mechanics. In fact, all such variables are treated in the same way. These quantities, such as position, momentum, energy, angular momentum, etc, are called "dynamical variables". Classically, these can all take on exact values (in principle) but not in quantum mechanics.

The handling of energy so far is the one which is closer to the general case so let's have a more detailed look at that. We solved the TISE to give "eigenstates" which had an associated energy

$$-\frac{\hbar^2}{2m}\frac{d^2u_n}{dx^2} + Vu_n = \hat{H}u_n = E_nu_n$$

It is clear that if the energy value is to have any meaning at all as an energy, then measuring the energy of a system with a total wavefunction

$$\psi(x,t) = \psi_n(x,t) = u_n(x)e^{-iE_nt/\hbar}$$

would given the value E_n. However, we also know we can make superpositions of the stationary solutions

$$\psi_s(x,t) = \sum_n a_n\psi_n(x,t) = \sum_n a_nu_n(x)e^{-iE_nt/\hbar}$$

for some arbitrary constants a_n, and these are still solutions of the full TDSE. To be clear, ψ_s is *not* a solution of the TISE. Most importantly for this lecture, it contains several different eigenstates u_n with different energy values E_n. What would we get if we measured the energy of such a system? The obvious answer would be that we get the average energy of the states which are added together, weighted appropriately depending on how much of each is included in the superposition. However, as I said back in Lecture 1, quantum mechanics is not obvious. In fact, we always measure the energy to be one of the E_n values, not the average. Which value? This is the critical issue; we get a random energy value chosen from all the possible E_n in the superposition. There is no way to predict which of the E_n will be the measured value for a particular measurement. Even if we make sure the system has exactly the same wavefunction before each energy measurement, then we will still not be able to say we will get the same result each time, even though Postulate 1 states we have all the possible information available. Hence, for a general energy measurement, i.e. when we are not in a particular energy eigenstate, then the result of an energy measurement is random and unpredictable. This is now looking more like the position measurement and in exactly the same way, the unpredictability is intrinsic to the theory and not just due to us being ignorant of all the facts.

4 Collapse of the wavefunction

We have not said what happens to the wavefunction after we do an energy or position measurement. The obvious thing to assume would be that as long as we do our measurement carefully, we could leave the wavefunction unchanged. However, yet again, quantum mechanics does not follow the obvious route and this assumption is not true.

Consider doing a position measurement, which results in the particle being measured at some x value x_a. Let's say we do another measurement of position very quickly after the first, which results in a position x_b. If the system was undisturbed by the first measurement, then we would have the original wavefunction still present. This would say that the probability of measuring the particle some distance away from x_a is still large. Hence, this would imply the second position measurement is completely unconnected with the first measurement. This would be very odd; it would mean "position" is not really measuring anything to do with particle motion. In fact, as we know about relativity, we should be even more sceptical about this. If we made the second measurement after a *very* short time δt, then we know the particle cannot have moved faster than light and so cannot possibly be found at a position $|x_b - x_a| > c\delta t$. However, if the wavefunction is undisturbed and we believe Postulate 1 tells us there is no more information, then we would have to conclude relativity could be violated. Rather than assume Einstein was wrong, we therefore need to drop the assumption that the wavefunction is unchanged. After the first measurement, it must be shaped so that it is peaked very close to x_a, so that x_b will always be close to x_a. In fact, in the limit of doing two measurements immediately one after the other, i.e. $\delta t \to 0$, then the wavefunction following the first must be only non-zero at x_a and zero everywhere else; this shape is called the "Dirac delta function". The wavefunction is said to collapse from the original wavefunction to the delta function because of the measurement at the instant the measurement is made

$$\psi(x) \to \delta(x - x_a)$$

Note, be clear that the random nature of this measurement process is not something which could be overcome by more careful design of the measuring equipment; it is intrinsic to quantum mechanics.

Similarly, we would expect that two consecutive energy measurements, E_a and E_b, would give the same value of energy, i.e. $E_b = E_a$. In fact, as energy is conserved, we would expect that we should get the same value for all later measurements no matter how long we wait before the second measurement. This implies that for the second measurement, we must be sure of getting the same energy value. This means the wavefunction must be such that we can only get E_a for all subsequent measurements and this means we must not be in a superposition any more. Hence, we have to conclude that following an energy measurement, the wavefunction must be in the energy eigenstate corresponding to the measured value. Again, the measurement has forced the wavefunction to collapse from the general superposition to one of the TISE solutions

$$\psi_s = \sum_n a_n \psi_n \to \psi_a$$

Note, in both cases, the collapse is down to a state which will give a definite value of the measured variable so a second measurement gives a consistent value, at least in the limit of being a very short time afterwards. Hence, the collapse depends on which dynamical variable is being measured and on the random outcome of that measurement. It is clear that during this change the wavefunction time dependence cannot be following the Schrödinger equation; the superposition *is* a solution of the full Schrödinger equation so the above change is completely different. Hence, we find measurements give a very different change of the wavefunction with time compared with the Schrödinger equation.

5 Operators

We need to formalise the above process. We said a measurement will give only particular values (at least of energy) and will result in a wavefunction which definitely gives that value for a second measurement. We need to find what those values can be and also what the wavefunction will collapse into. This needs us to consider "operators".

Consider an energy eigenstate u_n which we know satisfies

$$-\frac{\hbar^2}{2m}\frac{d^2 u_n}{dx^2} + V u_n = \hat{H} u_n = E_n u_n$$

where

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

This is trivial mathematically, but $\hat{H}$ is an example of an "operator", which by convention are labelled with the "hat". In this case, $\hat{H}$ is the energy operator, which "operates" on the wavefunction and "reveals" the energy value E_n, as in the equation above. For the particular case of the energy operator, $\hat{H}$ is often also called the "Hamiltonian operator" (or even just the "Hamiltonian" if people are being sloppy although strictly speaking there is a difference). This is because it corresponds to the function in classical mechanics called the Hamiltonian function which was mentioned in Lecture 1.

We will do a lot of work with operators. In some sense, they are a generalisation of functions. With a function $f(x)$, you plug in a value of x and get another value $f(x)$ back. Some functions can be so trivial you don't necessarily think of them as functions, e.g. $f(x) = 2x$ means "multiply by two" and $f(x) = 1$ means "return one". For the case of operators, you plug in a function and get another function back. For instance, d/dx is an operator and so appying this to $f(x)$ would give $g(x) = df/dx$, which is a different function. The first term in $\hat{H}$ is the double derivative and so has a similar effect. Some operators can also be very trivial, e.g. an operator might be simply "multiply by a function $h(x)$" so the result of applying it to $f(x)$ is $g(x) = hf$, which is indeed a different function. The second term of $\hat{H}$ is like this; the V (which we could write as $\hat{V}$ if we wanted to) simply means the result is Vu. Furthermore, "multiply by two" is also a perfectly good operator as it gives back another function $2f$ so the "operator" would be 2. We will tend to add the hat to such simple operators only when needed for emphasis. It can be complicated to keep track of what operators do mathematically as they are more complicated than simple functions. Firstly, always be clear in your mind about which function the operator is operating on; use brackets to remove any ambiguity if necessary. Also, be clear of the order of the operations with more than two operators; $\hat{Q}\hat{R}f$ is not necessarily the same as $\hat{R}\hat{Q}f$ as is easily demonstrated for $\hat{Q} = x$ and $\hat{R} = d/dx$. In several ways, operators have similar properties to matrices. Generally, when working with operators, it is always worth thinking "What would I have to do with a derivative?" as this will get you around lots of potential slip-ups and will always reduce to something simple anyway if the operator is more straightforward.

Second Year Quantum Mechanics - Lecture 9
Hermitian operators

Paul Dauncey, 28 Oct 2011

1 Eigenstates and eigenvalues

We are searching for a way to calculate the states which have definite measurement values, which are what the wavefunction collapses into. The equation

$$\hat{H}u_n(x) = E_n u_n(x)$$

is critical and contains a lot of the important points about the operators of dynamical variables in general. The structure of this equation shows that the operator operates on a wavefunction u (called the "eigenstate") and produces exactly the same wavefunction, but multiplied by a number E, the corresponding "eigenvalue". The prefix "eigen" comes from German, where is means "associated with"; each eigenfunction has an associated eigenvalue. Hence, the eigenstates of an operator are very special functions because they have the property of not being changed in form by the operator. For a general operator, most functions will not do this. However, there is not generally just one such function; we have seen for the energy operator that there are many such eigenstates and each has its own eigenvalue. The operator "digs out" the relevant energy for any eigenstate.

The obvious question is whether we can find equivalent operators for the other dynamical variables. It turns out this is the case; every dynamical variable has an operator associated with it. We need to find the eigenstates and eigenvalues for each operator in order that we can find the possible values resulting from a measurement of the associated variable. It turns out the operators cannot be arbitrary but to make physical sense, they have to be "Hermitian" operators (again due to Charles Hermite). We will discuss what this means next but everything we have so far can be summarised in the second postulate.

Postulate 2: Every dynamical variable may be represented by an Hermitian operator whose eigenvalues represent the possible results of carrying out a measurement of the value of the dynamical variable. Immediately after such a measurement, the wavefunction of the system will be identical to the eigenstate corresponding to the eigenvalue obtained as a result of the measurement.

2 Hermitian operators

The form of all the operators for dynamical observables must be such that the eigenvalues are guaranteed to be real numbers and not complex. Otherwise, we would not know how to interpret the predictions of quantum mechanics when comparing with the real values we measure.

An important class of operators that have real eigenvalues are the Hermitian operators, which are defined though

$$\int_{-\infty}^{\infty} \psi^* \hat{Q}\phi \, dx = \int_{-\infty}^{\infty} (\hat{Q}\psi)^* \phi \, dx$$

where $\hat{Q}$ is an Hermitian operator and ψ and ϕ are any two well-dehaved functions that vanish at infinity. (Note, Rae has a different, but equivalent, definition.) Note, all Hermitian operators have real eigenvalues but not all operators with real eigenvalues are Hermitian. The Hermitian condition is more restrictive that absolutely necessary. However, experimentally, Postulate 2 has been found to be true and no exceptions have ever been found.

We can use the Hermitian definition to get out two critical results about the eigenstates and eigenvalues of Hermitian operators, including the fact that the eigenvalues must be real. Consider the eigenvalue equation

$$\hat{Q}\phi_n = q_n\phi_n$$

Multiplying by the complex conjugate of another eigenstate ϕ_m^* and integrating gives

$$\int_{-\infty}^{\infty} \phi_m^* \hat{Q}\phi_n \, dx = \int_{-\infty}^{\infty} \phi_m^* q_n \phi_n \, dx = q_n \int_{-\infty}^{\infty} \phi_m^* \phi_n \, dx$$

Using the definition of a Hermitian operator, then the LHS is

$$\int_{-\infty}^{\infty} \phi_m^* \hat{Q}\phi_n \, dx = \int_{-\infty}^{\infty} (\hat{Q}\phi_m)^* \phi_n \, dx = \int_{-\infty}^{\infty} (q_m\phi_m)^* \phi_n \, dx = q_m^* \int_{-\infty}^{\infty} \phi_m^* \phi_n \, dx$$

so

$$q_m^* \int_{-\infty}^{\infty} \phi_m^* \phi_n \, dx = q_n \int_{-\infty}^{\infty} \phi_m^* \phi_n \, dx$$

Take the case of $m = n$; for this, then the integrals are

$$\int_{-\infty}^{\infty} \phi_n^* \phi_n \, dx = 1$$

by normalisation so we have

$$q_n^* = q_n$$

which proves q_n is real, as promised. Hence, by using Hermitian operators, we can guarantee that we will only ever predict real values for measurement outcomes.

Alternatively, by taking $m \neq n$, then we have

$$(q_m - q_n) \int_{-\infty}^{\infty} \phi_m^* \phi_n \, dx = 0$$

Assuming the eigenvalues are not the same, i.e. $q_m \neq q_n$, then we must have

$$\int \phi_m^* \phi_n \, dx = 0$$

This property means that the different eigenstates of Hermitian operators are said to be orthogonal. (The exception occurs if they have the same eigenvalues, which gets us to degeneracy, which is tackled later in the course when we consider quantum mechanics in 3D.) We have proved this in the general case without having to write down any explicit wavefunctions. It holds for all the wavefunctions we have seen so far, even though we didn't check this at the time. We can combine these two results as

$$\int \phi_m^* \phi_n \, dx = \delta_{mn}$$

where δ_{mn} is the Kronecker delta function, and is defined to be

$$\begin{aligned} \delta_{mn} &= 1 \text{ when } m = n, \\ &= 0 \text{ when } m \neq n \end{aligned}$$

This integral equation is the *orthonormality* condition and eigenstates of Hermitian operators are said to be orthonormal.

3 Momentum

The obvious question is what are the equivalent operators for the other dynamical variables; in particular for momentum? Let's think about the structure of $\hat{H}$ a little to get some clues. It is clear that of the two terms in $\hat{H}$, only V is related to the potential so it alone must be the potential operator, $\hat{V} = V(x)$. (Here, the operator is just a function, which you should think of as a very simple type of operator, just doing a multiplication.) Since the total energy E is the sum of the potential and kinetic energy T, then it is clear the kinetic energy operator must be the first term

$$\hat{T} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}$$

Classically, we can write

$$T = \frac{1}{2}mv^2 = \frac{p^2}{2m}$$

Of these, the second seems to have the right mass dependence so we will try this one. Let's make the association with a momentum operator $\hat{p}$ through

$$\frac{\hat{p}^2}{2m} = \hat{T} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}$$

so that

$$\hat{p}^2 = -\hbar^2\frac{d^2}{dx^2}$$

which has a solution

$$\hat{p} = -i\hbar\frac{d}{dx}$$

where the overall minus sign is not defined from the above but is required to be consistent with the de Broglie relations, below. If this is the momentum operator, then momentum eigenstates and eigenvalues should obey

$$\hat{p}\phi_n = -i\hbar\frac{d\phi_n}{dx} = p_n\phi_n$$

The solution of this equation is given by

$$\frac{d\phi_n}{\phi_n} = i\frac{p_n}{\hbar}dx$$

for which

$$\ln\phi_n = i\frac{p_n}{\hbar}x + C$$

or

$$\phi_n = (e^C)e^{ip_nx/\hbar} = Ae^{ip_nx/\hbar}$$

so the constant of integration C is seen to be related to the normalisation factor. Note, since $p = \hbar k$ for wavenumber k, then this is

$$\phi_n = Ae^{ik_nx}$$

Hence, the eigenstates of the momentum operator have a very simple form. Note, there are no restrictions on the allowed momentum values from the operator eigenvalue equation so any value of momentum is possible. Hence, the eigenvalues form a continuous spectrum, not a discrete set, and the label n is usually dropped, so we will just write p in future. (This is similar to the energy eigenvalues for an unbound state.) Note, we dropped $\hat{p} = +i\hbar d/dx$ because it would give waves going in the opposite direction to p, i.e. $p = -\hbar k$, so positive p values would have waves moving in the negative direction.

Let's check that $\hat{p}$ is Hermitian. Consider

$$\int_{-\infty}^{\infty} \psi^* \hat{p} \phi \, dx = \int_{-\infty}^{\infty} \psi^* \left(-i\hbar \frac{d\phi}{dx} \right) dx$$

Integrating by parts gives

$$-i\hbar \int_{-\infty}^{\infty} \psi^* \frac{d\phi}{dx} \, dx = -i\hbar \left([\psi^* \phi]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{d\psi^*}{dx} \phi \, dx \right)$$

The first term is zero as the functions go to zero at infinity and the second term gives

$$i\hbar \int_{-\infty}^{\infty} \frac{d\psi^*}{dx} \phi \, dx = \int_{-\infty}^{\infty} \left(i\hbar \frac{d\psi^*}{dx} \right) \phi \, dx = \int_{-\infty}^{\infty} \left(-i\hbar \frac{d\psi}{dx} \right)^* \phi \, dx = \int_{-\infty}^{\infty} (\hat{p}\psi)^* \phi \, dx$$

Hence, this shows $\hat{p}$ obeys the equation defining an Hermitian operator and so it is indeed Hermitian. All the other operators we have met so far are also Hermitian.

Finally, note that the operation for the potential energy "operator" is actually just "multiple by V", rather than something like "take a derivative in x" as for momentum. This suggests that dynamical variables which can be expressed as a function of x, $f(x)$, may be represented by the operator "multiply by $f(x)$". In particular, the operator for position x may be taken to be "multiply by x", i.e. $\hat{x} = x$. These developments are summarised in the third postulate.

Postulate 3: The operators representing the position and momentum of a particle are x and $-i\hbar\, d/dx$ respectively. Operators representing other dynamical variables bear the same function relation to these as do the corresponding classical quantities to the classical position and momentum variables.

As you know, momentum is constant for a free particle, i.e. when $V = 0$. In QM, it turns out that this means the free particle wavefunctions from the Schrödinger equation are momentum eigenstates as well as energy eigenstates. This means these are states with a definite value of energy *and* momentum. Hence, free particles have a wavefunction

$$\psi = Ae^{ikx}$$

for which the momentum is

$$\hat{p}\psi = -i\hbar \frac{d\psi}{dx} = -i\hbar(ik\psi) = \hbar k \psi = p\psi$$

and the energy is

$$\hat{H}\psi = \hat{T}\psi = -\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} = -\frac{\hbar^2}{2m}(-k^2\psi) = \frac{(\hbar k)^2}{2m}\psi = \frac{p^2}{2m}\psi$$

as would be expected classically.

Note also that, if $V \neq 0$, i.e. if there is a force, then the momentum eigenstates are no longer the same as the energy eigenstates. Hence, an energy eigenstate will not have definite momentum. You can think of this classically that if there is a force, the momentum is not constant with time. Also, this is in fact similar to the situation with position; we know energy eigenstates do not have definite position either.

Also note that the Hamiltonian operator is the only one which is dependent on the physical system under consideration. This is true for this due to its dependence on V so its eigenstates and eigenvalues vary from system to system. In contrast, all other operators are "simpler" in as far as they are identical in any system and so we only need to solve for the eigenstates and eigenvalues once and these are valid in any situation.

Second Year Quantum Mechanics - Lecture 10
Complete orthonormal sets

Paul Dauncey, 1 Nov 2011

1 Introduction

We have seen that measurements will only result in eigenvalues of the appropriate operator. However, there are two major issues which we still need to look into. Firstly, we saw for the case of an energy measurement that a wavefunction satisfying the Schrödinger equation could be considered to be in a superposition

$$\psi(x,t) = \sum_n a_n u_n(x) e^{-iE_n t/\hbar}$$

where the $u_n(x)$ are the solutions of the TISE, i.e. the eigenvalues of the energy operator $\hat{H}$. In this case, the possible outcomes of an energy measurement are the E_n corresponding to any of the $u_n(x)$ included in this sum. However, how does this generalise for the other dynamical variables? How do we find the possible values which could be the result for a particular measurement?

Secondly, we have not said anything about the probability of getting each of the values, e.g. for getting each of the E_n. Are they all equally likely or is there some weighting? To answer these questions, we need to find some properties of the eigenstates.

2 Completeness

You will have been learning about Fourier series (and transforms), which work because any periodic function, or function of limited range $0 \le x \le T$, can be written as a sum of sines and cosines

$$f(x) = \frac{1}{2}c_0 + \sum_{n=1}^{\infty} c_n \cos nkx + s_n \sin nkx$$

where $k = 2\pi/T$. The values of c_n and s_n are found using

$$c_0 = \frac{2}{T}\int_0^T f(x)\,dx, \qquad c_n = \frac{2}{T}\int_0^T \cos nkx f(x)\,dx, \qquad s_n = \frac{2}{T}\int_0^T \sin nkx f(x)\,dx$$

These are called "overlap integrals" and reflect the extend to which $f(x)$ looks like $\sin nkx$ or $\cos nkx$. Fourier transforms are the same thing but generalised to an infinite range.

It is possible to do Fourier series using complex functions also. In this case, we use

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n e^{inkx} + a_n^* e^{-inkx}$$

The a_n^* for the second term in the sum is required to ensure $f(x)$ is real but if the function being expanded is complex, then we can just write it as a sum from $n = -\infty$ to ∞ and only the first term is needed. Again, the overlap integral gives

$$a_n = \frac{2}{T}\int_0^T e^{-inkx} f(x)\,dx,$$

Because *any* periodic function can be represented in this way, then the sine and cosine functions, together with the "constant value" function 1/2 (multiplying the a_0 or c_0) are said to form a *complete set*.

Consider the infinite square well. Its energy eigenstates are just sines and cosines within the region $-a < x < a$. There is no constant term in this case as it is not allowed by the boundary conditions. Hence, our knowledge of Fourier series shows any arbitrary wavefunction for the infinite square well that obeys the boundary conditions can be written as a sum over the energy eigenstates.

$$\psi = \sum_{n=1}^{\infty} a_n u_n$$

We now take a huge jump and say this is generally possible using all sets of Hermitian eigenstates; this has been proved by Hilbert. This means the Fourier series complete set is not the only possible complete set; it fact there are many other complete sets which can be used and every operator will give us such a set if we solve for its eigenstates.

Going back to the Fourier case, then the reason the overlap integral gives the coefficients depends on the functions being orthogonal and normalised; specifically

$$\frac{2}{T}\int_0^T \cos nkx \cos mkx \, dx = \delta_{mn}, \qquad \frac{2}{T}\int_0^T \sin nkx \sin mkx \, dx = \delta_{mn}, \qquad \int_0^T \cos nkx \sin mkx \, dx = 0$$

where δ_{mn} is the Kronecker delta function. This means multiplying by a sine or cosine and integrating picks out one term in the Fourier series. We previously found the orthonormality relation also holds for any Hermitian operator eigenstates

$$\int \phi_m^* \phi_n \, dx = \delta_{mn}$$

Hence, we are saying the eigenstates of Hermitian operators are complete *and* orthonormal; in the jargon they form a “complete orthonormal set”, sometimes written as CONS.

Note, not all complete sets are orthogonal. E.g. a Taylor series expanded around $x = 0$ (i.e. a Maclaurin series) is

$$f(x) = f(0) + \left.\frac{df}{dx}\right|_0 x + \frac{1}{2}\left.\frac{d^2 f}{dx^2}\right|_0 x^2 + \dots$$

Since the derivatives are all evaluated at $x = 0$, then they are constant values so this is effectively

$$f(x) = a_0 + a_1 x + a_2 x^2 + \dots = \sum_n a_n x^n$$

This expansion can be done for any function with well-behaved derivatives and so is complete in this sense, but the expansion functions, which are the polynomials x^n, are not orthogonal to each other.

Given a general Hermitian operator $\hat{Q}$ with eigenvalues q_n corresponding to eigenstates ϕ_n, then any arbitrary wavefunction can be written as an (in principle infinite) superposition of the eigenstates

$$\psi = \sum_n a_n \phi_n$$

In the Fourier series, a_n (or c_n and s_n) are called weighting factors and the coefficient plays a similar role in QM. Note, however, that as the ϕ_n and ψ are complex in general, then so must the a_n. How do we find the values of the a_n? Since

$$\psi = \sum_m a_m \phi_m$$

then multipling by a particular eigenstate ϕ_n^* gives

$$\int \phi_n^* \psi \, dx = \sum_m a_m \int \phi_n^* \phi_m \, dx$$

However, we know the integral on the right hand side as we found it earlier; it is

$$\int \phi_n^* \phi_m \, dx = \delta_{nm}$$

Hence we find

$$\sum_m a_m \int \phi_n^* \phi_m \, dx = \sum_m a_m \delta_{nm} = a_1 \delta_{n1} + a_2 \delta_{n2} + a_3 \delta_{n3} + \ldots = a_n$$

and so

$$a_n = \int \phi_n^* \psi \, dx$$

which is the general equivalent of the Fourier overlap integral, with ϕ_n^* playing the roles of the sines and cosines, i.e. the expansion functions, and ψ the role of $f(x)$, i.e. the function being expanded. The expansion coefficients a_n are often called the *amplitudes*.

Given this, then it is clear how we generalise the statement about measuring the energy eigenvalues E_n. If we are going to do a measurement of a variable on a system with a wavefunction ψ, we need to expand ψ in terms of the eigenstates of the measurement variable. The eigenstates included in the resulting expansion tell us which eigenvalues can result from the measurement.

3 Measurement probabilities

Finally, to complete the picture, we now need to know how likely each value is. Although it is not obvious from the above result, the a_n amplitude values calculated using the overlap integral result in a correctly normalised ψ wavefunction, assuming the ϕ_n are normalised of course. There is a important property of the amplitudes which arises from the normalisation of ψ. Since we need

$$\int |\psi|^2 \, dx = \int \psi^* \psi \, dx = 1$$

then we must have

$$\int (\sum_m a_m^* \phi_m^*)(\sum_n a_n \phi_n) \, dx = \sum_{m,n} a_m^* a_n \int \phi_m^* \phi_n \, dx = 1$$

Again using the orthonormality of the eigenstates

$$\int \phi_m^* \phi_n \, dx = \delta_{mn}$$

then this becomes

$$\sum_{m,n} a_m^* a_n \delta_{mn} = \sum_n a_n^* a_n = \sum_n |a_n|^2 = 1$$

This is very suggestive; we have a wavefunction ψ which can be decomposed into a sum of eigenstates ϕ_n of an operator with weights a_n. If we do a measurement of the dynamical variable of the operator, we will get one of the eigenvalues of those states. The a_n are larger if the eigenstate looks like the wavefunction and they basically tell us how much of the eigenstate is in the wavefunction. Hence, it would seem reasonable to assume that $|a_n|^2$ is the *probability* that the result would be the eigenvalue of the state ϕ_n. The fact that the sum of the $|a_n|^2$ gives one would then mean we have a probability of one of measuring any of the possible eigenvalues, i.e. we will get *some* value out of the total list of possible values. This is of course exactly equivalent to the reason we need the normalisation $\int |\psi|^2 \, dx = 1$ in the first place.

Postulate 4: When a measurement of a dynamic variable represented by a Hermitian operator is carried out on a system whose wavefunction is ψ, then the probability of the result being equal to a particular eigenvalue q_n will be $|a_n|^2$, where $\psi = \sum_n a_n\phi_n$ and the ϕ_n are the eigenstates of the operator corresponding to the q_n.

Beware; the overlap integral gives a_n, while the probability is given by $|a_n|^2$; this is easy to forget. Note, the amplitudes are so-called because in the case of photons, the probability density is proportional to the light intensity which goes as the square of the electromagnetic fields, i.e. the square of the field amplitudes. Hence, squaring the amplitudes here to get the probability is analogous.

4 Vectors and functions

There are mathematical similarities between functions and vectors which help get an intuitive feel for what is happening with these expansions. A normal 3D vector $\boldsymbol{a}$ has components a_i which could be written $a(i)$. A function is written $\psi(x)$ and so the similarity is that vectors are like functions with a discrete and finite set of integer input values; alternatively, functions are like vectors with an infinite number of components. Taking this similarity further, then a dot product is

$$\boldsymbol{a}.\boldsymbol{b} = \sum_i a(i)b(i)$$

which in the limit of taking integer i to be continuous x becomes

$$\int \psi(x)\phi(x)\,dx$$

The functions we actually use in QM are complex so it actually becomes $\int \psi^*(x)\phi(x)\,dx$ but otherwise these expressions are similar.

The vector $\boldsymbol{a}$ can be written in terms of some axis vectors as

$$\boldsymbol{a} = a_1\boldsymbol{i} + a_2\boldsymbol{j} + a_3\boldsymbol{k} = a_1\hat{\boldsymbol{n}}_1 + a_2\hat{\boldsymbol{n}}_2 + a_3\hat{\boldsymbol{n}}_3 = \sum_i a_i\hat{\boldsymbol{n}}_i$$

Any vector can be written in this way. The axis vectors are unit vectors, i.e. $\hat{\boldsymbol{n}}_i^2 = 1$, and are orthogonal to each other, i.e. $\hat{\boldsymbol{n}}_i.\hat{\boldsymbol{n}}_j = 0$ for $i \neq j$. Hence, it is seen that the $\hat{\boldsymbol{n}}_i$ form a CONS for 3D vectors. This means that the components are given by

$$a_i = \boldsymbol{a}.\hat{\boldsymbol{n}}_i$$

This is analogous to

$$\psi(x) = \sum_i a_i\phi_i(x)$$

with

$$a_i = \int \phi_i^*(x)\psi(x)\,dx$$

You can think of the eigenstates as forming a base set of “axes”, and we use those to resolve any wavefunction into components along those axes. Different Hermitian operators will give different sets of eigenstates, and hence axes, but any set form a CONS; the different CONS correspond to rotations of the axes.

Second Year Quantum Mechanics - Lecture 11
Time dependence

Paul Dauncey, 3 Nov 2011

1 Introduction

We have seen that a measurement of a variable will give back one of the eigenvalues of the Hermitian operator associated with the variable. The actual eigenvalue is randomly chosen from the possible ones included in the wavefunction expansion in terms of the operator eigenstates. These could be energy eigenstates, meaning solutions of the TISE $u_n(x)$, or momentum eigenstates, $e^{ipx/\hbar}$, for example. However, in all cases, the eigenstates are a function of x but not of t so the collapse does not tell us anything about the future time dependence.

We now want to think about the time dependence of the wavefunction in general, which requires us to consider both the cases with and without measurements. We will also need to start thinking a bit about conserved quantities as these are important in mechanics.

2 Time dependence without measurements

The final postulate is one which we have effectively been assuming all along.

Postulate 5: Between measurements, the development of the wavefunction with time is governed by the time-dependent Schrödinger equation.

This has some direct implications. When we make a measurement, the wavefunction collapses to some eigenstate. Following the measurement, Postulate 5 says it then propagates according to the TDSE. This means the eigenstate is the initial condition for the solution to the TDSE. Hence, the future value for the wavefunction is completely determined by the measured eigenvalue (and hence eigenstate).

The term "mechanics" in both classical mechanics and quantum mechanics means that we are trying to describe how the system changes with time. In the case of classical mechanics, then things are very simple (in principle). We want to know $x(t)$ and $p(t)$. Once we have set up the system with the initial conditions, i.e. $x(0)$ and $p(0)$, then for any later time, $x(t)$ and $p(t)$ are well determined, whether we "observed" the system or not.

In quantum mechanics, following a measurement, we set up the system with a definite set of initial conditions $\psi(x, t = 0)$, which is two (real and imaginary parts) functions rather than two variables, but the principle is the same. If we make no measurements, the value of $\psi(x, t)$ at any future time is also well determined. Classical mechanics, and quantum mechanics without any measurements, are both said to be *deterministic*. If this was the only difference between classical and quantum mechanics then they would still be quite different theories (and the world would look *very* different!) but if intelligent beings could still exist, there would not be a philosophical problem in moving conceptually from one to the other.

However, unlike the classical case, if we try to observe a quantum system, i.e. make measurements of it, then its motion will be affected in a random, unpredictable way. We can only then make predictions of the probability of things happening. Such systems are not deterministic, but are said to be *probabilistic*. It is this part of quantum mechanics which causes all the big questions; what causes nature to choose one eigenvalue rather than the other, how can two identical experiments give different outcomes, etc.

It would clearly be nice to do experiments to check directly the deterministic part of quantum mechanics by itself but of course this is impossible; any experiment cannot tell us anything unless we look at the system under consideration, at which point we are doing a measurement and we disturb it.

There is a second implication of Postulate 5. If the future wavefunction only depends on the initial condition, i.e. on the eigenstate of the measured eigenvalue, then it cannot have any dependence on what the wavefunction was *before* the collapse. Hence, any wavefunction which could have given the measured eigenvalue may have existed before, but this information is completely lost. Hence, the previous wavefunction is in general "forgotten", along with any properties it may have had. This also means that, unlike the classical case, we cannot put negative times into our solution to get the wavefunction before the measurement.

3 TDSE solutions

There is a special application of the expansion for the energy case. Say we have a wavefunction at $t = 0$, $\psi(x, 0)$. This is a function of x so, as always, we can do an expansion in terms of energy eigenstates (as they are a complete set) at that time

$$\psi(x, 0) = \sum_n a_n u_n(x)$$

However, we know that, if there are no further measurements, a wavefunction which is initially an energy eigenstate evolves as $u_n(x)e^{-iE_nt/\hbar}$, so we might expect the time development of the general wavefunction for any later time is given by

$$\psi(x, t) = \sum_n a_n u_n(x) e^{-iE_nt/\hbar}$$

We already saw this superposition is a solution of the complete time-dependent Schrödinger equation and so this is indeed the correct result. Hence, the above is the general solution for the time dependence of a wavefunction if it is not measured. Sometimes it is possible to write a solution to the Schrödinger equation in a closed form without using an expansion. If so, this is often very convenient, but it will always turn out to be just another way of expressing the expansion above; there is only one solution of the Schrödinger equation given an initial condition $\psi(x, 0)$. Hence, while we have been trying to show that all dynamical variables can be handled the same way, this particular result shows the energy operator is special as it controls the time dependence. This is not really suprising as the Hamiltonian operator $\hat{H}$ gives the time dependence in the Schrödinger equation

$$i\hbar \frac{\partial \psi}{\partial t} = \hat{H}\psi$$

Note, if the wavefunction at $t = 0$ was created by making an energy measurement which gave E_n, then it results in the wavefunction collapsing into an energy state u_n. Hence, by definition, the expansion has only one term and the wavefunction at a later time is then

$$\psi(x, t) = u_n(x) e^{-iE_nt/\hbar}$$

You will see this is then one of our separable solutions and so is a stationary state, meaning an energy eigenstate, so the probability density never changes with time and a future energy measurement will always give back E_n.

You could alternatively think of the general case in a different way. Since the wavefunction evolves with time, then if we did the expansion at some later time, we would get different values of the expansion coefficients. Hence, the coefficients can be considered as a function of time. We can simply rewrite the above as

$$\psi(x, t) = \sum_n \left(a_n e^{-iE_nt/\hbar}\right) u_n(x) = \sum_n a_n(t) u_n(x)$$

which shows this explicitly. You can think of the time evolution either way; both are valid. Of course, in terms of the probabilities for measuring the energy values

$$P_n(t) = |a_n(t)|^2 = |a_n(0)|^2 e^{iE_nt/\hbar} e^{-iE_nt/\hbar} = |a_n(0)|^2 = P_n(0)$$

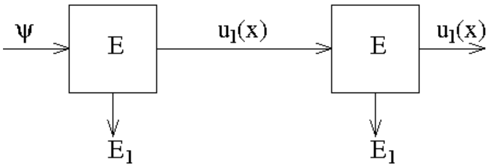
and so are constant, i.e. the probabilities to get each of the E_n values never change, as the "amount" of each of the a_n, i.e. their modulus, never changes.

4 Effect of measurements

Let's think through an example set of measurements and see how this all fits together. We will first consider the case of a series of rapid measurements, so that the time dependence can be neglected. To pick a specific example, we want the system to be in a well-defined initial state, so we will make an energy measurement and only if it gives a particular energy, then we use it. Otherwise we disturb the system again and try to measure the energy again. Picking systems (e.g. atoms in a beam) which give a particular measurement value like this is called "preparing" a state. Anyway, following this, we know the system is in a definite energy state, e.g.

$$\psi(x, 0) = u_1(x)$$

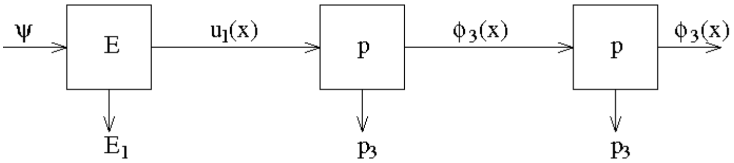
If the energy is now rapidly measured a second time, then because there is only one energy eigenstate in the wavefunction, we know the measurement outcome must be E_1 again. This is equivalent to saying the "sum" of states has only one term and so its amplitude must satisfy $|a_1|^2 = 1$, i.e. have a probability of one. These two measurements can be represented by the diagram below.



Let's now change the second measurement to be a momentum measurement. This will have the state $u_1(x)$ as its input so to find the probabilities of the various possible momentum eigenvalues, then we need to expand in momentum eigenstates

$$u_1(x) = \sum_p a_p \phi_p(x)$$

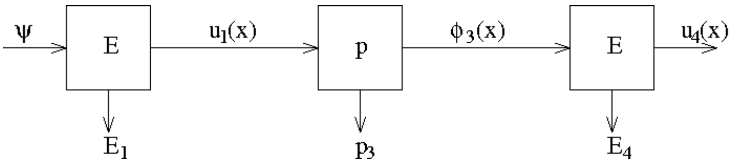
(For simplicity the states are written as if discrete, although as we saw, the p values are actually continuous.) The probability of getting any p is $|a_p|^2$ and in general many of these could be non-zero so the outcome of the measurement will be uncertain. Say the measurement happens to give p_3; the outgoing wavefunction will then have collapsed to $\phi_3(x)$. If another measurement of p is then done, as shown in the diagram below, then as it is now a momentum eigenstate, we will definitely get p_3.



However, consider yet another case, shown below, where we measure the energy, not the momentum, as the third measurement. The measurement of momentum is as before but now the input to the third measurement, $\phi_3(x)$, is not an energy eigenstate. Hence, we need to expand this function in terms of energy eigenstates

$$\phi_3(x) = \sum_n b_n u_n(x)$$

where the $|b_n(x)|^2$ give the probability of each E_n. In general, there will be several non-zero amplitudes here and hence an uncertainty in the energy measurement outcome. This shows an initially measured energy of E_1 may not be replicated later, and a different value, e.g. E_4, might be found in the last measurement.

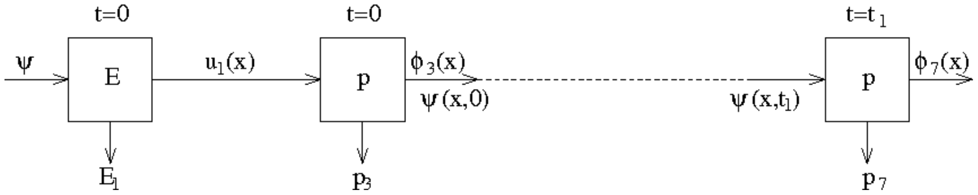


Does this mean energy is not conserved? No; what it means is that in measuring the momentum, the measuring device (whatever it is) has interfered with the system, changing its state. In doing this, it can exchange energy (and indeed momentum) with the system and so we would not necessarily expect the energy afterwards to be the same as before. In other words, the energy of the system is certainly not conserved but it is not an isolated system so we don't expect it to be; only the total energy, including that of the measurement apparatus, etc, is conserved. Hence, what we think of as conservation of variables is more subtle within quantum mechanics than classical mechanics.

It is worth considering when we would be sure of the last measurement result being E_1 again. This can only happen if the input to this measurement is $u_1(x)$ so in terms of this diagram, it would require $\phi_3(x)$ and $u_1(x)$ to have the same functional form.

5 Measurements with time dependence

Finally, we should consider the extra complication of allowing the time dependence, i.e. having measurements spaced apart in time. Consider just the repeated momentum measurements discussed above, but with a non-negligible time between them.



The wavefunction out of the first measurement is $\psi(x, 0) = \phi_3(x)$ and we need to allow this to propagate in time until the second momentum measurement at time t_1. Hence, we need to express $\psi(x, 0)$ in terms of energy eigenstates and then add the time dependence, as discussed at the beginning of the lecture. Since we already had

$$\psi(x, 0) = \phi_3(x) = \sum_n b_n u_n(x)$$

then

$$\psi(x,t_1)=\sum_n b_n u_n(x) e^{-iE_nt_1/\hbar}$$

This is the input to the second momentum measurement so it needs to be decomposed in terms of momentum eigenstates

$$\psi(x,t_1)=\sum_p c_p\phi_p(x)$$

The amplitudes can be found by the standard method of doing the overlap integral, which here gives

$$c_p=\int \phi_p(x)^*\psi(x,t_1)\,dx=\sum_n b_n e^{-iE_nt_1/\hbar}\int \phi_p(x)^* u_n(x)\,dx$$

It is clear that calculating the c_p will be messy but the critical issue is that in general, several of the c_p will be non-zero, so the momentum measurement is uncertain, even though it follows a previous momentum measurement with no other measurements in between.

In the same way as before, this arises because the momentum eigenstates and the energy eigenstates are not the same. If they were, all the sums above would contain only one term and all measurements would be certain. It is not impossible for two operators to have the same eigenstates and in fact we already saw an example. The momentum eigenstates are $e^{ipx/\hbar}=e^{ikx}$ which are also the energy eigenstates for a free particle, i.e. when there is no force. This is exactly the case when we would expect classically that the momentum would be conserved so there is clearly a similar result in QM. Hence we find there is a connection between conservation laws and shared eigenstates, which we will come back to in future lectures.

Second Year Quantum Mechanics - Lecture 12
Commutators and anticommutators

Paul Dauncey, 4 Nov 2011

1 Introduction

Postulate 3 says the classical combination of dynamical variables gives the correct combination of operators in quantum mechanics. We have seen some examples; classically the kinetic energy is

$$T = \frac{p^2}{2m}$$

and we have said the quantum kinetic energy operator is

$$\hat{T} = \frac{\hat{p}^2}{2m}$$

The potential energy is another example, although not so obvious. E.g. for the simple harmonic oscillator, the classical potential energy is

$$V = \frac{1}{2} m\omega_0^2 x^2$$

Postulate 3 says we should then write

$$\hat{V} = \frac{1}{2} m\omega_0^2 \hat{x}^2$$

but because $\hat{x} = x$ then this is simply

$$\hat{V} = \frac{1}{2} m\omega_0^2 x^2 = V(x)$$

so we end up with the operator being the original function.

2 Commutators

Unfortunately things are not always so simple. Consider the two operators $\hat{x}$ and $\hat{p}$. Let's say we want to construct the operator corresponding to the classical product xp. In classical mechanics this is unambiguous, but not in quantum mechanics. These operators need a function to work on so we shall write this explicitly as ψ to start with. Consider the action of $\hat{x}\hat{p}$ on some wavefunction

$$\hat{x}\hat{p}\psi = x\left(-i\hbar\frac{d\psi}{dx}\right) = -i\hbar x\frac{d\psi}{dx}$$

The operators in the other order give

$$\hat{p}\hat{x}\psi = -i\hbar\frac{d}{dx}(x\psi) = -i\hbar\psi - i\hbar x\frac{d\psi}{dx}$$

and so the two combinations are indeed not the same. This means that when trying to find quantum operators for products of dynamical variables, it is not always clear which to use.

The difference of these two combinations plays a critical role in quantum mechanics and is

$$\hat{x}\hat{p}\psi - \hat{p}\hat{x}\psi = i\hbar\psi$$

The actual function used ψ is arbitrary here, i.e. the result is independent of ψ, so we can write a general equation for the operators themselves

$$\hat{x}\hat{p} - \hat{p}\hat{x} = i\hbar$$

provided that we always remember that a function on the right is implied. This combination occurs so often in quantum mechanics that it has its own notation

$$[\hat{x}, \hat{p}] = \hat{x}\hat{p} - \hat{p}\hat{x}, \qquad \text{so} \qquad [\hat{x}, \hat{p}] = i\hbar$$

and this is called the *commutator* of the two operators. It has meaning only when operating on a function. This definition of a commutator is not limited to $\hat{x}$ and $\hat{p}$; the same form can be used for any two operators. The commutators have some basic properties. For a general pair of operators $\hat{Q}$ and $\hat{R}$, then

$$[\hat{Q}, \hat{R}] = \hat{Q}\hat{R} - \hat{R}\hat{Q} = -[\hat{R}, \hat{Q}]$$

The commutator of an operator with itself is

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

and similarly

$$[\hat{Q}, \hat{Q}^n] = \hat{Q}\hat{Q}^n - \hat{Q}^n\hat{Q} = \hat{Q}^{n+1} - \hat{Q}^{n+1} = 0$$

You should compare these results with those of the classical Poisson brackets we discussed in Lecture 1. You will find the properties of the commutators are very similar and this is a sign of a connection back to classical mechanics.

Of course, the value of the commutator is not always $i\hbar$; some operators have a commutator which is zero, in which case they are said to *commute.* Some commutators result in expressions involving other operators. The ones which do give $i\hbar$ are specially paired and are called *conjugate variables.*

We will see there is a deep connection between the commutation relation

$$[\hat{x}, \hat{p}] = i\hbar$$

and the Uncertainty Principle in a few lectures. The commutation relation above is so fundamental to QM that it is often taken as one of the postulates instead of assuming the form of $\hat{x}$ and $\hat{p}$, as we did for Postulate 3. Note that the right hand side has a size of $\hbar$. Therefore, in large (classical) systems, this is effectively zero and so the operators approximately commute. Hence, the effects of quantum mechanics are limited to small systems.

3 Hermitian combinations

A commutator is itself an operator so we need to know its properties. One important property of any commutator is that is it not Hermitian, but has a property called anti-Hermitian. Consider the product of any two Hermitian operators $\hat{Q}$ and $\hat{R}$ within an integral

$$\int \phi^* \hat{Q}\hat{R}\psi \, dx = \int (\hat{Q}\phi)^* \hat{R}\psi \, dx = \int (\hat{R}\hat{Q}\phi)^* \psi \, dx$$

Clearly, also

$$\int \phi^* \hat{R}\hat{Q}\psi \, dx = \int (\hat{Q}\hat{R}\phi)^* \psi \, dx$$

Hence, subtracting

$$\int \phi^* (\hat{Q}\hat{R} - \hat{R}\hat{Q})\psi \, dx = \int (\hat{R}\hat{Q} - \hat{Q}\hat{R}\phi)^* \psi \, dx$$

so

$$\int \phi^*[\hat{Q},\hat{R}]\psi\,dx = \int([\hat{R},\hat{Q}]\phi)^*\psi\,dx = -\int([\hat{Q},\hat{R}]\phi)^*\psi\,dx$$

A Hermitian operator would have both sides equal to each other; if they are the negative of each other, as here, then they are called anti-Hermitian. Hence, any commutator is anti-Hermitian and so the commutator cannot represent a dynamical variable. However, it is trivial to convert an anti-Hermitian operator to a Hermitian one. Multiplying by i gives

$$\int \phi^* i[\hat{Q},\hat{R}]\psi\,dx = \int -i([\hat{Q},\hat{R}]\phi)^*\psi\,dx = \int(i[\hat{Q},\hat{R}]\phi)^*\psi\,dx$$

and so i times the commutator is indeed Hermitian and hence can correspond to an observable. This is general; any anti-Hermitian operators can be made Hermitian by multiplying by i.

Although not so fundamental to QM, for completeness we need to mention anticommutators. The definition of an anticommutator is

$$\{\hat{Q},\hat{R}\} = \hat{Q}\hat{R} + \hat{R}\hat{Q}$$

This is sometimes a source of confusion; note the commutator has a negative sign and the anticommutator has a positive sign. Note, again trivally

$$\{\hat{Q},\hat{R}\} = \{\hat{R},\hat{Q}\}$$

and

$$\{\hat{Q},\hat{Q}\} = 2\hat{Q}^2$$

As might be expected, if the anticommutator for two operators is zero, then they are said to *anticommute*. Note, don't confuse the anticommutator with the Poisson bracket; although they both have the same bracket symbol, the former works with operators and the latter with classical variables.

The usefulness of the anticommutator is that it is also Hermitian. By adding, rather than subtracting, the two equations at the start of the section, then we find

$$\int \phi^*(\hat{Q}\hat{R} + \hat{R}\hat{Q})\psi\,dx = \int(\hat{R}\hat{Q} + \hat{Q}\hat{R}\phi)^*\psi\,dx$$

which means

$$\int \phi^*\{\hat{Q},\hat{R}\}\psi\,dx = \int(\{\hat{R},\hat{Q}\}\phi)^*\psi\,dx = \int(\{\hat{Q},\hat{R}\}\phi)^*\psi\,dx$$

which shows the anticommutator is indeed a Hermitian variable and hence can also correspond to an observable.

Note, if the operators happen to commute, i.e. $[\hat{Q},\hat{R}] = 0$, then this means $\hat{Q}\hat{R} = \hat{R}\hat{Q}$ so in this case $\{\hat{Q},\hat{R}\} = 2\hat{Q}\hat{R}$. Conversely if they happen to anticommute, i.e. $\{\hat{Q},\hat{R}\} = 0$, then this means $\hat{Q}\hat{R} = -\hat{R}\hat{Q}$ so in this case $i[\hat{Q},\hat{R}] = 2i\hat{Q}\hat{R}$. Hence, overall, if the operator corresponding to the product of two classical dynamical variables is needed, then it could be $i[\hat{Q},\hat{R}]/2$ or $\{\hat{Q},\hat{R}\}/2$ or even some combination of the two. Only if one of these is zero is the answer unambiguous. This shows it is not always trivial to go from a classical to the equivalent quantum system and only by comparing with experiment can we determine which combination is correct.

Finally, note it is actually straightforward to write the product of two operators in terms of commutators and anticommutators. Consider the product of $\hat{Q}$ and $\hat{R}$

$$\begin{aligned}\hat{Q}\hat{R} &= \frac{1}{2}\left(\hat{Q}\hat{R} + \hat{Q}\hat{R}\right) = \frac{1}{2}\left(\hat{Q}\hat{R} + \hat{R}\hat{Q} + \hat{Q}\hat{R} - \hat{R}\hat{Q}\right) \\ &= \frac{1}{2}\left(\{\hat{Q},\hat{R}\} + [\hat{Q},\hat{R}]\right) = \frac{1}{2}\{\hat{Q},\hat{R}\} + \frac{1}{2}[\hat{Q},\hat{R}]\end{aligned}$$

Note that the above is the commutator, which is anti-Hermitian, so this makes it again clear that the product $\hat{Q}\hat{R}$ is not itself Hermitian unless the commutator is zero. The above can also be written as

$$\hat{Q}\hat{R} \;=\; \left(\frac{1}{2}\{\hat{Q},\hat{R}\}\right) - i\left(\frac{i}{2}[\hat{Q},\hat{R}]\right)$$

which is a form we will use later when discussing the Uncertainty Principle.

4 Further examples

We know the commutator of $\hat{x}$ and $\hat{p}$. We have been using another operator, the Hamiltonian operator $\hat{H}$, so let's see what the commutators of $\hat{x}$ and $\hat{p}$ with $\hat{H}$ are. We know $\hat{H} = \hat{T} + \hat{V}$ so

$$[\hat{H},\hat{p}] = [\hat{T}+\hat{V},\hat{p}] = [\hat{T},\hat{p}] + [\hat{V},\hat{p}]$$

The first is easy

$$[\hat{T},\hat{p}] = \frac{1}{2m}[\hat{p}^2,\hat{p}] = 0$$

as we showed $[\hat{Q}^n,\hat{Q}] = 0$ above. The second requires an explicit calculation so let's put the wavefunction back in

$$[\hat{V},\hat{p}]\psi = V\left(-i\hbar\frac{d\psi}{dx}\right) + i\hbar\frac{d}{dx}(V\psi) = -i\hbar V\frac{d\psi}{dx} + i\hbar\frac{dV}{dx}\psi + i\hbar V\frac{d\psi}{dx} = i\hbar\frac{dV}{dx}\psi$$

and hence

$$[\hat{V},\hat{p}] = i\hbar\frac{dV}{dx}$$

This holds for any function $V(x)$ so note in particular the case of $V(x) = x$ which then reduces to the $[\hat{x},\hat{p}] = i\hbar$ result. Hence, overall

$$[\hat{H},\hat{p}] = i\hbar\frac{dV}{dx}$$

Commuting $\hat{x}$ with $\hat{H}$, then again

$$[\hat{H},\hat{x}] = [\hat{T},\hat{x}] + [\hat{V},\hat{x}]$$

This time, the second term is easy

$$[\hat{V},\hat{x}] = Vx - xV = 0$$

as they are both simply functions. The first term is

$$[\hat{T},\hat{x}] = \frac{1}{2m}[\hat{p}^2,\hat{x}]$$

This could be calculated using the explicit derivative for $\hat{p}^2$ similarly to before, but there is a neater way to find it, using

$$[\hat{p}^2,\hat{x}] = \hat{p}\hat{p}\hat{x} - \hat{x}\hat{p}\hat{p} = \hat{p}\hat{p}\hat{x} - \hat{p}\hat{x}\hat{p} + \hat{p}\hat{x}\hat{p} - \hat{x}\hat{p}\hat{p} = -\hat{p}[\hat{x},\hat{p}] - [\hat{x},\hat{p}]\hat{p} = -2i\hbar\hat{p}$$

so overall

$$[\hat{H},\hat{x}] = -\frac{i\hbar}{m}\hat{p}$$

In summary, $\hat{x}$ and $\hat{p}$ do not commute and $\hat{x}$ and $\hat{H}$ do not commute. Also, $\hat{p}$ and $\hat{H}$ do not commute *unless* $dV/dx = 0$. We know this is the case when momentum is classically conserved so this gives us a first hint that classical conservation is related to commutators. We also saw in this case the momentum and energy eigenstates were the same, i.e. proportional to $e^{ipx/\hbar} = e^{ikx}$. These are all related and we will see this more explicitly in the next few lectures.

Second Year Quantum Mechanics - Lecture 13
Compatibility, expectation values, Dirac notation

Paul Dauncey, 8 Nov 2011

1 Introduction

We will look at the definitions of compatibility and the expectation values. Also, we will introduce an alternative notation for QM.

2 Compatibility

A few lectures ago, we saw that, because a measurement of energy would result in an energy eigenstate which is a stationary state, then a second measurement would always give the same energy value. However, if a momentum measurement was done in between, then this no longer holds and the second energy measurement could result in other energy values. At a basic level, this means no quantity (except the normalisation) is ever conserved absolutely in quantum mechanics.

However, we would like to understand which sets of variables we can measure without one messing up the other. If we want to have a definite outcome of a measurement, we have to be in the eigenstate of the variable being measured before the measurement. Hence, for two subsequent measurements of different variables, the only way the second can give a definite outcome is if the operators have the same eigenstate.

Two observables are called *compatible* if their operators have a common set of eigenstates. One example we already saw was the eigenstates of the free particle Hamiltonian $\hat{H}_{\text{free}}$

$$\psi = Ae^{ikx} = Ae^{ipx/\hbar}$$

which were also eigenstates of $\hat{p}$. Note, although operators can share eigenstates, the eigenvalues are not (necessarily) the same; in this example these states have energy $E = \hbar^2k^2/2m$ and momentum $p = \hbar k$.

If we have have two general operators $\hat{Q}$ and $\hat{R}$ which share eigenstates, then

$$\hat{Q}\phi_n = q_n\phi_n, \qquad \hat{R}\phi_n = r_n\phi_n$$

For any general wavefunction ψ, we can expand in these eigenstates

$$\psi = \sum_n a_n\phi_n$$

If we make a measurement of the dynamical variable associated with $\hat{Q}$, then we (randomly) get a particular eigenvalue, say q_3. If we then immediately afterwards make a measurement of the variable associated with $\hat{R}$, then what will be the outcome? The wavefunction has collapsed into the state ϕ_3 and so we will get the eigenvalue r_3 with certainty. The wavefunction remains in ϕ_3 after the second measurement. This follows from Postulate 2. Hence, this is the condition required to be able to do two sets of measurements and be sure of getting the same answer each time. This means if we restrict ourselves to only making rapid measurements of variables which share the same eigenstates, we will always get predictable answers.

The question is how can we identify compatible operators? For compatible $\hat{Q}$ and $\hat{R}$, consider $\hat{Q}\hat{R}$ on a general wavefunction

$$\hat{Q}\hat{R}\psi = \sum_n a_n\hat{Q}\hat{R}\phi_n = \sum_n a_n\hat{Q}r_n\phi_n = \sum_n a_nr_n\hat{Q}\phi_n = \sum_n a_nr_nq_n\phi_n$$

Clearly, with the operators inverted

$$\hat{R}\hat{Q}\psi = \sum_n a_n q_n r_n \phi_n = \sum_n a_n r_n q_n \phi_n$$

since q_n and r_n are simply real numbers. Hence, subtracting

$$[\hat{Q}, \hat{R}]\psi = 0$$

i.e. if $\hat{Q}$ and $\hat{R}$ are compatible, they must commute. Note, we already saw in the last lecture that $\hat{H}$ and $\hat{p}$ commute in the particular case of $dV/dx = 0$ i.e. for a free particle, and this is why they can have common eigenstates as shown above.

Hence, commutation is a good test of whether operators are compatible or not. However, we also need to check the inverse, i.e. if two operators commute, are they compatible? Consider $\hat{Q}$ with eigenstates ϕ_n which are not (necessarily) eigenstates of $\hat{R}$. If $\hat{Q}$ and $\hat{R}$ commute, then for any of the $\hat{Q}$ eigenstates

$$\begin{aligned} [\hat{Q}, \hat{R}]\phi_n &= 0 \\ \hat{Q}\hat{R}\phi_n - \hat{R}\hat{Q}\phi_n &= 0 \\ \hat{Q}(\hat{R}\phi_n) &= q_n(\hat{R}\phi_n) \end{aligned}$$

This is the eigenstate equation for $\hat{Q}$ again. This means that $\phi'_n = \hat{R}\phi_n$ is also an eigenstate of $\hat{Q}$. However, it is an eigenstate with the same eigenvalue, q_n, as ϕ_n. When discussing orthonormality, we assumed each q_n was different, i.e. two different eigenstates would not have the same eigenvalue. This is usually a safe assumption when working in 1D. If we assume this again now, then it means ϕ'_n and ϕ_n cannot really be different eigenstates but ϕ'_n has to be just proportional to ϕ_n

$$\phi'_n \propto \phi_n \qquad \text{so} \qquad \phi'_n = r_n \phi_n$$

Hence

$$\hat{R}\phi_n = r_n \phi_n$$

which means $\hat{R}$ has the same eigenstates as $\hat{Q}$, i.e. they are compatible. So we have found that if operators commute *and* they have unique eigenvalues, then they are compatible. Hence, commuting operators are a good test of compatibility, at least in 1D.

What is the flip side? If $\hat{Q}$ and $\hat{R}$ don't commute, then they cannot have the same eigenstates, or else the calculation above would apply. Two immediate subsequent measurements do not result in a definite outcome for the second. In this case, a measurement of the variable associated with $\hat{Q}$ would give some value q_m and collapse the wavefunction to ϕ_m. However, this is not now an eigenstate of $\hat{R}$ so the outcome of the subsequent measurement is uncertain; we cannot predict in advance which eigenvalue of $\hat{R}$ we will get. Such operators are said to be *incompatible*; one example of this is $\hat{x}$ and $\hat{p}$ which do not share eigenstates and do not commute.

Warning: do not confuse compatibility and the Uncertainty Principle (UP), which we will shall discuss later. Compatibility concerns measurements of two variables, one following the other, and whether the second gives a definite outcome. The UP concerns the spread of measurement outcomes of a single measurement of either of two variables (but not both) on a given initial state, where the initial state needs to be prepared in the same way before each measurement. It has nothing to say about what a second measurement gives.

3 Expectation values

Imagine a beam of atoms such that the energy ground state and the first excited state are both populated to a significant degree. Hence, the total state is a mixture of the two and we would write the wavefunction as

$$\psi = a_1 u_1 + a_2 u_2$$

To make it simple, let's take the case of an equal mixture, when $a_1 = a_2 = 1/\sqrt{2}$. We then measure the energy of each atom one at a time. Sometimes the wavefunction collapses to u_1 and we measure E_1 and sometimes it collapses to u_2 and we measure E_2. The first case happens with probability $|a_1|^2 = 1/2$ and the second with probability $|a_2|^2 = 1/2$. We might want to know what the average energy which we will measure is. It is obviously $(E_1 + E_2)/2$ but more systematically, it is straightforward using Postulate 4

$$\langle E \rangle = |a_1|^2 E_1 + |a_2|^2 E_2 = \frac{1}{2} E_1 + \frac{1}{2} E_2$$

Since $|a_1|^2 + |a_2|^2 = 1$, then this average has to be somewhere between E_1 and E_2, as would be expected, and for the equal mixture it is clearly half way between.

This can be broadened to the general superposition case; for

$$\psi = \sum_n a_n \phi_n$$

then the average value of the measurement of the dynamical variable associated with $\hat{Q}$ will be given by

$$\langle Q \rangle = \sum_n |a_n|^2 q_n$$

The value $\langle Q \rangle$ is called the *expectation value* of the variable. As we will see in a future lecture, this turns out to correspond to the classical limit; measurements give an average close to the classical variable value.

Obviously, if there is only one term in the superposition, ϕ_n, then $\langle Q \rangle = q_n$. However, note in general that $\langle Q \rangle$ is the average value, *not* the most probable value; in the energy example it does not correspond to either E_1 or E_2.

Going back to the atomic beam example, then consider the expression

$$\int_{-\infty}^{\infty} \psi^* \hat{H} \psi \, dx$$

where, as usual, $\hat{H}$ is the Hamiltonian (energy) operator. Substituting in the expansion of ψ, then we get

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2}}(u_1^* + u_2^*) \hat{H} \frac{1}{\sqrt{2}}(u_1 + u_2) \, dx = \frac{1}{2} \int_{-\infty}^{\infty} (u_1^* + u_2^*)(E_1 u_1 + E_2 u_2) \, dx$$

using the fact that the u_n are eigenstates of $\hat{H}$. This then gives

$$\frac{1}{2} \left(E_1 \int_{-\infty}^{\infty} u_1^* u_1 \, dx + E_2 \int_{-\infty}^{\infty} u_2^* u_2 \, dx + E_1 \int_{-\infty}^{\infty} u_2^* u_1 \, dx + E_2 \int_{-\infty}^{\infty} u_1^* u_2 \, dx \right)$$

Using the orthonormality condition

$$\int_{-\infty}^{\infty} u_n^* u_m \, dx = \delta_{nm}$$

then we get

$$\int_{-\infty}^{\infty} \psi^* \hat{H} \psi \, dx = \frac{1}{2}(E_1 + E_2) = \langle H \rangle$$

Hence, this gives us directly the value of the energy expectation value without having to explicitly expand the wavefunction ψ into the energy eigenstates. This generalises to

$$\begin{aligned} \int_{-\infty}^{\infty} \psi^* \hat{Q} \psi \, dx &= \int_{-\infty}^{\infty} (\sum_m a_m \phi_m)^* \hat{Q} (\sum_n a_n \phi_n) \, dx = \int_{-\infty}^{\infty} (\sum_m a_m^* \phi_m^*)(\sum_n a_n q_n \phi_n) \, dx \\ &= \sum_{m,n} a_m^* a_n q_n \int_{-\infty}^{\infty} \phi_m^* \phi_n \, dx = \sum_{m,n} a_m^* a_n q_n \delta_{mn} = \sum_n |a_n|^2 q_n = \langle Q \rangle \end{aligned}$$

Hence, we have two ways of calculating the expectation value

$$\begin{aligned} \langle Q \rangle &= \int_{-\infty}^{\infty} \psi^* \hat{Q} \psi \, dx \\ &= \sum_n |a_n|^2 q_n \end{aligned}$$

and which to use is purely convenience. Note that the result is just a number, as the operators and functions have been integrated. Hence, an expectation value can be manipulated like any other simple number.

Finally, note the expression for the expectation value is similar to the integral involved in the definition of the Hermitian operator, but for the special case of the two functions being the same. Indeed

$$\langle Q \rangle = \int_{-\infty}^{\infty} \psi^* \hat{Q} \psi \, dx = \int_{-\infty}^{\infty} (\hat{Q}\psi)^* \psi \, dx$$

The two integrals above are complex conjugates of each other and so as they are equal, then expectation value must be real. This has to be true as it corresponds to something which is measurable. The fact that the expectation value must be real is also apparent from the other expression for it

$$\langle Q \rangle = \sum_n |a_n|^2 q_n$$

We showed all eigenvalues of Hermitian operators are real, so all the q_n are real, and $|a_n|^2$ is always real, so the expectation value is a sum of real numbers and hence is itself real.

4 Dirac notation

As this is used in more advanced courses, then we should consider Dirac "bra-ket" notation. For now, consider it purely as a shorthand, alternative notation for the sorts of expressions we use in QM, although it is really a more general way of representing states. Specifically, using the analogy of a function being a vector with an infinite number of components, then in writing $\psi(x)$, we are writing the equivalent of $A(i)$, i.e. the specific components of a vector. The more general way of writing a vector, $\underline{A}$, corresponds to Dirac notation.

The "ket" $|\rangle$ is equivalent to a wavefunction; specifically

$$\psi(x) \equiv |\psi\rangle$$

An eigenstate can therefore be written as

$$\phi_n(x) \equiv |\phi_n\rangle \qquad \text{or} \qquad |n\rangle$$

where the second form can be used if there is no ambiguity about which operator's eigenstates are being considered. Hence, for example, our usual wavefunction expansion could be written in two equivalent ways

$$\psi(x) = \sum_n a_n \phi_n(x), \qquad |\psi\rangle = \sum_n a_n |\phi_n\rangle$$

The "bra" $\langle|$ represents the complex conjugate of the wavefunction; specifically

$$\psi^*(x) \equiv \langle\psi|$$

so

$$\phi_n^*(x) \equiv \langle\phi_n| \qquad \text{or} \qquad \langle n|$$

Note, the complex conjugate is part of the definition of the bra; we do not write $\langle\psi^*|$.

Kets and bras can be combined; the combination of the bra and ket (to make a "bra-ket") is defined so it includes doing the integral. Hence, for example, the orthonormality condition can also be written in two equivalent ways

$$\int_{-\infty}^{\infty} \phi_m^* \phi_n \, dx = \delta_{mn}, \qquad \langle\phi_m|\phi_n\rangle = \delta_{mn}$$

A final example is the expression for the expectation value; this can be written as

$$\int_{-\infty}^{\infty} \psi^* \hat{Q} \psi \, dx, \qquad \langle\psi|\hat{Q}|\psi\rangle \equiv \langle Q\rangle$$

where the second form shows one reason why the pointed brackets are used to denote the expectation value.

A summary table of the notation is

$$\begin{array}{lrcl}
\text{ket} & \psi(x) & \equiv & |\psi\rangle \\
 & \phi_n(x) & \equiv & |\phi_n\rangle \qquad \text{or} \qquad |n\rangle \\
 & \psi(x) = \sum_n a_n \phi_n(x) & \equiv & |\psi\rangle = \sum_n a_n |\phi_n\rangle \\
\text{bra} & \psi^*(x) & \equiv & \langle\psi| \\
 & \phi_n^*(x) & \equiv & \langle\phi_n| \qquad \text{or} \qquad \langle n| \\
\text{bra} - \text{ket} & \int_{-\infty}^{\infty} \phi_m^*(x)\phi_n(x)\, dx = \delta_{mn} & \equiv & \langle\phi_m|\phi_n\rangle = \delta_{mn} \quad \text{or} \quad \langle m|n\rangle = \delta_{mn} \\
 & \int_{-\infty}^{\infty} \phi^*(x)\hat{Q}\psi(x)\, dx & \equiv & \langle\phi|\hat{Q}\psi\rangle \qquad \text{or} \qquad \langle\phi|\hat{Q}|\psi\rangle \\
 & \int_{-\infty}^{\infty} \psi^*(x)\hat{Q}\psi(x)\, dx = \langle Q\rangle & \equiv & \langle\psi|\hat{Q}|\psi\rangle = \langle Q\rangle
\end{array}$$

Second Year Quantum Mechanics - Lecture 14
The Ehrenfest theorem

Paul Dauncey, 10 Nov 2011

1 The correspondence principle

Clearly, the world around us looks classical, even though we believe it is fundamentally governed by quantum mechanics. This means that quantum mechanics must give classical mechanics as a limiting case. This is called the *correspondence principle* and is stated in several different ways. The limiting case can be considered as in the limit of large objects, meaning much bigger than any wavelengths of the particles. It can also be considered as the limit of $\hbar \to 0$; e.g. in this limit, the quantised spacing between energy levels $\Delta E_n = E_{n+1} - E_n$ goes to zero so we get no quantisation but a continuous range of possible energy values. Another way to consider the limit is that it holds when all quantum numbers are large, e.g. $n \to \infty$ so that the energy is much bigger than the spacing; $E_n \gg \Delta E_n$ so that the quantisation effects are negligible. In this lecture, we want to quantify this idea.

2 Expectation value time dependence

The situation we will consider is the following. We will prepare a large number of atoms to be in the same initial state, as discussed in a previous lecture. We leave them for a time t and then measure some quantity Q for each atom. We find the average and then ask, how does that average change with t, the time the atoms are left?

We have seen that the average value of a measurement of some operator $\hat{Q}$ is given by the expection value

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

If ψ is not an energy eigenstate, i.e. not a stationary state, then it will have some non-trivial time dependence and hence the expectation value of $\hat{Q}$ will in general change with time. Assuming the operator $\hat{Q}$ has no explicit time dependence, consider

$$\frac{d\langle Q \rangle}{dt} = \frac{d}{dt} \int \psi^* \hat{Q} \psi \, dx = \int \frac{\partial}{\partial t}(\psi^* \hat{Q} \psi) \, dx = \int \frac{\partial \psi^*}{\partial t} \hat{Q} \psi \, dx + \int \psi^* \hat{Q} \frac{\partial \psi}{\partial t} \, dx + \int \psi^* \frac{\partial \hat{Q}}{\partial t} \psi \, dx$$

where we drop the last term as the operators we will look at here have no explicit time dependence. Note that since $\langle Q \rangle$ has all the spatial dependence integrated out, then it is purely a function of t, i.e. $\langle Q \rangle = f(t)$, and so the derivative on the left hand side is a total derivative. However, once taken into the integral, then the time derivatives need to become partial derivatives, as $\psi = \psi(x, t)$.

We know the time dependence of the wavefunction; Postulate 5 says it is given by the Schrödinger equation

$$i\hbar \frac{\partial \psi}{\partial t} = \hat{H} \psi$$

so that

$$\frac{\partial \psi}{\partial t} = -\frac{i}{\hbar} \hat{H} \psi$$

The complex conjugate of this equation is

$$\frac{\partial \psi^*}{\partial t} = \frac{i}{\hbar} (\hat{H} \psi)^*$$

This means

$$\frac{d\langle Q\rangle}{dt} = \int \frac{i}{\hbar}(\hat{H}\psi)^*\hat{Q}\psi\,dx - \int \psi^*\hat{Q}\frac{i}{\hbar}\hat{H}\psi\,dx = \frac{i}{\hbar}\left[\int (\hat{H}\psi)^*\hat{Q}\psi\,dx - \int \psi^*\hat{Q}\hat{H}\psi\,dx\right]$$

Using the fact that the Hamiltonian is Hermitian, then this gives

$$\frac{d\langle Q\rangle}{dt} = \frac{i}{\hbar}\left[\int \psi^*\hat{H}\hat{Q}\psi\,dx - \int \psi^*\hat{Q}\hat{H}\psi\,dx\right] = \frac{i}{\hbar}\int \psi^*[\hat{H},\hat{Q}]\psi\,dx = \frac{1}{\hbar}\langle i[\hat{H},\hat{Q}]\rangle$$

Hence, the time dependence of the expectation value is given by $i/\hbar$ times the expectation value of the commutator with the Hamiltonian. We previously showed that i times the commutator is a Hermitian operator, so the time derivative is an observable, i.e. a dynamical variable.

Note, this equation holds for any operator $\hat{Q}$ but we must always take the commutator with the Hamiltonian $\hat{H}$. Again, this shows the Hamiltonian is special with regard to time dependence.

You should compare this form for the time dependence with the classical equation given in Lecture 1 using the Hamilton formulation

$$\frac{dQ}{dt} = \{Q,H\} = -\{H,Q\}$$

where the terms on the right are the Poisson brackets. (Don't confuse these with the anticommutator.) One of the ways to state the correspondence principle is that replacing a commutator by $i\hbar$ times the Poisson bracket gives the classical result and vice versa, e.g.

$$[\hat{H},\hat{Q}] \leftrightarrow i\hbar\{H,Q\}$$

which means

$$\frac{i}{\hbar}[\hat{H},\hat{Q}] \leftrightarrow -\{H,Q\}$$

corresponding to what we saw above.

3 The Ehrenfest theorem

Let's check this for the operators we know so far. We previously found the three commutators

$$[\hat{x},\hat{p}] = i\hbar, \qquad [\hat{H},\hat{x}] = -\frac{i\hbar\hat{p}}{m}, \qquad [\hat{H},\hat{p}] = i\hbar\frac{dV}{dx}$$

Note, these are all proportional to $\hbar$ so in the limit of $\hbar \to 0$, they all go to zero. Having all operators commute is yet another way of considering the correspondence principle. However, the thing of interest here is, what do the second and third commutators give for the expectation value time dependence? Since

$$[\hat{H},\hat{x}] = -\frac{i\hbar\hat{p}}{m}$$

then the time dependence of the expectation value of x is

$$\frac{d\langle x\rangle}{dt} = \frac{1}{\hbar}\left\langle i\left(-\frac{i\hbar\hat{p}}{m}\right)\right\rangle = \frac{\langle\hat{p}\rangle}{m}$$

Also, since

$$[\hat{H},\hat{p}] = i\hbar\frac{dV}{dx}$$

then

$$\frac{d\langle p\rangle}{dt} = \frac{1}{\hbar}\left\langle i\left(i\hbar\frac{dV}{dx}\right)\right\rangle = -\left\langle\frac{dV}{dx}\right\rangle$$

These two equations should look familiar; they are very similar to the classical Hamiltonian equations we studied in Lecture 1. Specifically, they are equivalent to the classical equations $p = mv$ and $F = dp/dt = ma = -dV/dx$. This shows that the average measurements of position and momentum obey the classical equations and is a quantitative expression of the correspondence principle. If you compare the above to the classical Poisson brackets for these pairs of variables, you will see they agree with the above correspondence relation.

This turns out to be general and was first stated by Ehrenfest: *The equations of motion for the expectations values of observables are the same as the equations of motion for their classical counterparts.* His theorem therefore states it is not just $\hat{x}$ and $\hat{p}$ which have expectation values with time dependences like the classical ones; it is every operator which does this.

4 Conserved variables

Given the Ehrenfest theorem, we can now see what a classically conserved variable actually corresponds to in quantum mechanics. Since the theorem says the classical equations of motion are the same as the expectation value equations of motion, then clearly a classically conserved quantity must also have a constant $\langle Q\rangle$ and so $d\langle Q\rangle/dt = 0$. It is clear that this must correspond to

$$\langle[\hat{H}, \hat{Q}]\rangle = 0$$

There are two conditions under which this occurs. The "trivial" one is when ψ happens to be an energy eigenstate. Explicitly, $\psi = u_n$ and so

$$\begin{aligned}\langle[\hat{H}, \hat{Q}]\rangle &= \int u_n^*(\hat{H}\hat{Q} - \hat{Q}\hat{H})u_n\,dx = \int u_n^*\hat{H}\hat{Q}u_n\,dx - \int u_n^*\hat{Q}\hat{H}u_n\,dx \\ &= \int(\hat{H}u_n)^*\hat{Q}u_n\,dx - \int u_n^*\hat{Q}\hat{H}u_n\,dx = \int E_n u_n^*\hat{Q}u_n\,dx - \int E_n u_n^*\hat{Q}u_n\,dx = 0\end{aligned}$$

This is an "state-dependent" case as it depends on the particular wavefunction. In fact, we have to expect this because an energy eigenstate is a stationary state and nothing changes with time, as was mentioned near the beginning of the lecture. An equivalent example classically would be a particle moving in a potential well. In general, the particle momentum is not conserved. However, if it is sitting at the bottom of the potential and not actually moving, then the momentum is constant, as it is zero at all times.

The more interesting case for $\langle Q\rangle$ being constant is when the expectation value of the commutator is zero for *any* wavefunction, which means the commutator itself must be zero. By definition, this is when the operator $\hat{Q}$ commutes with the Hamiltonian. As we saw, this happens when it is compatible with the Hamiltonian. In this case, the dynamical variable associated with $\hat{Q}$ is a *constant of the motion*, i.e. a classically conserved quantity. In terms of quantum mechanics, the Ehrenfest theorem only applies to the time dependence between measurements but does then say the expectation value of the quantity is constant if it commutes with $\hat{H}$. The most trivial case of this is for $\hat{Q} = \hat{H}$, since the commutator of an operator with itself is zero, i.e. $[\hat{H}, \hat{H}] = 0$. This means the energy is a classically conserved quantity, as we would expect for the conservative systems being considered. Another example is that of a free particle; since $V = 0$, the Hamiltonian is simply

$$\hat{H}_{\text{free}} = \frac{\hat{p}^2}{2m}$$

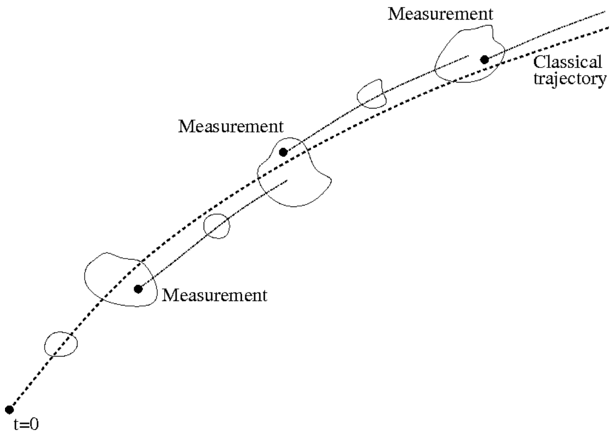
We already saw in the last lecture that $\hat{p}$ commutes with this, i.e.

$$[\hat{H}_{\text{free}}, \hat{p}] = 0$$

so momentum is a conserved quantity for a free particle, as we would expect.

5 Classical motion

We now know enough quantum mechanics to know how classical motion arises. If I throw a ball in the air and watch it rise and fall by shining light on it, I am effectively performing measurements of its position very frequently. What is happening at a quantum level? The first position measurement effectively collapses the wavefunction to a Dirac delta function, as we discussed previously. The wavefunction then time evolves for a while; the wavefunction may spread out but Ehrenfest says the average position follows the classical trajectory. A short time later, I observe its position again and it collapses again, but it will be close to the classical point. It again then moves and spreads but again stays close to the classical path. This continues but as long as I continue to measure it, then it will never spread too far. This means the probability of measuring it far off the classical path is small, and hence it will appear to approximately follow the classical path overall.



You might be wondering how long is "a short time" as the wavefunction can spread out more if more time is left between measurements and it will tend to look less and less like the classical trajectory. This is indeed the case, but for classical-sized objects, the spread (which depends on the wavelength) is minute and completely unobservable on any realistic timescale between measurements. The same is not true for quantum-sized objects, such as an electron. If left unobserved for some time, the wavefunction will spread significantly and so can be found in a very different position from the classical expectation. One striking example of this is the two-slit experiment. If the electron is observed (semi-)continuously as above by shining light on it, it will map out a classical path, either to hit the barrier or pass through one of the two slits. The screen will show no interference pattern in this case. If the light used to observe the electron is turned off, then the path will spread in a quantum manner rather than the electron moving like a classical particle. Without the light, we cannot tell which slit the electron went through and the screen will show the interference pattern again.

We have so far concentrated on the average position, which looks classical. We now see we have to think about the spread of measurements. This is clearly a purely quantum effect.

Second Year Quantum Mechanics - Lecture 15
Measurement spread

Paul Dauncey, 11 Nov 2011

1 Introduction

In the last lecture, we saw that the expectation values of all operators followed the classical time dependence. However, while the average measurement might obey classical equations, the spread of the measurements of position and momentum (or any other variable) are clearly only a quantum effect. We will look at this in this lecture.

2 RMS spread

We would like a measure of how spread out the measurement probability distribution for the variable associated with an operator is. A couple of lectures ago, we considered a state which was in an equal superposition of the first two energy states

$$\psi = \frac{1}{\sqrt{2}}(u_1 + u_2)$$

so a measurement will give either E_1 or E_2. Therefore, its energy expectation value is

$$\langle E\rangle = \frac{1}{2}(E_1 + E_2)$$

A measure of the spread of the measurements would be given by the root mean square (RMS) around this average, here labelled as ΔE. To find this, we first find the mean square, which means the average of the square of the difference of each measurement from the expectation value. A measurement of E_1 has a difference from the expectation value of

$$E_1 - \langle E\rangle = E_1 - \frac{1}{2}(E_1 + E_2) = \frac{1}{2}(E_1 - E_2)$$

and similarly

$$E_2 - \langle E\rangle = E_2 - \frac{1}{2}(E_1 + E_2) = \frac{1}{2}(E_2 - E_1)$$

Hence, the average of the squares of these is simply

$$(\Delta E)^2 = \Delta E^2 = \frac{1}{2}\left\{\left[\frac{1}{2}(E_1 - E_2)\right]^2 + \left[\frac{1}{2}(E_2 - E_1)\right]^2\right\} = \frac{1}{4}(E_2 - E_1)^2$$

and so the RMS is then

$$\Delta E = \frac{1}{2}|E_2 - E_1|$$

This gives some measure of how far off we would be likely to be from the expectation value for an average measurement.

The general case for this can be derived. For an operator $\hat{Q}$, then the wavefunction can be expressed as an expansion in eigenstates of $\hat{Q}$

$$\psi = \sum_n a_n \phi_n$$

and the expectation value is

$$\langle Q\rangle = \sum_n |a_n|^2 q_n$$

For each state, the probability of getting its eigenvalue is $|a_n|^2$ so the mean square around the expectation value is

$$\Delta q^2 = \sum_n |a_n|^2 (q_n - \langle Q\rangle)^2 = \sum_n |a_n|^2 (q_n^2 - 2q_n\langle Q\rangle + \langle Q\rangle^2) = \sum_n |a_n|^2 q_n^2 - 2\langle Q\rangle \sum_n |a_n|^2 q_n + \langle Q\rangle^2 \sum_n |a_n|^2$$

since $\langle Q\rangle$ is just a number. However, we know these last two sums are $\langle Q\rangle$ and 1, respectively, so this can be written as

$$\Delta q^2 = \langle Q^2\rangle - 2\langle Q\rangle^2 + \langle Q\rangle^2 = \langle Q^2\rangle - \langle Q\rangle^2$$

A similar expression occurs in statistics of measurements. The new thing we need here is $\langle Q^2\rangle$, which can be evaluated by applying the operator twice to the wavefunction and then integrating

$$\langle Q^2\rangle = \int_{-\infty}^{\infty} \psi^* \hat{Q}^2 \psi \, dx$$

However, there is a different expression for this which can be quite useful as it only requires applying the operator once. The above is

$$\langle Q^2\rangle = \int_{-\infty}^{\infty} \psi^* \hat{Q}(\hat{Q}\psi)\, dx = \int_{-\infty}^{\infty} (\hat{Q}\psi)^* (\hat{Q}\psi)\, dx = \int_{-\infty}^{\infty} |\hat{Q}\psi|^2 \, dx$$

Note, this also shows that the integral is over positive-definite values so that the expectation value $\langle Q^2\rangle \geq 0$ always, as expected for the average of a squared quantity.

Checking the above expression gives the right ΔE^2 for our two state system, then

$$\langle E^2\rangle = \frac{1}{2}(E_1^2 + E_2^2)$$

and hence it says

$$\begin{aligned} \Delta E^2 &= \langle E^2\rangle - \langle E\rangle^2 = \frac{1}{2}(E_1^2 + E_2^2) - \frac{1}{4}(E_1 + E_2)^2 = \frac{1}{4}(2E_1^2 + 2E_2^2 - E_1^2 - E_2^2 - 2E_1E_2) \\ &= \frac{1}{4}(E_1^2 + E_2^2 - 2E_1E_2) = \frac{1}{4}(E_2 - E_1)^2 \end{aligned}$$

as before.

There is no limit on how narrow or wide the spread of measurements around the expectation value can be. If the wavefunction is an eigenstate of the quantity being measured, then $\Delta q = 0$ as the only possible value we can measure is the eigenvalue so there is no spread. However, if the wavefunction has a large number of terms in the expansion of eigenstates of the variable, then the spread can be arbitrarily big.

3 The difference operator

It is often convenient to deal with a difference operator where we define

$$\hat{Q}' = \hat{Q} - \langle Q\rangle$$

We need to check this is Hermitian. Remembering the expectation value is just a real number, then consider the Hermitian condition equation

$$\begin{aligned} \int \psi^* (\hat{Q} - \langle Q\rangle)\phi \, dx &= \int \psi^* \hat{Q}\phi\, dx - \int \psi^* \langle Q\rangle \phi\, dx = \int (\hat{Q}\psi)^* \phi\, dx - \int (\langle Q\rangle\psi)^*\phi\, dx \\ &= \int [(\hat{Q} - \langle Q\rangle)\psi]^* \phi\, dx = \int (\hat{Q}'\psi)^*\phi\, dx \end{aligned}$$

and hence $\hat{Q}'$ is indeed Hermitian. It is clear the eigenvalues of $\hat{Q}'$ are $q_n - \langle Q \rangle$ so that we can write

$$\Delta q^2 = \sum_n |a_n|^2 (q_n - \langle Q \rangle)^2 = \langle Q'^2 \rangle = \int |Q'\psi|^2 \, dx$$

as shown above.

4 The SHO ground state

Let's see an example of all of this using the ground state of the simple harmonic oscillator. The wavefunction is

$$u_0(x) = A e^{-ax^2/2}$$

where the constant a is $m\omega_0/\hbar$ and A is the normalisation constant. The normalisation means

$$\int u_0^* x u_0 \, dx = |A|^2 \int e^{-ax^2} \, dx = 1$$

The expectation value of x is then

$$\langle x \rangle = \int u_0^* x u_0 \, dx = |A|^2 \int x e^{-ax^2} \, dx = 0$$

since this is an odd function. Note, this says the average position we will find the particle is at the centre, which makes sense as the SHO potential is symmetric around $x = 0$. For momentum then we need

$$\hat{p} u_0 = -i\hbar \frac{du_0}{dx} = -i\hbar \left(\frac{-2ax}{2} \right) A e^{-ax^2/2} = i\hbar a x u_0$$

and so

$$\langle p \rangle = \int u_0^* \hat{p} u_0 \, dx = i\hbar a \int u_0^* x u_0 \, dx = 0$$

as the integral is $\langle x \rangle = 0$ again. This also makes sense; the particle is as likely to be going left as right so the average momentum must be zero; this indeed holds for any bound state. For the spread in x, since $\langle x \rangle = 0$, then $\hat{x}' = \hat{x} = x$ and so

$$\Delta x^2 = \langle x'^2 \rangle = \int |x u_0|^2 \, dx = |A|^2 \int x^2 e^{-ax^2} \, dx = |A|^2 \frac{1}{2a} \int e^{-ax^2} \, dx = \frac{1}{2a}$$

where we have used a standard integral relation and the fact that u_0 is normalised. Using this, then

$$\Delta x = \frac{1}{\sqrt{2a}}$$

Note, the probability density is

$$|u_0|^2 \propto e^{-ax^2}$$

which looks like a Gaussian distribution. A Gaussian is normally written as

$$\propto e^{-x^2/2\sigma^2}$$

so in terms of the Gaussian width

$$a = \frac{1}{2\sigma^2} \qquad \text{so} \qquad \sigma = \frac{1}{\sqrt{2a}} = \Delta x$$

as would be expected. Finally, for p^2, then again $\hat{p}' = \hat{p}$ so

$$\Delta p^2 = \langle p'^2 \rangle = \int |\hat{p}u_0|\, dx = (i\hbar a)(-i\hbar a) \int |xu_0|\, dx = \hbar^2 a^2 \frac{1}{2a} = \frac{\hbar^2 a}{2}$$

Hence

$$\Delta p = \hbar \sqrt{\frac{a}{2}}$$

We therefore find

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

This should look familiar; it is related to the Heisenberg Uncertainty Principle which states

$$\Delta x \Delta p \geq \frac{\hbar}{2}$$

for any state. We see the ground state of the SHO is a "minimum uncertainty state", because the equality holds, rather than the more general inequality. If we had looked at the higher energy SHO states, we would have found that the uncertainty product was bigger than $\hbar/2$ and so only the ground state is at the minimum uncertainty. We will derive the full Uncertainty Principle in the next lecture and see the conditions for when this minimum uncertainty can occur.

Second Year Quantum Mechanics - Lecture 16
The Uncertainty Principle

Paul Dauncey, 14 Nov 2011

1 Introduction

Last year, you learned

$$\Delta x \Delta p \geq \frac{\hbar}{2}$$

meaning that there is a fundamental limit to the precision with which the outcome of position and momentum measurements on a QM wavefunction can be predicted. This has nothing to do with experimental resolution and holds even for an infinitely precise apparatus. Note, the Uncertainty Principle tells us nothing about the uncertainty of an individual measurement outcome of x or p alone.

We can show that the Uncertainty Principle is a consequence of the postulates of quantum mechanics. Specifically, using the commutator of $\hat{x}$ and $\hat{p}$

$$[\hat{x}, \hat{p}] = \hat{x}\hat{p} - \hat{p}\hat{x} = i\hbar$$

we can prove the Uncertainty Principle. Firstly, we need to introduce the Schwartz inequality.

2 The Schwartz inequality

To try to motivate this result, consider two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. You all know the dot product is

$$\boldsymbol{a}.\boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}| \cos\theta$$

where θ is the angle between the two vectors. Since $\cos\theta \leq 1$, then we know

$$|\boldsymbol{a}|^2 |\boldsymbol{b}|^2 \geq (\boldsymbol{a}.\boldsymbol{b})^2$$

It is always useful to consider in what circumstances an inequality becomes an equality. Here it is easy to see; it is when $\cos\theta = \pm 1$, which is when the two vectors are parallel or antiparallel (although not necessarily equal in length). This condition is given by

$$\boldsymbol{a} = \lambda \boldsymbol{b}$$

for any constant λ. We now have to take a conceptual leap to translate this result to functions and so get the Schwartz inequality. Consider

$$|\boldsymbol{a}|^2 = \sum_{i=1}^{3} a_i^2$$

The sum goes from 1 to 3 as vectors have three components as we live in 3D. However, it is perfectly possible to define vectors with many more components which then live in higher dimensions. Taking the number of components to infinity, then we get a continuous range of components, i.e.

$$|\boldsymbol{a}|^2 = \sum_{i=1}^{3} a_i^2 \to \int a(x)^2 \, dx$$

In fact, as we are needing to handle complex functions, we need a more general form of this

$$|\boldsymbol{a}|^2 = \sum_{i=1}^{3} a_i^2 \to \int a^*(x)a(x)\,dx = \int |a(x)|^2\,dx$$

Similarly, the dot product becomes something like

$$\boldsymbol{a}.\boldsymbol{b} = \sum_{i=1}^{3} a_i b_i \to \int a^*(x)b(x)\,dx$$

or the complex conjugate; strictly we need to consider $\boldsymbol{a}.\boldsymbol{b}^*$ and $\boldsymbol{a}^*.\boldsymbol{b}$ separately. With these forms, then it is plausible that for the functions $a(x)$ and $b(x)$, the equivalent of the inequality above is

$$\int |a(x)|^2\,dx \int |b(x)|^2\,dx \geq \left|\int b^*(x)a(x)\,dx\right|^2$$

This is the Schwartz inequality. Note, in the same way, the inequality becomes an equality if

$$a(x) = \lambda b(x)$$

3 The Uncertainty Principle

We already showed the mean square of measurement value around the average could be found using the difference operator, which we found was Hermitian,

$$\hat{Q}' = \hat{Q} - \langle\hat{Q}\rangle$$

so that the mean square is

$$\Delta q^2 = \int \psi^*(\hat{Q} - \langle\hat{Q}\rangle)^2\psi\,dx = \int \psi^*\hat{Q}'\hat{Q}'\psi\,dx = \int (\hat{Q}'\psi)^*(\hat{Q}'\psi)\,dx = \int |\hat{Q}'\psi|^2\,dx$$

If we have another operator $\hat{R}$, then similarly

$$\Delta r^2 = \int |\hat{R}'\psi|^2\,dx$$

Hence, the product of the two mean squares is

$$\Delta q^2 \Delta r^2 = \int |\hat{Q}'\psi|^2\,dx \int |\hat{R}'\psi|^2\,dx$$

If we now identify $a(x) = \hat{Q}'\psi$ and $b(x) = \hat{R}'\psi$, then we can use the Schwartz inequality to give

$$\Delta q^2 \Delta r^2 \quad \geq \quad \left|\int (\hat{Q}'\psi)^*\hat{R}'\psi\,dx\right|^2 = \left|\int \psi^*\hat{Q}'\hat{R}'\psi\,dx\right|^2 = \left|\left\langle\hat{Q}'\hat{R}'\right\rangle\right|^2$$

Hence, the product of the uncertainties is not arbitrary, but must be at least as big as some positive value. You can see they can be correlated as, when measuring q, we need to expand in the $\hat{Q}$ eigenstates $\phi_{q\,n}$

$$\psi = \sum_n a_n \phi_{q\,n}$$

while for r we need the wavefunction in terms of $\phi_{r\,n}$

$$\psi = \sum_n b_n \phi_{r\,n}$$

Hence, the coefficients a_n and b_n, which determine the measurement probabilities and so the mean squares here, are calculated from the *same* ψ, which means there will be some correlation within them. This is quantified by the Uncertainty Principle.

We need to manipulate the above inequality to a more useful form. We saw the product of two Hermitian operators, $\hat{Q}'\hat{R}'$ is not necessarily Hermitian, but that there are two combinations which are

$$\frac{1}{2}\{\hat{Q}', \hat{R}'\}, \qquad \frac{i}{2}[\hat{Q}', \hat{R}']$$

We have

$$\hat{Q}'\hat{R}' = \frac{1}{2}\left(\hat{Q}'\hat{R}' + \hat{R}'\hat{Q}' - \hat{R}'\hat{Q}' + \hat{Q}'\hat{R}'\right) = \frac{1}{2}\{\hat{Q}', \hat{R}'\} + \frac{1}{2}[\hat{Q}', \hat{R}'] = \left(\frac{1}{2}\{\hat{Q}', \hat{R}'\}\right) - i\left(\frac{i}{2}[\hat{Q}', \hat{R}']\right)$$

Hence, our inequality can be written as

$$\begin{aligned} \Delta q^2 \Delta r^2 \quad &\geq \quad \left| \int \psi^* \frac{1}{2}\{\hat{Q}', \hat{R}'\}\psi\, dx - i \int \psi^* \frac{i}{2}[\hat{Q}', \hat{R}']\psi\, dx \right|^2 \\ &\geq \quad \left| \left\langle \frac{1}{2}\{\hat{Q}', \hat{R}'\} \right\rangle - i \left\langle \frac{i}{2}[\hat{Q}', \hat{R}'] \right\rangle \right|^2 = |A - iC|^2 \end{aligned}$$

Since both these combinations are Hermitian, the expectation values of them are real so the LHS is of the form $|A - iC|^2$ for real A and C. This gives $|A - iC|^2 = (A - iC)(A + iC) = A^2 + C^2$. Hence we have

$$\Delta q^2 \Delta r^2 \quad \geq \quad A^2 + C^2 = \left\langle \frac{1}{2}\{\hat{Q}', \hat{R}'\} \right\rangle^2 + \left\langle \frac{i}{2}[\hat{Q}', \hat{R}'] \right\rangle^2$$

Note, since both terms are real and positive, then

$$\Delta q^2 \Delta r^2 \geq A^2 \qquad \text{and} \qquad \geq C^2$$

It is usually more convenient to deal with the original operators, not the prime versions. Consider

$$\hat{Q}'\hat{R}' = (\hat{Q} - \langle Q \rangle)(\hat{R} - \langle R \rangle) = \hat{Q}\hat{R} - \hat{Q}\langle R \rangle - \hat{R}\langle Q \rangle + \langle Q \rangle\langle R \rangle$$

Hence

$$\langle R'Q' \rangle = \langle RQ \rangle - \langle Q \rangle\langle R \rangle - \langle R \rangle\langle Q \rangle + \langle Q \rangle\langle R \rangle = \langle RQ \rangle - \langle Q \rangle\langle R \rangle$$

Similarly

$$\langle Q'R' \rangle = \langle QR \rangle - \langle Q \rangle\langle R \rangle$$

so subtracting gives

$$\langle [\hat{Q}', \hat{R}'] \rangle = \langle [\hat{Q}, \hat{R}] \rangle$$

while adding gives

$$\langle \{\hat{Q}', \hat{R}'\} \rangle = \langle \{\hat{Q}, \hat{R}\} \rangle - 2\langle Q \rangle\langle R \rangle$$

Hence

$$\Delta q^2 \Delta r^2 \quad \geq \quad \left(\left\langle \frac{1}{2}\{\hat{Q}, \hat{R}\} \right\rangle - \langle \hat{Q} \rangle\langle R \rangle \right)^2 + \left\langle \frac{i}{2}[\hat{Q}, \hat{R}] \right\rangle^2$$

This is the general form for the Heisenberg Uncertainty Principle. Note, the inequality becomes an equality if $a = \lambda b$ which means $\hat{Q}'\psi = \lambda\hat{R}'\psi$.

4 Position and momentum

Taking the case for $\hat{Q}=\hat{x}$ and $\hat{R}=\hat{p}$, then we know $[\hat{x},\hat{p}]=i\hbar$ so

$$\begin{aligned}\Delta x^2\Delta p^2 &\ge \left(\left\langle\frac{1}{2}\{\hat{x},\hat{p}\}\right\rangle-\langle x\rangle\langle p\rangle\right)^2+\left\langle\frac{i}{2}[\hat{x},\hat{p}]\right\rangle^2\\ &\ge \left(\left\langle\frac{1}{2}\{\hat{x},\hat{p}\}\right\rangle-\langle x\rangle\langle p\rangle\right)^2+\frac{\hbar^2}{4}\end{aligned}$$

The first term depends on the state under consideration. However, it is clear that the uncertainty product has to be at least as big as the second term alone, so this is often written as

$$\begin{aligned}\Delta x^2\Delta p^2 &\ge \frac{\hbar^2}{4}\\ \Delta x\Delta p &\ge \frac{\hbar}{2}\end{aligned}$$

as you saw previously. Because of dropping the first term, then there are now two conditions for the inequality to become an equality. Firstly, as always, the functions must satisfy $a(x)=\lambda b(x)$ which means

$$\begin{aligned}\hat{x}'\psi &= \lambda\hat{p}'\psi\\ (\hat{x}-\langle x\rangle)\psi &= \lambda(\hat{p}-\langle p\rangle)\psi\end{aligned}$$

Secondly, if the equality is to equal only the commutator term, then the anticommutator term must be zero, so we need

$$\left\langle\frac{1}{2}\{\hat{x},\hat{p}\}\right\rangle=\langle x\rangle\langle p\rangle$$

If both these conditions hold, then the state is said to be a “minimum uncertainty state”.

We calculated $\Delta x\Delta p$ for the SHO ground state in the last lecture. We found both $\langle x\rangle=0$ and $\langle p\rangle=0$ so $\hat{p}'=\hat{p}$ and $\hat{x}'=\hat{x}=x$. We also found $\Delta x^2=\langle x^2\rangle=1/2a$ and $\Delta p^2=\langle p^2\rangle=\hbar^2a/2$, so $\Delta x^2\Delta p^2=\hbar^2/4$ and hence found it was indeed a minimum uncertainty state.

To check the above conditions do hold, then noting we found

$$\hat{p}u_0=i\hbar a x u_0$$

then clearly $\hat{x}'u_0=\lambda\hat{p}'u_0$ with $\lambda=1/i\hbar a$. Also, using the above, then

$$x\hat{p}u_0=i\hbar a x^2u_0,\qquad \hat{p}(xu_0)=-i\hbar u_0+x\hat{p}u_0=-i\hbar u_0+i\hbar a x^2u_0,$$

so that

$$\langle\{\hat{x},\hat{p}\}\rangle=i\hbar(2a\langle x^2\rangle-1)=0$$

since we found $\langle x^2\rangle=1/2a$. Hence, this state does indeed satisfy both conditions, as expected.

Second Year Quantum Mechanics - Lecture 17
Continuous eigenvalues

Paul Dauncey, 17 Nov 2011

1 The Dirac delta function

The Dirac delta function is the continuous equivalent of the discrete-case Kronecker delta function. A Dirac delta function $\delta(x)$ is zero everywhere except for at the point $x = 0$ where it is infinite. The only sensible definition is when it is integrated, for which it satisfies

$$\int_{-\infty}^{\infty} \delta(x-a) f(x)\, dx = f(a)$$

Note, for the particular case of $f(x) = 1$, this means

$$\int_{-\infty}^{\infty} \delta(x-a)\, dx = 1$$

which means it has an area of one below the function. Some other useful properties are

$$\delta(-x) = \delta(x), \qquad \delta(\alpha x) = \delta(x)/\alpha$$

Finally, there is a standard integral form for the delta function

$$\int_{-\infty}^{\infty} e^{ikx}\, dk = 2\pi\delta(x)$$

You can think of it as the limit of various functions, such as a normalised Gaussian

$$G(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-a)^2/2\sigma^2}$$

with the width σ being taken to zero, or a square box of width w and height $1/w$, with w being taken to zero.

2 Continuous eigenvalues

Not all dynamical variables have discrete, i.e. quantised, eigenvalues; we know the most common ones we have used, momentum and position, do not. Also, unbound energy eigenstates do not have discrete energies, but have a continuous spectrum. We need to generalise what we have done to deal with these cases. Firstly, for the eigenvalue equation, we need to replace

$$\hat{Q}\phi_n(x) = q_n\phi_n(x)$$

with

$$\hat{Q}\phi(q,x) = q\phi(q,x)$$

where q is now a continuous variable and so the eigenstate can be considered to be a function of both x and q. We then replace the general expansion of a wavefunction

$$\psi(x) = \sum_n a_n\phi_n(x)$$

with

$$\psi(x) = \int a(q)\phi(q,x)\, dq$$

One critical issue with the expansion is that we need to be able to find the values of the coefficients a_n or $a(q)$. We saw how this is done for the discrete case; consider the overlap integral for a particular eigenstate ϕ_m

$$\int \phi_m(x)^* \psi(x)\, dx = \sum_n a_n \int \phi_m(x)^* \phi_n(x)\, dx$$

but we know the eigenstates of any Hermitian operator are orthonormal, meaning

$$\int \phi_m(x)^* \phi_n(x)\, dx = \delta_{mn}$$

so the overlap integral becomes

$$\int \phi_m(x)^* \psi(x)\, dx = \sum_n a_n \delta_{mn} = a_m$$

Hence, we can find the coefficients by doing the overlap integrals. How does this work for continuous eigenvalues? If we calculate the overlap integral for a particular eigenstate $\phi(q', x)$, then we get

$$\int \phi(q', x)^* \psi(x)\, dx = \int \phi(q', x)^* \left[\int a(q) \phi(q, x)\, dq \right] dx = \int a(q) \left[\int \phi(q', x)^* \phi(q, x)\, dx \right] dq$$

If we want this to give $a(q')$, then we have to assume

$$\int \phi(q', x)^* \phi(q, x)\, dx = \delta(q' - q)$$

because then

$$\int \phi(q', x)^* \psi(x)\, dx = \int a(q) \delta(q' - q)\, dq = a(q')$$

Hence, it seems that the generalisation of the orthogonality condition needs to be that

$$\int \phi_m(x)^* \phi_n(x)\, dx = \delta_{mn}$$

becomes

$$\int \phi(q', x)^* \phi(q, x)\, dx = \delta(q' - q)$$

This relation does indeed hold for Hermitian operator eigenstates.

Since the wavefunction is normalised, previously we had

$$1 = \int |\psi(x)|^2\, dx = \int \left[\sum_n a_n^* \phi_n(x)^* \right] \left[\sum_m a_m \phi_m(x) \right] dx = \sum_{nm} a_n^* a_m \left[\int \phi_n(x)^* \phi_m(x)\, dx \right]$$

and using the orthonormality condition

$$\int \phi_n(x)^* \phi_m(x)\, dx = \delta_{nm}$$

this gave

$$1 = \sum_{nm} a_n^* a_m \delta_{nm} = \sum_n a_n^* a_n = \sum_n |a_n|^2$$

which shows that the total probability does indeed add to one. How do we get this to work for the continous case? We can do something very similar

$$\begin{aligned}
1 &= \int |\psi(x)|^2\,dx = \int \left[\int a(q)^*\phi(q,x)^*\,dq\right]\left[\int a(q')\phi(q',x)\,dq'\right]\,dx \\
&= \int\int a(q)^*a(q')\left[\int \phi(q,x)^*\phi(q',x)\,dx\right]\,dq\,dq' = \int\int a(q)^*a(q')\delta(q-q')\,dq\,dq' \\
&= \int a(q)^*\left[\int a(q')\delta(q-q')\,dq'\right]\,dq = \int a(q)^*a(q)\,dq = \int |a(q)|^2\,dq
\end{aligned}$$

This is easy to interpret; instead of $|a_n|^2$ being the probability of getting the result q_n, we now have $|a(q)|^2\,dq$ as being the probability of getting a result in the range q to $q+dq$, i.e. $|a(q)|^2$ is the probability density of the measurement giving q. The fact that $\int |a(q)|^2\,dq = 1$ ensures the total probability to measure any value is one, as required.

Finally, let's check the expectation value also makes sense. We have shown for discrete eigenstates that the expectation value can be written in two ways

$$\langle Q\rangle = \int \psi^*\hat{Q}\psi\,dx = \sum_n |a_n|^2 q_n$$

so given the continuous probability, the second expression clearly becomes

$$\langle Q\rangle = \int |a(q)|^2 q\,dq$$

However, the first expression for $\langle Q\rangle$ doesn't depend on whether the eigenstates are continuous or discrete so we need to check this is still consistent. Consider this first expression when expanded into the eigenstates

$$\begin{aligned}
\langle Q\rangle &= \int \psi^*\hat{Q}\psi\,dx = \int \left[\int a(q)\phi(q,x)\,dq\right]^*\hat{Q}\left[\int a(q')\phi(q',x)\,dq'\right]\,dx \\
&= \int\int a(q)^*a(q')q'\left[\int \phi(q,x)^*\phi(q',x)\,dx\right]\,dq\,dq' \\
&= \int\int a(q)^*a(q')q'\delta(q-q')\,dq\,dq' \\
&= \int a(q)^*\left[\int a(q')q'\delta(q-q')\,dq'\right]\,dq \\
&= \int a(q)^*a(q)q\,dq = \int |a(q)|^2 q\,dq
\end{aligned}$$

as required. Hence, the equivalent expressions are

	Discrete	Continuous
Eigenstate equation	$\hat{Q}\phi_n(x) = q_n\phi_n(x)$	$\hat{Q}\phi(q,x) = q\phi(q,x)$
Orthonormality	$\int \phi_m(x)^*\phi_n(x)\,dx = \delta_{mn}$	$\int \phi(q',x)^*\phi(q,x)\,dx = \delta(q'-q)$
Expansion	$\psi(x) = \sum_n a_n\phi_n(x)$	$\psi(x) = \int a(q)\phi(q,x)\,dq$
Overlap integral	$a_n = \int \phi_n^*(x)\psi(x)\,dx$	$a(q) = \int \phi^*(q,x)\psi(x)\,dx$
Probability	$\|a_n\|^2$	$\|a(q)\|^2\,dq$
Total probability	$\sum_n \|a_n\|^2 = 1$	$\int \|a(q)\|^2\,dq = 1$
Expectation value	$\langle Q\rangle = \sum_n \|a_n\|^2 q_n$	$\langle Q\rangle = \int \|a(q)\|^2 q\,dq$
	$\langle Q\rangle = \int \psi^*(x)\hat{Q}\psi(x)\,dx$	$\langle Q\rangle = \int \psi^*(x)\hat{Q}\psi(x)\,dx$

3 Eigenstates of momentum

We have already seen the eigenstates of momentum; they are

$$\phi(p,x) = Ae^{ipx/\hbar}$$

We can now fix the constant A using the continuous orthonormality condition

$$\delta(p-p') = \int_{-\infty}^{\infty} \phi(p',x)^*\phi(p,x)\,dx = |A|^2\int_{-\infty}^{\infty} e^{-ip'x/\hbar}e^{ipx/\hbar}\,dx = |A|^2\int_{-\infty}^{\infty} e^{i(p-p')x/\hbar}\,dx$$

This is a standard integral; the general form is

$$\int_{-\infty}^{\infty} e^{ikx}\,dx = 2\pi\delta(k)$$

so with $k = (p-p')/\hbar$ here, this gives

$$\delta(p-p') = |A|^2 2\pi\delta\left(\frac{p'-p}{\hbar}\right) = |A|^2 2\pi\hbar\delta(p'-p)$$

Hence, choosing A to be real, then $A = 1/\sqrt{2\pi\hbar}$ and

$$\phi(p,x) = \frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar}$$

Given this, then the general expansion in momentum eigenstates is

$$\psi(x) = \int a(p)\phi(p,x)\,dp = \frac{1}{\sqrt{2\pi\hbar}}\int a(p)e^{ipx/\hbar}\,dp$$

and the expansion coefficients are given by the overlap integrals

$$a(p) = \int \phi(p,x)^*\psi(x)\,dx = \frac{1}{\sqrt{2\pi\hbar}}\int \psi(x)e^{-ipx/\hbar}\,dx$$

which is (give or take the $\hbar$'s) identical to a complex Fourier transform.

4 Eigenstates of position

We have said that the position operator is simply $\hat{x} = x$ and that the eigenstates would need to be like delta functions. Specifically, if we measure position, then we know where the particle is, and a further measurement immediately afterwards would give the same result, as usual. This would imply that the eigenstate must be highly peaked at the position measured, or else there would be some probability of measuring the particle somewhere else. In fact, in the limit of an exact position measurement, then it will have to be singular, i.e. like a Dirac delta function. Mathematically, we are looking for an eigenstate which obeys

$$x\phi(x_m, x) = x_m \phi(x_m, x)$$

where $\phi(x_m, x)$ is the eigenstate for the particle to be at a particular position x_m. The easiest way to check that a delta function satisfies this equation is to integrate it, so

$$\int x\phi(x_m, x)\, dx = x_m \int \phi(x_m, x)\, dx$$

If we take

$$\phi(x_m, x) = \delta(x - x_m)$$

then the LHS is

$$\int x\delta(x - x_m)\, dx = x_m$$

and the RHS is

$$x_m \int \delta(x - x_m)\, dx = x_m$$

also, hence satisfying the equation. We should check this has the right properties. The orthonormality condition is

$$\delta(x_m - x'_m) = \int_{-\infty}^{\infty} \phi(x_m, x)^* \phi(x'_m, x)\, dx$$

This is most easily done by initially only substituting in the delta function (which is real) for the first eigenstate above in order to do the integral, and then substituting in the second delta function afterwards. This gives

$$\int_{-\infty}^{\infty} \delta(x_m - x)\phi(x'_m, x)\, dx = \phi(x'_m, x_m) = \delta(x_m - x'_m)$$

so they do indeed satisfy the orthonormality condition. Finally, we need to see what an expansion in terms of the eigenstates gives. For this

$$\psi(x) = \int a(x_m)\phi(x_m, x)\, dx_m = \int a(x_m)\delta(x - x_m)\, dx_m = a(x)$$

Hence the expansion coefficients *are* the wavefunction; the probability density for measuring x_m is $|a(x_m)|^2 = |\psi(x_m)|^2$ as we have assumed all along. Hence, this is not a special case for position, but just a consequence of the particular form of the position eigenstates.

5 Normalisation

This all seems fine until we consider the normalisation of the eigenstate themselves. The orthonormality condition for the discrete case was

$$\int \phi_n(x)^* \phi_m(x)\, dx = \delta_{nm}$$

which means for $n = m$, then

$$\int \phi_n(x)^* \phi_n(x)\, dx = \int |\phi_n(x)|^2\, dx = \delta_{nn} = 1$$

which is what we mean by normalised. However, for the continuous case, then with $q' = q$, we get

$$\int \phi(q,x)^* \phi(q,x)\, dx = \int |\phi(q,x)|^2\, dx = \delta(q-q) = \delta(0) = \infty$$

This is very general; it means *all* eigenstates with continuous eigenvalues cannot be normalised.

To check that the normalisation really is infinite e.g. for the momentum case, doing it explicitly gives

$$\int |\phi(p,x)|^2\, dx = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} e^{-ipx/\hbar} e^{ipx/\hbar}\, dx = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dx = \infty$$

as we would expect from the above.

Second Year Quantum Mechanics - Lecture 18
Wavepackets

Paul Dauncey, 18 Nov 2011

1 Physical limitations

We found the orthonormality condition for eigenstates with continuous eigenvalues, such as for x and p, was

$$\int \phi(q', x)^* \phi(q, x)\, dx = \delta(q' - q)$$

which is infinite for $q' = q$, so eigenstates with continuous eigenvalues cannot be normalised. As we saw, these include all position and momentum eigenstates and in addition, any unbound energy eigenstates.

This means if a particle wavefunction was equal to one of these eigenstates, then the probability density would no longer make sense and our whole structure would have to be rethought. To avoid this, we have to conclude that the eigenstates cannot correspond to physical states. The other way to look at this is that the wavefunction cannot ever collapse to one of these eigenstates.

Why can't we make a pure momentum eigenstate? This would have a single value for k or equivalently for the wavelength λ. You should know from Fourier analysis that a pure wavelength is only possible if the wave has an infinite extent throughout all space. Hence, any apparatus which measures momentum and collapsed the wavefunction to a perfect pure momentum state would have to be infinitely large. Another way to see this is from the HUP; the only way to get $\Delta p = 0$ is to have $\Delta x = \infty$, i.e. a infinite sized system. This is clearly impossible in practise so no measurement can ever give an exact momentum value.

What about a pure position eigenstate? If we could make this, then the expansion in terms of momentum eigenstates would give

$$\delta(x_m - x) = \frac{1}{\sqrt{2\pi\hbar}} \int a(p) e^{ipx/\hbar}\, dp$$

and hence the overlap integral would be

$$a(p) = \frac{1}{\sqrt{2\pi\hbar}} \int e^{-ipx/\hbar} \delta(x_m - x)\, dx = \frac{1}{\sqrt{2\pi\hbar}} e^{-ipx_m/\hbar}$$

which means the probability density for momentum $|a(p)|^2$ is

$$|a(p)|^2 = \frac{1}{2\pi\hbar} e^{-ipx_m/\hbar} e^{ipx_m/\hbar} = \frac{1}{2\pi\hbar}$$

i.e. is constant with p, meaning the same for all p no matter how big, which gives an infinitely wide range of momentum. This is again expected from the HUP; the only way to have $\Delta x = 0$ is to have $\Delta p = \infty$. This is a problem as at some later time we could measure an arbitrarily big momentum, which of course means arbitrarily large energy. Hence, a position eigenstate takes infinite energy to create. This can also be seen directly as the curvature $d^2\psi/dx^2$ at the top of the delta function is infinitely negative, which means the kinetic energy operator $\hat{T} = -(\hbar^2/2m)d^2/dx^2$ has an infinite positive expectation value. In fact, we can even see this result by considering an explicit experiment, where we shine a laser beam on an electron and try to find its position by looking at the location of scattering photons. You all know the resolution will be limited by the photon wavelength so to get an infinitely precise location, we would need an infinitely short wavelength, which is an infinitely high frequency and so, by de Broglie, an infinite energy.

2 Wavepackets

Hence, we have to conclude we cannot actually make eigenstates of position or momentum. This leaves us to wonder about what *does* happen when we make a measurement of these values, as we would normally expect the wavefunction to collapse to the eigenstate. If the wavefunction did collapse to a position delta function when measuring x, then the equal spread of all momenta would mean there would be no memory of the original particle momentum. However, we certainly can measure both position and momentum of single particles to some accuracy, as easily seen in many experiments in high energy physics.

We have to think about what happens when we measure position (or momentum). No experimental apparatus is perfect so we can never measure a continuous variable infinitely accurately. Hence, we will always only know the position of a particle with some resolution Δx. This means an immediate second measurement only needs to be reproducible within this range, not exactly reproducible as we assumed before. (Note, the Δx here is the intrinsic (physically limiting) resolution of the measurement, such as the wavelength of the laser mentioned above, not some degraded resolution due to poor optics, for example.) The removal of exact reproducibility of the measurement means the wavefunction does not need to collapse to an infinitely thin Dirac delta function, but to something with a finite width. Given the central limit theorem, then it is reasonable to assume this would be something like a Gaussian in the position probability around the measured value x_m

$$|\psi(x)|^2 = |A|^2 e^{-(x-x_m)^2/2\Delta x^2}$$

Naively, we might just take the real square root of this function to get $\psi(x)$, i.e.

$$\psi(x) = A e^{-(x-x_m)^2/4\Delta x^2}$$

but things are a bit more subtle. Let's assume for a moment that before the position measurement, we measured momentum p_m to extremely high accuracy so that we can say we have an (effectively) pure momentum state before the position measurement, i.e. $\sim e^{ip_m x/\hbar}$. The probability of getting an output wavefunction ψ from the position measurement is then given by modulus squared of the overlap integral, where the integral is

$$\frac{1}{\sqrt{2\pi\hbar}} \int \psi(x)^* e^{ip_m x/\hbar}\, dx$$

and so the most likely wavefunction after the x measurement will be the one with the biggest overlap. Taking the naive

$$\psi(x) = A e^{-(x-x_m)^2/4\Delta x^2}$$

gives a very small overlap as multiplying by the $e^{ip_m x/\hbar}$ gives both positive and negative contributions which cancel out. A much bigger overlap is given if we chose

$$\psi(x) = A e^{-(x-x_m)^2/4\Delta x^2} e^{ip_m x/\hbar}$$

which still gives a Gaussian probability distribution. This looks like a travelling wave modulated by a Gaussian envelope and so the exponentials cancel each other, giving only positive contributions to the integral and hence giving a large overlap. Such a wavefunction is called a *wavepacket.*

Although it looks like it only contains one momentum p_m, because of the Gaussian modulation, in reality it must contain a spread of momenta; a pure momentum eigenstate must extend throughout all space and this does not. We can find the probability density for p of this wavefunction by taking the Fourier transform

$$\begin{aligned} a(p) &= \frac{1}{\sqrt{2\pi\hbar}}\int \psi(x)e^{-ipx/\hbar}\,dx = \frac{A}{\sqrt{2\pi\hbar}}\int e^{-(x-x_m)^2/4\Delta x^2}e^{ip_mx/\hbar}e^{-ipx/\hbar}\,dx \\ &= \frac{A}{\sqrt{2\pi\hbar}}\int e^{-(x-x_m)^2/4\Delta x^2}e^{-i(p-p_m)x/\hbar}\,dx \end{aligned}$$

This integral can be done and results in

$$a(p) = Be^{-(p-p_m)^2\Delta x^2/\hbar^2}e^{-ipx_m/\hbar}$$

Note, this has a similar form for $a(p)$ as for $\psi(x)$. By squaring this, we find it also has a Gaussian spread in momentum probability density

$$|a(p)|^2 = |B|^2 e^{-2(p-p_m)^2\Delta x^2/\hbar^2} = |B|^2 e^{-(p-p_m)^2/2\Delta p^2}$$

with a central value of p_m and a spread size Δp inversely proportional to the position spread. Explicitly

$$\frac{1}{2\Delta p^2} = \frac{2\Delta x^2}{\hbar^2}$$

so

$$\Delta x^2\Delta p^2 = \frac{\hbar^2}{4} \qquad \text{so} \qquad \Delta x\Delta p = \frac{\hbar}{2}$$

Hence, a Gaussian wavepacket like this is a minimum uncertainty wavepacket in the HUP sense. If you test it against the two conditions necessary for this, you will find it satisfies both.

An “ideal” measurement would have resulted in the wavefunction collapsing to a delta function and the momentum getting an infinite spread, with all values equally probable. This would mean it has no memory of the previous momentum at all. However, by measuring with a finite resolution, then we end up with a normalisable wavepacket, and it will usually preserve the average momentum of the state before the measurement. This generally gives a larger spread in momentum than before the measurement, but it is not infinite and so can “remember” the previous momentum on average. Of course, it is still fundamentally probabilistic and we could get any p in principle, but values of p more than a few times Δp away from p_m have a very small probability as the overlap integral becomes small. This does mean that the original Postulate 2 is now seen to be economical with the truth; it applies to eigenstates with discrete eigenvalues, but the situation is more complicated for the continuous case.

3 Time development

The above considers the wavepacket when measured at $t = 0$. At later times, the wavefunction follows the Schrödinger equation and, for the case of a free particle, the wavepacket spreads out further. We wrote the general time dependent solution for the discrete case in terms of the expansion in the energy eigenstates

$$\psi(x, t = 0) = \sum_n a_n u_n(x)$$

so that at later times

$$\psi(x, t) = \sum_n a_n u_n(x)e^{-iE_nt/\hbar}$$

For the continuous case, this generalises to

$$\psi(x, t) = \int a(E)u(E, x)e^{-iEt/\hbar}\,dE$$

For the free particle, the momentum eigenstates are also the energy eigenstates so in fact since

$$\psi(x, t = 0) = \frac{1}{\sqrt{2\pi\hbar}} \int a(p) e^{ipx/\hbar} \, dp$$

then as $E = p^2/2m$, the time dependent wavefunction is given by

$$\psi(x,t) = \frac{1}{\sqrt{2\pi\hbar}} \int a(p) e^{ipx/\hbar} e^{-iEt/\hbar} \, dp = \frac{1}{\sqrt{2\pi\hbar}} \int a(p) e^{ipx/\hbar} e^{-ip^2 t/2m\hbar} \, dp$$

Note, this means that we can consider the expansion coefficients as time-dependent, just as we did before

$$a(p,t) = a(p) e^{-ip^2 t/2m\hbar} \qquad \text{so} \qquad |a(p,t)|^2 = |a(p)|^2$$

which means the probability density in momentum is constant with time. This is expected; the momentum eigenstates are also energy eigenstates and the energy probabilities are always constant with time.

The above can actually be solved; substituting in $a(p)$ explicitly, this gives

$$\psi(x,t) = \frac{B}{\sqrt{2\pi\hbar}} \int e^{-\Delta x^2 (p-p_m)^2/\hbar^2} e^{ip(x-x_m)/\hbar} e^{-ip^2 t/2m\hbar} \, dp$$

This integral can be evaluated with a long calculation and indeed is one of the (few) cases where the TDSE can be solved in a closed form. The result is that the average position moves with $p_m t/m$ and the position Gaussian width increases with time. Note, the momentum Gaussian width does not increase with time as the momentum states are energy eigenstates, so their probability density is constant. This means, since the position spread increases and the momentum spread is constant, that the uncertainty product becomes more than $\hbar/2$. Hence for $t > 0$, the state is no longer a minimum uncertainty state.

Note, this is for a free particle; other potentials will give a different rate of expansion, or shrink the wavepacket size, or even keep the packet the same size.

Second Year Quantum Mechanics - Lecture 19
Representations and ladder operators

Paul Dauncey, 21 Nov 2011

1 Representations

What we have done in the previous two lectures make x and p look more and more symmetric in how they are handled. The expansion of a wavefunction in terms of momentum eigenstates is

$$\psi(x) = \int a(q)\phi(p,x)\,dp = \frac{1}{\sqrt{2\pi\hbar}}\int a(q)e^{ipx/\hbar}\,dp$$

and the overlap integral gives the inverse

$$a(p) = \frac{1}{\sqrt{2\pi\hbar}}\int \psi(x)e^{-ipx/\hbar}\,dx$$

These are just Fourier transforms of each other. As always, the probability density for an x measurement is $|\psi(x)|^2$ and for a p measurement is $|a(p)|^2$. Hence, there seems to be a complete symmetry between x and p with regard to how they are handled.

It turns out we could do the whole of quantum mechanics using p as the variable we are working with, rather than x, and hence would express all our operators in terms of p and d/dp. The function $a(p)$ can be thought of as the wavefunction of the state, but in momentum, rather than position, space; specifically, we could write $a(p) = \Psi(p)$. We could then calculate an expansion of $\Psi(p)$ in position eigenstates and the overlap integral would give us $\psi(x)$. Again squaring $\Psi(p)$ gives the momentum probability density while squaring $\psi(x)$ gives the position probability density. Working with a wavefunction which is a function of p rather than x is called working in the "momentum representation", rather than the "position representation" (sometimes called the "Schrödinger representation") which we have used in the rest of the course. Changing between representations is analogous to changing between sets of coordinate axes for normal vectors; this can make a calculation more convenient but the results you get should be independent of which set you use.

In the momentum representation, then we take $\hat{p} = p$, the variable, but what do we take for $\hat{x}$? It turns out that commutators are the same in any representation. In particular, the expression

$$[\hat{x},\hat{p}] = i\hbar$$

is always true. This allows us to find $\hat{x}$ easily; it must be

$$\hat{x} = i\hbar\frac{d}{dp}$$

Note the positive sign here, as the commutator depends on order. To check the form of $\hat{x}$, then

$$[\hat{x},\hat{p}]\Psi(p) = i\hbar\frac{d}{dp}(p\Psi) - pi\hbar\frac{d}{dp}(\Psi) = i\hbar a + i\hbar\frac{d}{dp}(\Psi) - pi\hbar\frac{d}{dp}(\Psi) = i\hbar\Psi$$

so that

$$[\hat{x},\hat{p}] = i\hbar$$

as required. Hence, in the momentum repesentation, then the momentum eigenstates are

$$\delta(p_m - p)$$

and the position eigenstates are

$$\frac{1}{\sqrt{2\pi\hbar}}e^{-ipx/\hbar}$$

This is why we have been writing $\hat{x}$ rather than just x in the rest of the course; while they are the same in the position representation, this is not true in other representations so we do sometimes need to specify that we mean the operator.

To check this is all consistent, then we can calculate the position probability density from the overlap of these position eigenstates with the wavefunction $\Psi(p)$, which gives

$$a(x) = \frac{1}{\sqrt{2\pi\hbar}}\int \Psi(p)e^{ipx/\hbar}\,dp = \psi(x)$$

i.e. the same expression as we started with.

Dirac notation is actually more than just a shorthand; the kets stand for the states themselves, meaning they are independent of any particular representation. A physical state is clearly the same no matter how we choose to write it mathematically. E.g. if the ground state is $u_0(x)$ then this can be written as $|u_0\rangle$ or even $|0\rangle$, which does not specify whether it is a function of x or p or any other variable. To get the position amplitudes, we need to do the overlap integral, which in Dirac notation is

$$\psi(x) = \langle x|0\rangle$$

where $|x\rangle$ are the x eigenstates. Similarly

$$\Psi(p) = \langle p|0\rangle$$

These give the right expressions irrespective of whether the actual calculation is done in the position or the momentum representation so they are also representation independent.

Finally, we would need other operators, such as the Hamiltonian, if we want to work in the momentum representation. This would formally be

$$\hat{H} = \frac{\hat{p}^2}{2m} + \hat{V}(\hat{x}) = \frac{p^2}{2m} + \hat{V}\left(i\hbar\frac{d}{dp}\right)$$

For example, the SHO potential would give

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega_0^2\hat{x}^2 = \frac{p^2}{2m} - \frac{1}{2}m\hbar^2\omega_0^2\frac{d^2}{dp^2}$$

which is mathematically very similar to the position representation.

2 SHO ladder operators

Let's now look at the SHO again but in a way which is independent of the representation. It would seem that this would be not very useful as we cannot have specific forms for any operators. However, it turns out we can get a long way with these methods.

We saw the Hamiltonian for the harmonic oscillator was

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega_0^2\hat{x}^2}{2} = \frac{1}{2}\left(\frac{\hat{p}^2}{m} + m\omega_0^2\hat{x}^2\right) = \hbar\omega_0\left(\frac{1}{2m\hbar\omega_0}\hat{p}^2 + \frac{m\omega_0}{2\hbar}\hat{x}^2\right)$$

and we found the energy eigenvalues were

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega_0$$

with the ground state being

$$u_0 \propto e^{-m\omega_0 x^2/2\hbar}$$

Consider the operators

$$\hat{a} = \sqrt{\frac{m\omega_0}{2\hbar}}\hat{x} + i\sqrt{\frac{1}{2m\hbar\omega_0}}\hat{p}, \qquad \hat{a}^\dagger = \sqrt{\frac{m\omega_0}{2\hbar}}\hat{x} - i\sqrt{\frac{1}{2m\hbar\omega_0}}\hat{p}$$

Note, these are not Hermitian and so do not correspond to observables; the particle has no "a-like" quantity which we can measure. It doesn't mean they aren't useful as mathematical tools though, as we shall see. Consider

$$\hat{a}\hat{a}^\dagger = \frac{m\omega_0}{2\hbar}\hat{x}^2 + \frac{1}{2m\hbar\omega_0}\hat{p}^2 - \frac{i}{2\hbar}\hat{x}\hat{p} + \frac{i}{2\hbar}\hat{p}\hat{x} = \frac{1}{\hbar\omega_0}\hat{H} - \frac{i}{2\hbar}[\hat{x},\hat{p}] = \frac{\hat{H}}{\hbar\omega_0} + \frac{1}{2}$$

Similarly

$$\hat{a}^\dagger\hat{a} = \frac{\hat{H}}{\hbar\omega_0} - \frac{1}{2}$$

Hence, subtracting and adding gives

$$[\hat{a},\hat{a}^\dagger] = 1, \qquad \{\hat{a},\hat{a}^\dagger\} = \frac{2\hat{H}}{\hbar\omega_0}$$

Note, we can write the Hamiltonian in one of three forms

$$\hat{H} = \hbar\omega_0\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) = \hbar\omega_0\left(\hat{a}\hat{a}^\dagger - \frac{1}{2}\right) = \frac{\hbar\omega_0}{2}(\hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger)$$

Now, consider the commutator of $\hat{a}$ with $\hat{H}$ where $\hat{H}$ is expressed using the first of the above three forms

$$[\hat{H},\hat{a}] = \hbar\omega_0[\hat{a}^\dagger\hat{a},\hat{a}] + \frac{1}{2}\hbar\omega_0[1,\hat{a}] = \hbar\omega_0(\hat{a}^\dagger\hat{a}\hat{a} - \hat{a}\hat{a}^\dagger\hat{a}) + 0 = \hbar\omega_0[\hat{a}^\dagger,\hat{a}]\hat{a} = -\hbar\omega_0\hat{a}$$

Similarly

$$[\hat{H},\hat{a}^\dagger] = \hbar\omega_0\hat{a}^\dagger$$

This now allows us to get from one eigenstate to the next. Let's say we have some eigenstate u_n

$$\hat{H}u_n = E_n u_n$$

Consider

$$-\hbar\omega_0(\hat{a}u_n) = [\hat{H},\hat{a}]u_n = \hat{H}\hat{a}u_n - \hat{a}\hat{H}u_n = \hat{H}(\hat{a}u_n) - E_n(\hat{a}u_n)$$

so

$$\hat{H}(\hat{a}u_n) = (E_n - \hbar\omega_0)(\hat{a}u_n)$$

Hence, we have shown that $\hat{a}u_n$ is an eigenstate of $\hat{H}$ with an energy eigenvalue $E_n - \hbar\omega_0$. Hence, applying the $\hat{a}$ operator has converted u_n to u_{n-1}. The operator $\hat{a}$ is called a *lowering operator*. A similar calculation shows

$$\hat{H}(\hat{a}^\dagger u_n) = (E_n + \hbar\omega_0)(\hat{a}^\dagger u_n)$$

so $\hat{a}^\dagger$ converts u_n to u_{n+1} and is called a *raising operator*. The two operators are generically refered to as *ladder operators*, as they step up or down the "ladder" of energy eigenvalues.

We know the states cannot continue downwards for ever as there is a lowest energy state. This means for the lowest state that applying the lowering operator must fail; specifically, it must give zero

$$\hat{a}u_0 = 0$$

However, we can apply $\hat{a}^\dagger$ to this equation to give

$$0 = \hat{a}^\dagger \hat{a} u_0 = \left(\frac{\hat{H}}{\hbar\omega_0} - \frac{1}{2}\right) u_0 = \left(\frac{E_0}{\hbar\omega_0} - \frac{1}{2}\right) u_0$$

For this to hold, then

$$\frac{E_0}{\hbar\omega_0} - \frac{1}{2} = 0$$

so

$$E_0 = \frac{\hbar\omega_0}{2}$$

which we indeed know is the ground state energy. Hence as each energy is increased using $\hat{a}^\dagger$ by $\hbar\omega_0$, we have shown the energy spectrum is

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega_0$$

Note that the whole of the derivation of the energy eigenvalues above only used the commutation properties and we did not have to use $\hat{p} = -i\hbar d/dx$ or write down a wavefunction explicitly at all. Since commutors are the same in any representation, this means the derivation above is representation independent. It is of course always better to prove things as generally as possible, and this is a specific example of doing this.

Because we can convert from one state to the next, this means that if we can solve for any one of the eigenstates, we can calculate all of the others by repeated application of $\hat{a}$ or $\hat{a}^\dagger$. To find explicit eigenstates, we have to work in a specific representation, of course, so in the usual coordinate representation, then consider the equation above

$$0 = \hat{a} u_0 = \left(\sqrt{\frac{m\omega_0}{2\hbar}}\hat{x} + i\sqrt{\frac{1}{2m\hbar\omega_0}}\hat{p}\right) u_0 = \sqrt{\frac{m\omega_0}{2\hbar}} x u_0 + \sqrt{\frac{\hbar}{2m\omega_0}}\frac{du_0}{dx}$$

This is easy to solve; rearranging gives

$$\frac{du_0}{dx} = -\frac{m\omega_0}{\hbar} x u_0$$

so

$$\frac{du_0}{u_0} = -\frac{m\omega_0}{\hbar} x\, dx$$

for which

$$\ln u_0 = -\frac{m\omega_0}{2\hbar} x^2 + C$$

and so

$$u_0 = e^C e^{-m\omega_0 x^2/2\hbar}$$

where the e^C becomes the normalisation factor. This is what we found previously. We can get higher states by applying $\hat{a}^\dagger$ multiple times to u_0, e.g. $u_1 \propto \hat{a}^\dagger u_0$, but note the ladder operators do not give normalised eigenstates; they have to be normalised afterwards.

Second Year Quantum Mechanics - Lecture 20
Time independent perturbation theory

Paul Dauncey, 22 Nov 2011

1 TISE perturbations

It is often the case that we cannot solve a TISE exactly and so need to take some approximations. The *perturbation* approach is used when solutions to a similar Hamiltonian are known and we can treat the difference of the two as a small effect. Hence, we write the total Hamiltonian as

$$\hat{H}' = \hat{H}_0 + \hat{H}_1$$

where the eigenstates and eigenvalues of $\hat{H}_0$ are known

$$\hat{H}_0 u_n = E_n u_n$$

The extra part, $\hat{H}_1$, can be treated as small, in some sense; this extra term is said to "perturb" the original Hamiltonian solutions. We can write the full solutions as

$$E'_n = E_n + \delta E_n, \qquad u'_n(x) = u_n(x) + \delta u_n(x)$$

where the changes from the original solutions, δE_n and $\delta u_n(x)$, are taken to be small. In this case, it makes sense to do an expansion in powers of the correction to get a "perturbation series".

2 First order approximation

We are trying to solve

$$\hat{H}' u'_n = E'_n u'_n$$

The first order approximation to solve for the energy E'_n is to take

$$u'_n \approx u_n$$

In this approximation, the equation becomes

$$(\hat{H}_0 + \hat{H}_1) u_n = E'_n u_n$$

so

$$E_n u_n + \hat{H}_1 u_n = E'_n u_n$$

Multiplying by u_n^* and integrating gives

$$E_n \int u_n^* u_n \, dx + \int u_n^* \hat{H}_1 u_n \, dx = E'_n \int u_n^* u_n \, dx$$

and since the states are normalised, then this gives

$$E'_n = E_n + \int u_n^* \hat{H}_1 u_n \, dx$$

i.e.

$$\delta E_n = \int u_n^* \hat{H}_1 u_n \, dx = \langle H_1 \rangle_n$$

Hence, the first order correction to the energy is the expectation value of the extra Hamiltonian term.

To solve for the states u'_n, then we do something similar; namely, take the approximation that $E'_n \approx E_n$. In this case the equation becomes

$$(\hat{H}_0 + \hat{H}_1)(u_n + \delta u_n) = E_n(u_n + \delta u_n)$$

which gives

$$E_n u_n + \hat{H}_0 \delta u_n + \hat{H}_1 u_n + \hat{H}_1 \delta u_n = E_n u_n + E_n \delta u_n$$

The first term on each side cancels and the fourth term on the LHS is the small Hamiltonian term acting on the small correction and so is in fact second order and can be neglected in this approximation. Hence

$$(E_n - \hat{H}_0)\delta u_n = \hat{H}_1 u_n$$

is the equation which δu_n must satisfy. In order to express this in a more useful way, then remembering that the u_n form a CONS, then we can always write the change as an expansion

$$\delta u_n(x) = \sum_{k \neq n} a_{nk} u_k(x)$$

where $k \neq n$ as u_n is the original state and is already included to this order. Using this expansion

$$\hat{H}_0 \delta u_n(x) = \sum_k a_{nk} \hat{H}_0 u_k(x) = \sum_k a_{nk} E_k u_k(x)$$

Hence

$$\sum_k a_{nk}(E_n - E_k) u_k = \hat{H}_1 u_n$$

Multiplying by u_m^* and integrating gives

$$\sum_k a_{nk}(E_n - E_k) \int u_m^* u_k \, dx = \int u_m^* \hat{H}_1 u_n \, dx$$

Using orthonormality, this becomes

$$a_{nm}(E_n - E_m) = \int u_m^* \hat{H}_1 u_n \, dx$$

or

$$a_{nm} = \frac{\int u_m^* \hat{H}_1 u_n \, dx}{E_n - E_m}$$

Hence, the total eigenstate is a mixture of the original state plus small amounts of the other states, according to the values of the a_{nm} above. One thing to note is that the states are weighted by $1/(E_n - E_m)$ so that states close in energy will tend to be mixed in more than states further apart in energy.

3 Higher order approximations

The above procedure can be continued to higher orders. For the second order approximation, then the above procedure can be repeated with the state or energy being assumed to be, not just u_n or E_n as above, but the corrected values just derived. The next order approximation is then done in an equivalent way, etc. This can be repeated until the required accuracy is obtained.

Second Year Quantum Mechanics - Lecture 21
QM in two dimensions

Paul Dauncey, 24 Nov 2011

1 Three dimensions

Generally, the basic changes in the quantities we use in going from 1D to 3D are as follows.

	1D	3D
Position operator	$\hat{x}$	$\hat{\boldsymbol{r}} = \hat{x}\boldsymbol{i} + \hat{y}\boldsymbol{j} + \hat{z}\boldsymbol{k}$
Momentum operator	$\hat{p} = -i\hbar\frac{d}{dx}$	$\hat{\boldsymbol{p}} = \hat{p}_x\boldsymbol{i} + \hat{p}_y\boldsymbol{j} + \hat{p}_z\boldsymbol{k} = -i\hbar\left(\frac{\partial}{\partial x}\boldsymbol{i} + \frac{\partial}{\partial y}\boldsymbol{j} + \frac{\partial}{\partial z}\boldsymbol{k}\right) = -i\hbar\boldsymbol{\nabla}$
Kinetic energy	$\hat{T} = \frac{\hat{p}^2}{2m} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}$	$\hat{T} = \frac{\hat{p}^2}{2m} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) = -\frac{\hbar^2}{2m}\nabla^2$
Potential energy	$V(x)$	$V(\boldsymbol{r})$
Hamiltonian	$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)$	$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r})$
Wavefunction	$\psi(x,t)$	$\psi(\boldsymbol{r},t)$
Normalisation	$\int \lvert\psi(x,t)\rvert^2\,dx = 1$	$\int \lvert\psi(\boldsymbol{r},t)\rvert^2\,dx\,dy\,dz = \int \lvert\psi(\boldsymbol{r},t)\rvert^2\,d^3r = 1$

There are two new effects we shall look at, degeneracy and angular momentum and we shall use the example of the 2D SHO to illustrate them.

2 The two dimensional SHO

Although to describe the real world we clearly need to work in three dimensions, we shall start by looking at two dimensions. This gives most of the extra QM effects compared to one dimension but mathematically, two dimensions is simpler than three.

We start by writing down the energy eigenstate equation, i.e. the TISE, of the 2D SHO

$$\hat{H}u = \left[-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) + \frac{m\omega_x^2 x^2}{2} + \frac{m\omega_y^2 y^2}{2}\right]u = Eu$$

We will try using separation of variables, as we did for the time dependent Schrödinger equation where $\psi(x,t)$. Here we shall write

$$u(x,y) = X(x)Y(y)$$

so that the above equation becomes

$$-\frac{\hbar^2}{2m}\left(Y\frac{d^2X}{dx^2} + X\frac{d^2Y}{dy^2}\right) + \frac{m\omega_x^2 x^2}{2}XY + \frac{m\omega_y^2 y^2}{2}XY = EXY$$

Dividing throughout by XY gives

$$\left(-\frac{\hbar^2}{2m}\frac{1}{X}\frac{d^2X}{dx^2}++\frac{m\omega_x^2x^2}{2}\right)+\left(-\frac{\hbar^2}{2m}\frac{1}{Y}\frac{d^2Y}{dy^2}+\frac{m\omega_y^2y^2}{2}\right)=E$$

The first set of terms depend on x only and the second set on y only, so these must be constants; let these be E_x and E_y, respectively. Note $E=E_x+E_y$ so only one of these is really independent. Using this, then

$$-\frac{\hbar^2}{2m}\frac{1}{X}\frac{d^2X}{dx^2}+\frac{m\omega_x^2x^2}{2}=E_x$$

so

$$-\frac{\hbar^2}{2m}\frac{d^2X}{dx^2}+\frac{m\omega_x^2x^2}{2}X=E_xX$$

and similarly for Y. However, the above is the one dimensional SHO energy eigenstate equation. Hence, the two dimensions act as independent one dimensional SHO systems. This means we immediately know the eigenvalues

$$E_x=\left(n_x+\frac{1}{2}\right)\hbar\omega_x,\qquad E_y=\left(n_y+\frac{1}{2}\right)\hbar\omega_y$$

and hence the total energy is

$$E=E_x+E_y=\left(n_x+\frac{1}{2}\right)\hbar\omega_x+\left(n_y+\frac{1}{2}\right)\hbar\omega_y$$

and the eigenstate is

$$u_{n_xn_y}=u_{n_x}(x)u_{n_y}(y)$$

where the u_n are the 1D states. We find the energy depends on two quantum numbers, n_x and n_y, rather than just one so we need to label the eigenstates $u_{n_xn_y}$. This is very general and is due to moving to 2D; it is not particular to the SHO. In 3D, this means we need three quantum numbers to label eigenstates, e.g. u_{nlm}, as we shall see in the later lectures.

3 Degeneracy

What happens if the two dimensions have the same potential, i.e. $\omega_x=\omega_y=\omega_0$? The potential takes the form

$$V(x,y)=\frac{m\omega_x^2x^2}{2}+\frac{m\omega_y^2y^2}{2}=\frac{m\omega_0^2(x^2+y^2)}{2}=\frac{m\omega_0^2r^2}{2}=V(r)$$

and so becomes a central potential, i.e. it has circular symmetry. Also, the energies are given by

$$E=E_x+E_y=(n_x+n_y+1)\,\hbar\omega_0$$

The ground state clearly has $n_x=n_y=0$, for which $E_0=\hbar\omega_0$. However, there are now two first excited states, given by $n_x=1$, $n_y=0$ and $n_x=0$, $n_y=1$, both of which have $E_1=2\hbar\omega_0$. This effect, whereby there is more than one eigenstate with the same eigenvalue, is called *degeneracy.* This clearly results from the symmetry of the potential, i.e. because the value of ω_0 is the same for x and y. This is a general result; a symmetry of the system gives rise to degeneracy in the eigenstates. In three dimensions, there is often spherical symmetry of potentials, e.g. in the hydrogen atom, and so we will see there is a lot of degeneracy of the energies in such systems.

There is one critical result which needs to be emphasised about degeneracy. For the two first excited states above, let u_{10} be the first and u_{01} the second. These satisfy

$$\hat{H}u_{10}=E_1u_{10},\qquad \hat{H}u_{01}=E_1u_{01}$$

However, consider the superposition

$$\psi = \alpha u_{10} + \beta u_{01}$$

for any constants α and β, for which

$$\hat{H}\psi = \alpha\hat{H}u_{10} + \beta\hat{H}u_{01} = \alpha E_1 u_{10} + \beta E_1 u_{01} = E_1\left(\alpha u_{10} + \beta u_{01}\right) = E_1\psi$$

Hence *any superposition of degenerate eigenstates is also a degenerate eigenstate.* This means we can choose any relevant combinations that we wish, depending on the problem in hand. Note, this is true of degenerate eigenstates for any operator, not just the Hamiltonian.

4 Angular momentum

In two dimensions, there is only one component of angular momentum

$$L = xp_y - yp_x$$

so angular momentum is a scalar, rather than a vector, in two dimensions. However, it is clear this scalar is the equivalent of the L_z component in three dimensions. We want to look at the QM operator, which is

$$\hat{L} = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x = -i\hbar x\frac{\partial}{\partial y} + i\hbar y\frac{\partial}{\partial x}$$

Let's see the effect of this operator on the u_{10} state for the case of $\omega_x = \omega_y = \omega_0$, when the potential is central and so we expect angular momentum to be conserved. A reminder that the 1D solutions are

$$u_0 = Ae^{-ax^2/2}, \qquad u_1 = Bxe^{-ax^2/2}$$

Hence, since $\omega_x = \omega_y$ then the a constant is the same for x and y so

$$u_{10} = \left(Bxe^{-ax^2/2}\right)\left(Ae^{-ay^2/2}\right) = ABxe^{-a(x^2+y^2)/2}$$

and similarly

$$u_{01} = ABye^{-a(x^2+y^2)/2}$$

For u_{10}, then

$$\frac{\partial u_{10}}{\partial x} = ABe^{-a(x^2+y^2)/2} - ABax^2e^{-a(x^2+y^2)/2}$$

while

$$\frac{\partial u_{10}}{\partial y} = -ABaxye^{-a(x^2+y^2)/2}$$

Hence

$$\begin{aligned}\hat{L}u_{10} &= i\hbar ABax^2ye^{-a(x^2+y^2)/2} + i\hbar ABye^{-a(x^2+y^2)/2} - i\hbar ABax^2ye^{-a(x^2+y^2)/2}\\ &= i\hbar ABye^{-a(x^2+y^2)/2} = i\hbar u_{01}\end{aligned}$$

Similarly

$$\hat{L}u_{01} = -i\hbar u_{10}$$

This means neither u_{10} nor u_{01} are eigenstates of $\hat{L}$, even though we would expect L to be conserved classically. This seems to contradict our expectations. However, the above can be written as

$$\hat{L}u_{10} = i\hbar u_{01}, \qquad i\hat{L}u_{01} = \hbar u_{10}$$

Hence, adding these two equations gives

$$\begin{aligned}\hat{L}u_{10} + i\hat{L}u_{01} &= \hbar u_{10} + i\hbar u_{01}\\ \hat{L}(u_{10} + iu_{01}) &= \hbar(u_{10} + iu_{01})\end{aligned}$$

Hence, $u_{10}+iu_{01}$ is an eigenstate of $\hat{L}$ with eigenvalue $\hbar$. Similarly, by subtracting the equations, we get

$$\hat{L}(u_{10} - iu_{01}) = -\hbar(u_{10} - iu_{01})$$

which means this combination is an eigenstate with eigenvalue $-\hbar$. Remember, these states are linear combinations of the u_{01} and u_{10} and so because of degeneracy are therefore also eigenstates of energy. Hence, there are states which are indeed eigenstates of both energy and angular momentum, as we would expect, and this means the two operators must be compatible. It must be the case that $[\hat{H}, \hat{L}] = 0$ when $\omega_x = \omega_y$, as can be shown by direct calculation.

It is illustrative to look at this in terms of plane (circular) polar coordinates. For these, then

$$x = r\cos\phi, \qquad y = r\sin\phi$$

Consider $\partial/\partial\phi$, which is found by a standard calculation

$$\frac{\partial}{\partial\phi} = \frac{\partial x}{\partial\phi}\frac{\partial}{\partial x} + \frac{\partial y}{\partial\phi}\frac{\partial}{\partial y} = -r\sin\phi\frac{\partial}{\partial x} + r\cos\phi\frac{\partial}{\partial y} = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}$$

Hence

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

This looks intuitive; for the (linear) coordinate x, the (linear) momentum operator associated with it is $\hat{p} = -i\hbar\partial/\partial x$, while we have just found that for the angular coordinate ϕ, the associated angular momentum operator is $\hat{L} = -i\hbar\partial/\partial\phi$. The eigenstates and eigenvalues of angular momentum are easily found, exactly as for the linear momentum. Writing the eigenstate equation as

$$-i\hbar\frac{d\psi}{d\phi} = m_l\hbar\psi$$

so

$$\frac{d\psi}{\psi} = im_l\phi$$

so

$$\psi = C(r)e^{im_l\phi}$$

where $C(r)$ is any function of the radial coordinate r; it appears as the equivalent of a constant of integration as it is not a function of ϕ.

Unlike for the momentum case, there is a boundary condition, namely that the eigenstate has to be single-valued, i.e. adding 2π to ϕ must give the same state, so $\psi(\phi+2\pi) = \psi(\phi)$. This means

$$e^{im_l(\phi+2\pi)} = e^{im_l\phi}e^{im_l2\pi} = e^{im_l\phi}$$

so

$$e^{im_l2\pi} = 1$$

so this restricts m_l to be an integer, which can be either positive or negative. This means angular momentum eigenstates are discrete, *not* continuous.

The $\hat{L}$ eigenstates $u_{10} \pm i u_{01}$ found above had eigenvalues of $\pm\hbar$ and therefore must correspond to $m_l = \pm 1$. Explicitly

$$\begin{aligned} u_{10} \pm i u_{01} &= AB(x \pm iy)e^{-a(x^2+y^2)/2} = AB(r\cos\phi \pm ir\sin\phi)e^{-ar^2/2} \\ &= ABr(\cos\phi \pm i\sin\phi)e^{-ar^2/2} = ABre^{-ar^2/2}e^{\pm i\phi} \end{aligned}$$

which is exactly the form we found for $m_l = \pm 1$.

Finally, let's put in the time dependence explicitly; for an angular momentum eigenstate

$$C(r)e^{im_l\phi}e^{-iEt/\hbar} = C(r)e^{i(\hbar m_l\phi - Et)/\hbar} = C(r)e^{i(L\phi - Et)/\hbar}$$

The r part does not depend on time and so is constant. The ϕ part looks just like our previous travelling wave $e^{i(px-Et)/\hbar}$ except it is travelling in ϕ not x, i.e. it is going round in a circle, either in the positive ϕ direction if $m_l > 0$ or the negative ϕ direction if $m_l < 0$. Hence, it physically corresponds to waves circulating around the origin and hence to angular momentum.

Second Year Quantum Mechanics - Lecture 22
Degeneracy

Paul Dauncey, 29 Nov 2011

1 Introduction

Up until the previous lecture, all the bound energy eigenstates have had different energy eigenvalues. However, we saw last time that for the 2D SHO, this is not the case. This is also true for 3D; we will see that many distinct eigenstates have the same energy. Degeneracy forces us to re-examine some of our previous results as we assumed no degeneracy in deriving them. In particular, we need to check for orthogonality and the meaning of compatibility. We will also look at the collapse of the wavefunction when measuring degenerate states and the symmetry properties of degenerate energy eigenstates.

2 Orthogonality

We have used the fact that the 1D eigenstates of any Hermitian operator are orthogonal, i.e.

$$\int \phi_m^* \phi_n \, dx = \delta_{mn}$$

many times and indeed many of our results would not hold if this was not true. However, to prove orthogonality, we had to assume the eigenstates had different eigenvalues. The proof was given in Lecture 9 and rested on having derived

$$(q_n - q_m) \int \phi_m^* \phi_n \, dx = 0$$

and so, for $m \neq n$, concluding that as long as $q_n \neq q_m$, then the integral had to be zero. However, if the states can be degenerate, then q_n can equal q_m so we cannot be sure the integral is zero any more. Specifically, this means that if a set of eigenstates have the same eigenvalue, they may not necessarily be orthogonal. This looks like a disaster as we would have to go back to the beginning and rederive all our results allowing for degeneracy.

However, if states are degenerate, it turns out it is always possible to combine them to make a new set of eigenstates which *are* orthogonal. As we saw, these new states are still degenerate eigenstates, with the same eigenvalue, and so they are equally valid to use. We can then continue as before, with all our previous results still being valid. This procedure to make the states orthogonal is called the *Schmidt orthogonalisation procedure.*

Let's first draw a parallel with normal vectors in 2D. Let's say we have two non-parallel unit vectors, $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$. We could use these to describe any vector in space, but if they are not orthogonal to each other, it can be a bit painful. Hence, it would be useful to change to using $\boldsymbol{n}_1' = \boldsymbol{n}_1$ and $\boldsymbol{n}_2'$, where the $\boldsymbol{n}_i'$ are constructed to be orthogonal to each other.

The result can be obtained algebraically by saying that $\boldsymbol{n}_2'$ will lie in the 2D plane of $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$, and so can be constructed by some linear sum of $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$

$$\boldsymbol{n}_2' = a_1 \boldsymbol{n}_1 + a_2 \boldsymbol{n}_2$$

It needs to be orthogonal to $\boldsymbol{n}_1$ so

$$\boldsymbol{n}_1.\boldsymbol{n}_2' = 0 = a_1 \boldsymbol{n}_1.\boldsymbol{n}_1 + a_2 \boldsymbol{n}_1.\boldsymbol{n}_2 = a_1 + a_2 \boldsymbol{n}_1.\boldsymbol{n}_2$$

Hence

$$a_1 = -a_2(\boldsymbol{n}_1.\boldsymbol{n}_2)$$

and so

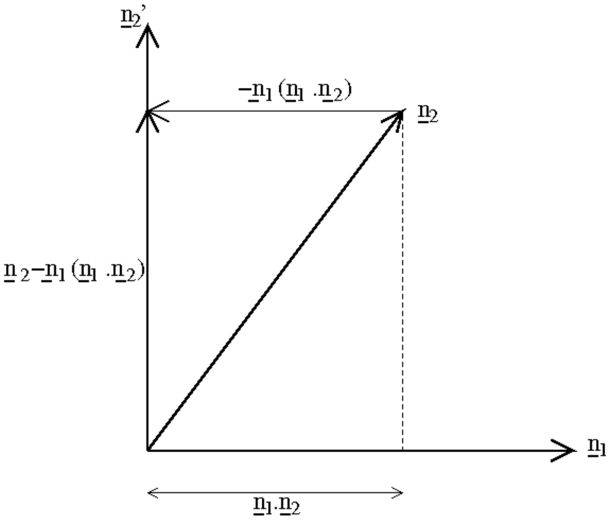
$$\boldsymbol{n}_2' = a_2\left[\boldsymbol{n}_2 - (\boldsymbol{n}_1.\boldsymbol{n}_2)\boldsymbol{n}_1\right]$$

where a_2 is fixed by the requirement that it should be a unit vector.

We can interpret this result geometrically; the length of the projection of $\boldsymbol{n}_2$ onto $\boldsymbol{n}_1$ is $\boldsymbol{n}_1.\boldsymbol{n}_2$ and so the vector component of $\boldsymbol{n}_2$ parallel to $\boldsymbol{n}_1$ is $\boldsymbol{n}_1(\boldsymbol{n}_1.\boldsymbol{n}_2)$. Subtracting this from $\boldsymbol{n}_2$ will give the vector component of $\boldsymbol{n}_2$ perpendicular to $\boldsymbol{n}_1$, which is what we want

$$\boldsymbol{n}_2' \propto \boldsymbol{n}_2 - \boldsymbol{n}_1(\boldsymbol{n}_1.\boldsymbol{n}_2)$$

It is only proportional as it will in general not be a unit vector and so the length will need to be scaled to one.



If we have three vectors, then there is a further $\boldsymbol{n}_3$ from which we can make a $\boldsymbol{n}_3'$ orthogonal to the other two. This could clearly be continued for as many vectors as required if we were in more than 3D.

The Schmidt orthogonalisation procedure is analogous to this. Let's say we have a set of degenerate non-orthogonal eigenstates ϕ_n and we want to construct an orthogonal set ϕ_n'. We can generally choose

$$\phi_1' = \phi_1$$

as there is nothing else for this to be orthogonal to yet. We then make ϕ_2' out of a combination of ϕ_1 and ϕ_2. Putting

$$\phi_2' = a_1\phi_1 + a_2\phi_2$$

then the requirement that it is orthogonal to ϕ_1' gives

$$\int \phi_1'^*\phi_2'\,dx = 0 = \int a_1\phi_1^*\phi_1\,dx + a_2\int \phi_1^*\phi_2\,dx = a_1 + a_2\int \phi_1^*\phi_2\,dx$$

Hence

$$a_1 = -a_2\int \phi_1^*\phi_2\,dx$$

and so

$$\phi_2' = a_2\left[\phi_2 - \left(\int \phi_1^*\phi_2\,dx\right)\phi_1\right]$$

Clearly ϕ_2' reduces to ϕ_2 if the original eigenstates are orthogonal as the integral is then zero. A similar procedure then fixes ϕ_3' in terms of ϕ_1, ϕ_2 and ϕ_3, with two constraints from requiring orthogonality with both ϕ_1' and ϕ_2'. This can be continued until the complete set of eigenstates have been orthogonalised.

While we rarely need to actually *do* the Schmidt orthogonalisation procedure, it is important that is can always be done in general, as this means we can always construct orthogonal states. Hence all our previous results can be made to hold, even for the case of degeneracy.

3 Compatibility

We discussed the 2D symmetric SHO in the last lecture. If we had done the calculation, we would have found that $[\hat{H}, \hat{L}] = 0$ so we might expect $\hat{H}$ and $\hat{L}$ would share eigenstates. However, we found u_{01} and u_{10} were not eigenstates of $\hat{L}$, but that the combinations $u_{01} \pm iu_{10}$ were. This means that although $\hat{H}$ and $\hat{L}$ commute, if the states are degenerate, it is no longer true that the eigenstates are automatically the same for both operators; we saw u_{01} and u_{10} are not. However, it turns out that there is always some combination of the degenerate eigenstates, $u_{01} \pm iu_{10}$ in this case, which *are* eigenstates of both operators.

In Lecture 13, we discussed the relationship between compatibility and commuting operators. We showed that if two operators were compatible (i.e shared all their eigenstates), then they always commuted. However, we also found that if they commuted, we could only deduce they were definitely compatible if there was no degeneracy. We showed that if ϕ_n were eigenstates of a general Hermitian operator $\hat{Q}$

$$\hat{Q}\phi_n = q_n\phi_n$$

and $\hat{Q}$ commuted with another operator $\hat{R}$

$$[\hat{Q}, \hat{R}] = 0$$

then

$$\hat{Q}(\hat{R}\phi_n) = q_n(\hat{R}\phi_n)$$

If q_n is unique to ϕ_n, then since this is another eigenstate equation for $\hat{Q}$, we concluded

$$\hat{R}\phi_n \propto \phi_n \qquad \text{so} \qquad \hat{R}\phi_n = r_n\phi_n$$

and so the operators were compatible. However, with degeneracy, then q_n is not unique; we could have some other state ϕ_m

$$\hat{Q}\phi_m = q_n\phi_m$$

so that it could be that

$$\hat{R}\phi_n \propto \phi_m$$

which means ϕ_n is not an eigenstate of $\hat{R}$. We saw an example of this in the last lecture where we found $\hat{L}u_{10} \propto u_{01}$ for the 2D SHO eigenstates, even though $[\hat{H}, \hat{L}] = 0$.

However, it turns out the degenerate eigenstates form a complete sub-set which can be used to construct *any* other degenerate state $\psi = \sum_n a_n\phi_n$. It also happens that there is always some sum of the degenerate eigenstates which forms an eigenstate of $\hat{R}$ and, since we know any sum of degenerate eigenstates is another eigenstate, then this sum is still an eigenstate of $\hat{Q}$ so

$$\hat{R}\left(\sum_n a_n\phi_n\right) = r\left(\sum_n a_n\phi_n\right)$$

Again, we saw an example of this in the last lecture, where we found $\hat{L}(u_{10}+iu_{01}) = \hbar(u_{10}+iu_{01})$. Hence, a commutator being zero is still a good test of the possibility of compatibility, but any given set of degenerate eigenstates may not exhibit it directly.

4 Collapse

Let's consider an example; a wavefunction is a superposition of the three 2D SHO normalised lowest energy states

$$\psi = \frac{1}{\sqrt{6}}(u_{00} + 2u_{10} + u_{01})$$

where the $1/\sqrt{6}$ is required to ensure the state is normalised. If we now measure the energy, then the result must be either the ground state energy or the first excited state energy as these are the only energy eigenstates in the superposition. The ground state energy is $E_0 = \hbar\omega_0$ and it is clear that the probability of getting this is $|1/\sqrt{6}|^2 = 1/6$. After this measurement, the wavefunction would collapse to $\psi' = u_{00}$; this is the same as we have seen before.

However, degeneracy complicates the other case. For the first excited state, the energy is $E_1 = 2\hbar\omega_0$ and since this is the only other outcome, then its probability must be $1-(1/6) = 5/6$, which could alternatively have been deduced from adding the probabilities for u_{10} (i.e. 4/6) and u_{01} (i.e. 1/6). However, what does the wavefunction collapse to in this case? Nothing has been done to favour u_{10} or u_{01} so would it collapse to either of these, or in fact to something else? To see this, define a new state

$$U_1 = \frac{1}{\sqrt{5}}(2u_{10} + u_{01})$$

Because this is a combination of degenerate eigenstates, it is itself also a degenerate eigenstate, with energy $E_1 = 2\hbar\omega_0$. Writing the original wavefunction in terms of U_1, then

$$\psi = \frac{1}{\sqrt{6}}(u_{00} + \sqrt{5}U_1)$$

It is now clear that a measurement resulting in E_1 has a probability of $|\sqrt{5}/\sqrt{6}|^2 = 5/6$ as before but also that the collapse should be to the part of the wavefunction which has that energy, i.e. the collapse is to $\psi' = U_1$. Hence, the resulting wavefunction will be some combination of u_{10} and u_{01} in general, not either of these separately.

This combination arises because we don't have complete information on the state; each state has two quantum numbers but we have only measured one quantity to try to fix these. If we want a definite wavefunction, then we need to do two measurements to fix both quantum numbers. In this case, following the energy measurement, consider doing an energy difference measurement $\Delta E = E_x - E_y$. Since u_{10} is excited in x while u_{01} is excited in y, the value of ΔE is different for u_{10} and u_{01} and so we can then tell them apart. The eigenstate u_{10} has $\Delta E = \hbar\omega_0$ and given $\psi' = U_1$, then the probability of this is $|2/\sqrt{5}|^2 = 4/5$. If this is the measurement result, the waverfunction will definitely collapse to u_{10}. Alternatively, u_{01} has $\Delta E = -\hbar\omega_0$ and the probability of this is $|1/\sqrt{5}|^2 = 1/5$. With this result, the wavefunction will collapse to u_{01}. Knowing both the total energy and the energy difference restricts us to a particular, well-defined eigenstate.

In general, we need to do two different measurements to "fully" collapse a wavefunction to a well-defined state, because there are two quantum numbers. In 3D, there are three quantum numbers and hence we will normally need to do three measurements to be sure that the wavefunction is in a particular state.

5 Symmetry

We have seen symmetry leads to degeneracy in general. However, degeneracy also complicates the way symmetry is exhibited in the eigenstates.

For the 1D bound energy eigenstates, we always worked with even parity potentials, i.e. potentials which had "mirror" symmetry. We found the eigenstates had a definite parity which was either even or odd; $u(-x) = \pm u(x)$. Indeed, the ground state was always even parity and then the states alternated odd, even, odd, etc., as they went up in energy. However, when taking the probability density, then $|u(-x)|^2 = |u(x)|^2$. Hence, both even and odd eigenstates give an even parity for the probability density, so the probability density always reflected the parity symmetry of the potential.

However, with degeneracy, this simple picture no longer holds. We saw in the last lecture that for the 2D SHO, when the x and y potentials were put equal to each other, then the total potential became

$$V(x,y) = \frac{m\omega^2}{2}x^2 + \frac{m\omega^2}{2}y^2 = \frac{m\omega^2}{2}(x^2+y^2) = \frac{m\omega^2}{2}r^2 = V(r)$$

and so is radially symmetric. This was reflected in the ground state, which was

$$u_{00} = A^2 e^{-a(x^2+y^2)/2}$$

for which

$$|u_{00}|^2 = |A|^4 e^{-ar^2}$$

which shows the same symmetry. However, this state is the only one which is not degenerate. The two first excited states could be written as

$$u_{10} = ABxe^{-a(x^2+y^2)/2}, \qquad u_{01} = ABye^{-a(x^2+y^2)/2}$$

which give

$$|u_{10}|^2 = |AB|^2 x^2 e^{-ar^2}, \qquad |u_{01}|^2 = |AB|^2 y^2 e^{-ar^2}$$

Neither of these probability densities is radially symmetric. However, the sum of the two probability densities gives

$$|u_{10}|^2 + |u_{01}|^2 = |AB|^2 x^2 e^{-ar^2} + |AB|^2 y^2 e^{-ar^2} = |AB|^2(x^2+y^2)e^{-ar^2} = |AB|^2 r^2 e^{-ar^2}$$

and so *is* radially symmetric.

This is a general result and is required physically; if the potential has a symmetry, then there is no reason for the probability density to be asymmetric when considering all the possible solutions. We find that while the probability densities of each of the degenerate eigenstates do not (necessarily) reflect the symmetry of the system, the sum of them does. In 3D, then generally

$$\sum |u_{nlm}(r)|^2$$

will show the symmetry of the system, where the sum is over all the orthogonal degenerate eigenstates of a particular energy value.

Second Year Quantum Mechanics - Lecture 23
Properties of 3D angular momentum operators

Paul Dauncey, 1 Dec 2011

1 Introduction

We now need to look explicitly at the 3D angular momentum operators. Classically

$$\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$$

We know that central potentials, where $V(\boldsymbol{r}) = V(r)$ produce radial forces, i.e. $\boldsymbol{F}$ is always along $\boldsymbol{r}$. The torque is $\boldsymbol{\tau} = \boldsymbol{r} \times \boldsymbol{F}$ so if $\boldsymbol{F}$ is parallel to $\boldsymbol{r}$, then $\boldsymbol{\tau} = 0$. Hence, these potentials give no torque to the particle and hence angular momentum is conserved. Angular momentum plays a central role in classical mechanics and the same is true in quantum mechanics.

2 Angular momentum components

The angular momentum component operators are easily found using Postulate 3; we just substitute $\boldsymbol{r} \to \hat{\boldsymbol{r}}$ and $\boldsymbol{p} \to \hat{\boldsymbol{p}}$, so

$$\hat{\boldsymbol{L}} = \hat{\boldsymbol{r}} \times \hat{\boldsymbol{p}}$$

Explicitly in Cartesians

$$\begin{aligned} \hat{L}_x &= \hat{y}\hat{p}_z - \hat{z}\hat{p}_y \\ \hat{L}_y &= \hat{z}\hat{p}_x - \hat{x}\hat{p}_z \\ \hat{L}_z &= \hat{x}\hat{p}_y - \hat{y}\hat{p}_x \end{aligned}$$

Note the cyclic ordering here, i.e. the components labels cycle round as

$$x \to y \to z \to x \to y \ldots$$

Since this is built into the definition of the $\hat{L}_i$, then it holds for all the results we will derive from the operators also. Hence, we normally only have to derive a result for one component and using the cyclic ordering, we can then find the equivalent result for the other two components.

You may be worried about operator ordering here as each component of $\hat{\boldsymbol{L}}$ has products like $\hat{x}\hat{p}_y$. Consider the commutator for one particular combination

$$\begin{aligned} [\hat{x}, \hat{p}_y]\psi(x, y) &= x\left(-i\hbar\frac{\partial\psi}{\partial y}\right) + i\hbar\frac{\partial}{\partial y}(x\psi) \\ &= -i\hbar x\frac{\partial\psi}{\partial y} + i\hbar x\frac{\partial\psi}{\partial y} = 0 \end{aligned}$$

Compared to $[\hat{x}, \hat{p}_x]$, the extra term which makes the latter $i\hbar$ is missing for the former since $\partial/\partial y$ does not act on x. Hence all such combinations of a coordinate component and a momentum of a different component give a zero commutator, so generally

$$[\hat{r}_i, \hat{p}_j] = i\hbar\delta_{ij}$$

Since only terms with different i and j appear in the $\hat{\boldsymbol{L}}$ components, then these all commute so the ordering doesn't matter and the above definition is unambiguous. Since they commute, this also means that the operator products like $\hat{x}\hat{p}_y$ can be seen to be equal to the anticommutator

$\{\hat{x}, \hat{p}_y\}/2$. We know this is one of the two Hermitian combinations for operators, so we can see the $\hat{\boldsymbol{L}}$ components are all Hermitian.

Note that these operators involve both the spatial and momentum components. If the particle is located in some region of space, so Δx, Δy and Δz are finite, then we know the momentum components will not be defined exactly. Hence, we might expect to have some uncertainty relation between the $\hat{L}_i$ and not be able to measure them all exactly. We therefore need to consider if they are compatible; variables can be compatible, and so can have the same eigenstates and hence have definite values simultaneously, if they commute. Let's see if the angular momentum components can be compatible. Consider

$$[\hat{L}_x, \hat{L}_y] = \hat{L}_x\hat{L}_y - \hat{L}_y\hat{L}_x$$

Explicitly

$$\begin{aligned}
\hat{L}_x\hat{L}_y &= (\hat{y}\hat{p}_z - \hat{z}\hat{p}_y)(\hat{z}\hat{p}_x - \hat{x}\hat{p}_z) \\
&= \hat{y}\hat{p}_x\hat{p}_z\hat{z} - \hat{y}\hat{x}\hat{p}_z^2 - \hat{p}_x\hat{p}_y z^2 + \hat{x}\hat{p}_y\hat{z}\hat{p}_z \\
\hat{L}_y\hat{L}_x &= \hat{y}\hat{p}_x\hat{z}\hat{p}_z - \hat{y}\hat{x}\hat{p}_z^2 - \hat{p}_x\hat{p}_y z^2 + \hat{x}\hat{p}_y\hat{p}_z\hat{z}
\end{aligned}$$

Hence

$$\begin{aligned}
[\hat{L}_x, \hat{L}_y] &= \hat{y}\hat{p}_x[\hat{p}_z, \hat{z}] + \hat{x}\hat{p}_y[\hat{z}, \hat{p}_z] \\
&= -i\hbar\hat{y}\hat{p}_x + i\hbar\hat{x}\hat{p}_y = i\hbar\hat{L}_z
\end{aligned}$$

Hence, these operators are not compatible and so do not have the same eigenstates. If we measure L_x and so collapse the wavefunction to one of its eigenstates, it will *not* be a simultaneous eigenstate of $\hat{L}_y$ and hence will *not* have a definite L_y value. Clearly, there is nothing special about x and y in the above so this holds for the other combinations also; specifically keeping the components in cyclic order

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z, \qquad [\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x, \qquad [\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y$$

Hence at most, only *one* of the L_i components can have a definite value at any one time, i.e. we can find eigenstates of $\hat{L}_x$ or $\hat{L}_y$ or $\hat{L}_z$, but not of any two of them, nor all three of them.

There is a useful way to remember the ordering of the component labels here as these three equations are equivalent to writing

$$i\hbar\hat{\boldsymbol{L}} = \hat{\boldsymbol{L}} \times \hat{\boldsymbol{L}}$$

Explicitly, for e.g. the z component, this gives

$$i\hbar\hat{L}_z = \hat{L}_x\hat{L}_y - \hat{L}_y\hat{L}_x = [\hat{L}_x, \hat{L}_y]$$

as before. This looks a little odd as the cross product of any vector with itself is zero but this is an equation for operators, not simple vectors, and so the operator ordering matters.

3 Angular momentum magnitude

We saw in 2D that the angular momentum operator depended on just the angle ϕ (and not r), so there was one variable and we found the states needed one quantum number, m_l. We will find that in 3D, the angular momentum operators depend on both the angles θ and ϕ (but again, not r). Hence the eigenstates will be functions of θ and ϕ and so with two coordinates, we expect we need two quantum numbers to define such a state. We cannot use two of the three components as they are not compatible so at most one of the three has a well defined quantum number for any state. Hence, we need to find another variable which *is* compatible with at least one of the components to define the other quantum number. Consider

$$\hat{L}^2 = \hat{\boldsymbol{L}}.\hat{\boldsymbol{L}} = \hat{L}_x\hat{L}_x + \hat{L}_y\hat{L}_y + \hat{L}_z\hat{L}_z = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$$

Does this commute with the $\hat{L}_i$? Trying

$$[\hat{L}^2, \hat{L}_x] = [\hat{L}_x^2, \hat{L}_x] + [\hat{L}_y^2, \hat{L}_x] + [\hat{L}_z^2, \hat{L}_x]$$

We know

$$[\hat{L}_x^2, \hat{L}_x] = \hat{L}_x^3 - \hat{L}_x^3 = 0$$

so we need the other two. The first gives

$$\begin{aligned} [\hat{L}_y^2, \hat{L}_x] &= \hat{L}_y^2\hat{L}_x - \hat{L}_x\hat{L}_y^2 = \hat{L}_y\hat{L}_y\hat{L}_x - \hat{L}_y\hat{L}_x\hat{L}_y + \hat{L}_y\hat{L}_x\hat{L}_y - \hat{L}_x\hat{L}_y\hat{L}_y \\ &= \hat{L}_y[\hat{L}_y, \hat{L}_x] + [\hat{L}_y, \hat{L}_x]\hat{L}_y = -i\hbar\hat{L}_y\hat{L}_z - i\hbar\hat{L}_z\hat{L}_y = -i\hbar\{\hat{L}_y, \hat{L}_z\} \end{aligned}$$

Similarly

$$[\hat{L}_z^2, \hat{L}_x] = i\hbar\{\hat{L}_z, \hat{L}_y\} = i\hbar\{\hat{L}_y, \hat{L}_z\}$$

since the anticommutator does not depend on the operator order. Hence, overall

$$[\hat{L}^2, \hat{L}_x] = [\hat{L}_x^2, \hat{L}_x] + [\hat{L}_y^2, \hat{L}_x] + [\hat{L}_z^2, \hat{L}_x] = 0 - i\hbar\{\hat{L}_y, \hat{L}_z\} + i\hbar\{\hat{L}_y, \hat{L}_z\} = 0$$

This clearly holds for any of the other two combinations of $[\hat{L}^2, \hat{L}_i]$ also, so we conclude that we can find simultaneous eigenstates of $\hat{L}^2$ and one of $\hat{L}_x$, $\hat{L}_y$ or $\hat{L}_z$. Hence, we can find the magnitude and one component, conventionally L_z, which then leaves an uncertainty in the direction of the rest of the magnitude $L^2 - L_z^2$ in the other two directions.

The fact that $\hat{L}^2$ can be compatible with all of the $\hat{L}_i$ but the $\hat{L}_i$ are not compatible with each other seems very wierd and, if we did not know about degeneracy, it would appear impossible. If there was no degeneracy among the eigenvalues of these operators, then saying e.g. $[\hat{L}^2, \hat{L}_x] = 0$ would mean all the eigenstates of $\hat{L}^2$ are also eigenstates of $\hat{L}_x$. However, we also have $[\hat{L}^2, \hat{L}_y] = 0$ so these same eigenstates of $\hat{L}^2$ would also be eigenstates of $\hat{L}_y$. The same holds for $\hat{L}_z$ so we would conclude all four operators must have the same eigenstates and so all four are compatible. This of course directly contradicts the fact that $[\hat{L}_i, \hat{L}_j] \neq 0$ for $i \neq j$ and so cannot be true.

Hence, the commutators of this system of operators *require* there to be degeneracy to make sense. There must be some degeneracy of the $\hat{L}^2$ eigenstates so that one combination of these give $\hat{L}_x$ eigenstates, a different combination give $\hat{L}_y$ eigenstates and yet a third different combination give $\hat{L}_z$ eigenstates. All these combinations are of course still eigenstates of $\hat{L}^2$ with the same eigenvalue, as is always true for superpositions of degenerate states. Hence, when we solve for the eigenvalues in the next lecture, we will expect to find e.g. several $\hat{L}_z$ eigenvalues for a given $\hat{L}^2$ eigenvalue. Of course, this physically makes sense; for a given length of $\hat{\boldsymbol{L}}$, then we can have components of different lengths simply by rotating the vector around in space. Of course, we must always have $L_z \leq |\boldsymbol{L}|$ so this would imply that the square of the eigenvalues of $\hat{L}_z$ for the compatible states have to always be equal to or smaller than the (degenerate) eigenvalue of $\hat{L}^2$ for those states. We shall find both of these expectations hold true when we solve for the eigenvalues explicitly in the next lecture.

Second Year Quantum Mechanics - Lecture 24
Eigenvalues of angular momentum

Paul Dauncey, 2 Dec 2011

1 Introduction

In 2D, we found the single angular momentum operator depended on ϕ only when written in plane polar coordinates. In 3D, we will find the angular momentum operators depend on θ and ϕ when written in spherical polar coordinates. Unfortunately, solving the resulting eigenvalue equations for the eigenstates is quite messy.

When we did the 1D SHO, we found that solving for the eigenstates in terms of Hermite polynomials was also messy. However, we saw later that we could find the eigenvalues using only the operators and commutators and it was a lot less complicated.

We will do things in the opposite order when solving for the angular momentum eigenvalues. We will first use the operators and get the eigenvalues without explicitly finding eigenstates in terms of θ and ϕ. Only after this, in the next lecture, will we sketch out the explicit solutions.

2 Angular momentum ladders

We want to solve for the eigenvalues of $\hat{L}^2$ and one of the components, for which we will choose $\hat{L}_z$. These can have shared eigenstates and the usual convention is to write these shared eigenstates as $Y(\theta,\phi)$. Hence, we have

$$\hat{L}^2 Y = \alpha Y, \qquad \hat{L}_z Y = \beta Y$$

where α and β are the eigenvalues we want to find. We saw we could move between the 1D SHO energy states using raising and lowering operators. For angular momentum, we saw in the last lecture that we expect degeneracy and hence several β values for a given α value. It turns out there is a similar set of ladder operators for angular momentum which allow us to move between these β values for a given α value. In the 1D SHO case, there was a lower limit, for which the lowering operator applied to the minimum state gave zero, but there was no upper limit. However, we know that β must be limited to be $\beta^2 \leq \alpha$ so we expect both a lower and an upper limit in the angular momentum case.

We define two (non-Hermitian) operators

$$\hat{L}_+ = \hat{L}_x + i\hat{L}_y, \qquad \hat{L}_- = \hat{L}_x - i\hat{L}_y$$

Consider the commutator

$$[\hat{L}_z, \hat{L}_+] = [\hat{L}_z, \hat{L}_x] + i[\hat{L}_z, \hat{L}_y] = (i\hbar\hat{L}_y) + i(-i\hbar\hat{L}_x) = i\hbar\hat{L}_y + \hbar\hat{L}_x = \hbar\hat{L}_+$$

Similarly

$$[\hat{L}_z, \hat{L}_-] = -\hbar\hat{L}_-$$

These have an identical structure to the relations we saw for the SHO. In that case, we found

$$[\hat{H}, \hat{a}^\dagger] = \hbar\omega\hat{a}^\dagger, \qquad [\hat{H}, \hat{a}] = -\hbar\omega\hat{a}$$

for the raising and lowering operators, respectively. Here, the operator for which we are finding eigenvalues is the first in the commutator ($\hat{H}$ or $\hat{L}_z$), the raising or lowering operators ($\hat{a}^\dagger$ and $\hat{a}$ or $\hat{L}_+$ and $\hat{L}_-$) appear in the commutator and on the right and the multiplier ($\hbar\omega$ or $\hbar$) is the amount the eigenvalue changes by. Given that in both cases the equations have an identical structure, then we know $\hat{L}_+$ and $\hat{L}_-$ will act in the same way. Explicitly, given a state Y with eigenvalue β, then

$$\hbar\hat{L}_+Y = [\hat{L}_z, \hat{L}_+]Y = \hat{L}_z\hat{L}_+Y - \hat{L}_+\hat{L}_zY = \hat{L}_z\hat{L}_+Y - \hat{L}_+\beta Y$$

Hence

$$\hat{L}_z(\hat{L}_+Y) = \beta(\hat{L}_+Y) + \hbar(\hat{L}_+Y) = (\beta+\hbar)(\hat{L}_+Y)$$

Hence, $\hat{L}_+Y$ is a state which has an L_z eigenvalue of $\beta+\hbar$, i.e. $\hat{L}_+$ has incremented the eigenvalue of Y by $\hbar$, as expected. Similarly

$$\hat{L}_z(\hat{L}_-Y) = (\beta-\hbar)(\hat{L}_-Y)$$

decreases the value by $\hbar$. $\hat{L}_+$ and $\hat{L}_-$ are ladder operators which move up and down the β values by one unit of $\hbar$. Note, like $\hat{a}$ and $\hat{a}^\dagger$, they do not give normalised states; the normalisation has to be done by hand afterwards.

How do we know the α value of the state is unchanged? Consider

$$[\hat{L}^2, \hat{L}_+] = [\hat{L}^2, \hat{L}_x] + i[\hat{L}^2, \hat{L}_y]$$

However, we know $\hat{L}^2$ commutes with all the $\hat{L}_i$ so both the above commutators are zero and so $[\hat{L}^2, \hat{L}_+] = 0$. Hence

$$[\hat{L}^2, \hat{L}_+]Y = 0 = \hat{L}^2\hat{L}_+Y - \hat{L}_+\hat{L}^2Y = \hat{L}^2(\hat{L}_+Y) - \alpha(\hat{L}_+Y)$$

Hence

$$\hat{L}^2(\hat{L}_+Y) = \alpha(\hat{L}_+Y)$$

and so the state formed by $\hat{L}_+Y$ does indeed have the same α value as the original Y state. The same holds for $\hat{L}_-Y$. Hence the ladder operators do not change α. If Y is an eigenfunction of $\hat{L}^2$ with eigenvalue α then $\hat{L}_+Y$ and $\hat{L}_-Y$ are *also* eigenfunctions of $\hat{L}^2$ with eigenvalue α. This means the $\hat{L}_+$ and $\hat{L}_-$ keep the magnitude of the angular momentum constant but rotate its direction to increment or decrement, respectively, the z component.

3 Angular momentum eigenvalues

To get the actual eigenvalues, we need to derive another relation. Consider

$$\hat{L}_+\hat{L}_- = (\hat{L}_x + i\hat{L}_y)(\hat{L}_x - i\hat{L}_y) = \hat{L}_x^2 + \hat{L}_y^2 - i[\hat{L}_x, \hat{L}_y] = \hat{L}^2 - \hat{L}_z^2 + \hbar\hat{L}_z$$

Similarly

$$\hat{L}_-\hat{L}_+ = \hat{L}^2 - \hat{L}_z^2 - \hbar\hat{L}_z$$

For the 1D SHO, we found the ground state energy by requiring that the lowering operator gave zero when operating on the ground state, i.e. $\hat{a}u_0 = 0$. For angular momentum, we know there must be a maximum and minimum β for a given α and so in a similar way, the raising operator on the maximum β state must give zero, as must the lowering operator on the minimum β state. Hence, we must have

$$\hat{L}_+Y_{\max} = 0, \qquad \hat{L}_-Y_{\min} = 0$$
$$\hat{L}_zY_{\max} = \beta_{\max}Y_{\max}, \qquad \hat{L}_zY_{\min} = \beta_{\min}Y_{\min}$$

Operating on the first of these equations with $\hat{L}_-$ we find

$$\hat{L}_-\hat{L}_+Y_{\text{max}} = (\hat{L}^2 - \hat{L}_z^2 - \hbar\hat{L}_z)Y_{\text{max}} = 0$$

We therefore find

$$(\alpha - \beta_{\text{max}}^2 - \hbar\beta_{\text{max}})Y_{\text{max}} = 0$$

and since we are looking for solutions for which $Y_{\text{max}} \neq 0$, we need

$$\alpha = \beta_{\text{max}}(\beta_{\text{max}} + \hbar)$$

Similarly

$$\alpha = \beta_{\text{min}}(\beta_{\text{min}} - \hbar) = (-\beta_{\text{min}})[(-\beta_{\text{min}}) + \hbar]$$

These two equations are only consistent if $\beta_{\text{min}} = -\beta_{\text{max}}$.

The raising and lowering operators show that neighbouring values of β are separated by $\hbar$ so we have

$$\beta_{\text{max}} - \beta_{\text{min}} = n\hbar$$

where n is an integer which is greater than or equal to zero, giving the number of raising operations needed to get from Y_{min} to Y_{max}. Since $\beta_{\text{min}} = -\beta_{\text{max}}$, the above means

$$\begin{aligned} \beta_{\text{max}} + \beta_{\text{max}} &= n\hbar \\ 2\beta_{\text{max}} &= n\hbar \\ \beta_{\text{max}} &= n\hbar/2 \end{aligned}$$

and so $\beta_{\text{min}} = -n\hbar/2$.

The value of n can in principle be any non-negative integer. If n is *even*, then we can put $n = 2l$ for integer l and get $\beta_{\text{max}} = l\hbar$ so that

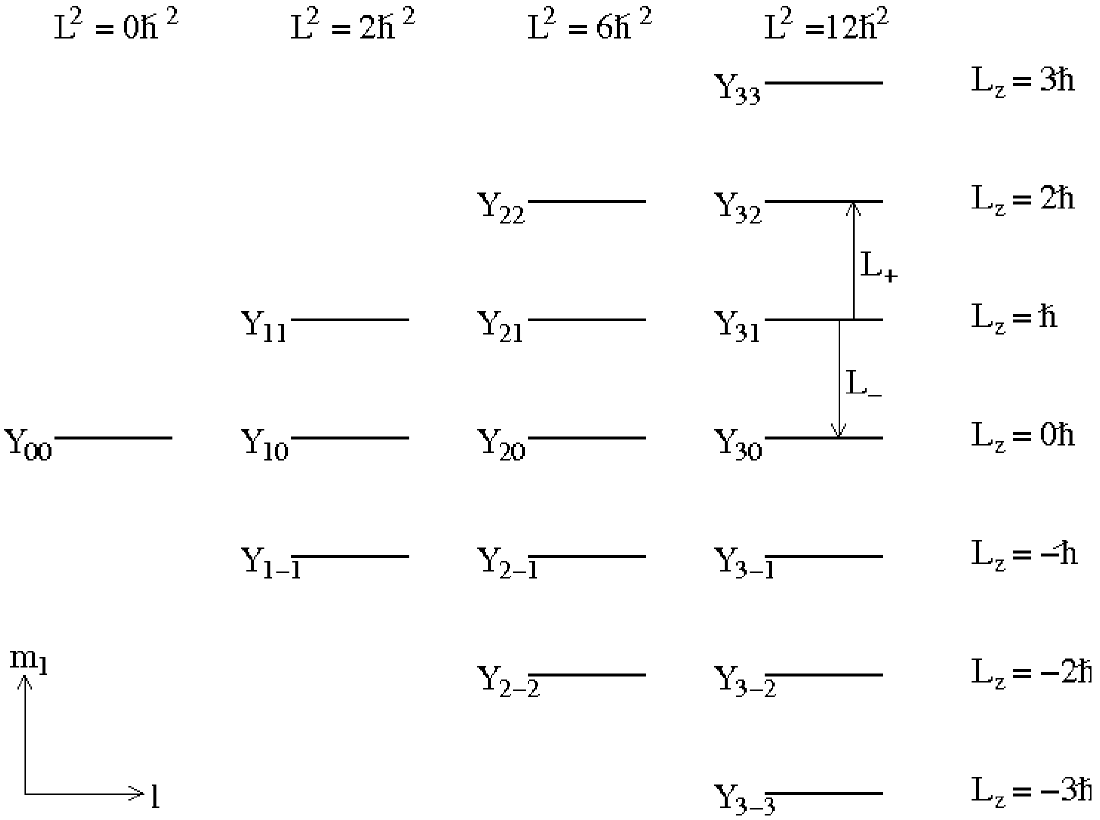
$$\alpha = \beta_{\text{max}}(\beta_{\text{max}} + \hbar) = l\hbar(l\hbar + \hbar) = l(l+1)\hbar^2$$

and

$$\beta = m_l\hbar, \qquad m_l = 0, \pm1, \pm2, \ldots \pm l$$

noting the restriction that $-l \leq m_l \leq l$. There are $2l+1$ different allowed m_l and hence β values so this is the degeneracy of the states of a given l. These β values correspond directly to what we found for the 2D angular momentum case, where we were restricted to integer values of m_l due to the boundary conditions of ϕ wrapping around by 2π.

The eigenstates themselves are labelled as $Y_{l\,m_l}(\theta, \phi)$ and the lowest values of $l = 0$ to 3 are shown in the figure below.



4 Half-integer angular momentum

For the above derivation of the eigenvalues, we took n to be even at the end. The derivation itself has no such restriction on n and it would seem that it could be odd. If this was the case, then $\beta_{\text{max}} = n\hbar/2$ would give half-integer multiples of $\hbar$, rather than integer multiples. These would not satisfy the angular boundary conditions on ϕ and so are not allowed as $\hat{L}_z$ eigenvalues. However, it turns out that these *are* realised physically but only as the eigenvalues for a quantity called "spin", which we shall discuss later in the course.

Second Year Quantum Mechanics - Lecture 25
Angular momentum in spherical polar coordinates

Paul Dauncey, 6 Dec 2011

1 Spherical polar coordinates

We will want to work with central potentials, so we will need to find the angular momentum operators and eigenstates explicitly in spherical polars. This is messy but the component operators turn out to be

$$\begin{aligned}\hat{L}_x &= -i\hbar\left(-\sin\phi\frac{\partial}{\partial\theta} - \cos\phi\cot\theta\frac{\partial}{\partial\phi}\right)\\ \hat{L}_y &= -i\hbar\left(\cos\phi\frac{\partial}{\partial\theta} - \sin\phi\cot\theta\frac{\partial}{\partial\phi}\right)\\ \hat{L}_z &= -i\hbar\frac{\partial}{\partial\phi}\end{aligned}$$

Note there is no r dependence. Note also that $\hat{L}_z$ looks less ugly than the other two and has the same form as in the 2D case. This is not because of any physical difference in x, y and z but because the definition of spherical polars picks out the z axis in an asymmetric way. The angle ϕ is the angle around the z axis but there is no equivalent in spherical polars for the x and y axes. Purely for this reason, it is standard to always pick out the z component as the component which shares eigenstates with $\hat{L}^2$, as we did in the last lecture. We actually already know the ϕ dependence of the $\hat{L}_z$ eigenstates as we solved for that in the 2D case; they are

$$\propto e^{im_l\phi}$$

where m_l must be an integer to keep the wavefunction single-valued for $\phi + 2\pi$. This has an eigenvalue given by

$$\hat{L}_z e^{im_l\phi} = -i\hbar\frac{\partial}{\partial\phi}(e^{im_l\phi}) = m_l\hbar e^{im_l\phi}$$

i.e. the eigenvalue is $m_l\hbar$, as we found in the last lecture.

2 The magitude-squared eigenstates

Using the above components, with some work then the $\hat{L}^2$ operator can be found to be

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

The usual convention for the $\hat{L}^2$ and $\hat{L}_z$ eigenstates is to write them as $Y_{l\,m_l}(\theta,\phi)$, although note some books write this as $Y_l^{m_l}(\theta,\phi)$. In principle, to solve for the $\hat{L}^2$ eigenstates, we would now need to solve the eigenvalue equation

$$\hat{L}^2 Y(\theta,\phi) = \alpha Y(\theta,\phi)$$

using the above operator. This would involve separation of variables to get separate θ and ϕ differential equations. However, we can actually bypass a lot of this formal procedure. We already know the functional form of ϕ in $\hat{L}_z$ eigenstates and we know from the last lecture that the eigenvalues of $\hat{L}^2$ are $l(l+1)\hbar^2$. Given that we know the ϕ dependence, we can write

$$Y_{l\,m_l}(\theta,\phi) = T(\theta)e^{im_l\phi}$$

Hence, putting this into $\hat{L}^2$ eigenvalue equations gives

$$\hat{L}^2 Y_{l\,m_l} = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]T(\theta)e^{im_l\phi} = l(l+1)\hbar^2 T(\theta)e^{im_l\phi}$$

Operating with the ϕ derivative explicitly reduces this to

$$\frac{e^{im_l\phi}}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{dT}{d\theta}\right) - m_l^2\frac{e^{im_l\phi}}{\sin^2\theta}T(\theta) = -l(l+1)T(\theta)e^{im_l\phi}$$

Cancelling the ϕ part and multiplying by $\sin^2\theta$ then gives

$$\sin\theta\frac{d}{d\theta}\left(\sin\theta\frac{dT}{d\theta}\right) + l(l+1)\sin^2\theta T(\theta) = m_l^2 T(\theta)$$

This is now an equation for θ only so we have effectively separated the variables. However, even in this form, to solve this equation is not trivial. The usual method changes the variable to $\epsilon = \cos\theta$ so that

$$\frac{d}{d\theta} = \frac{d\epsilon}{d\theta}\frac{d}{d\epsilon} = -\sin\theta\frac{d}{d\epsilon}$$

The function of ϵ is different from that of θ so so we write $T(\theta)$ as $P(\epsilon)$. For $m_l = 0$ this equation then reduces to

$$-\sin^2\theta\frac{d}{d\epsilon}\left[-\sin^2\theta\frac{dP}{d\epsilon}\right] + l(l+1)\sin^2\theta P = 0$$

so cancelling the $\sin^2\theta$ gives

$$\frac{d}{d\epsilon}\left[(1-\epsilon^2)\frac{dP}{d\epsilon}\right] + l(l+1)P = 0$$

This is the *Legendre equation* and the solutions are the *Legendre polynomials* which you may have met in other lecture courses. In a similar way to how we solved the 1D SHO, we will assume a series solution

$$P = \sum_p a_p\epsilon^p$$

which gives us

$$\sum_p\left[(p+1)(p+2)a_{p+2} - p(p+1)a_p + l(l+1)a_p\right]\epsilon^p = 0$$

For this to be zero for all values of ϵ, then there must be a recurrence relation for the coefficients

$$a_{p+2} = \frac{p(p+1) - l(l+1)}{(p+1)(p+2)}a_p$$

As for the SHO case, this will only give a finite series if one of the coefficients goes to zero. It is clear this is the case for $p = l$ so the series only goes up to ϵ^l or equivalently $(\cos\theta)^l$. The solutions are therefore labelled by the quantum number $P_l^0(\epsilon)$.

So far we only did the $m_l = 0$ solution, hence the superscript zero labelling P_l^0 above. As the original equation only depends on m_l^2, then the sign of m_l is irrelevant and the solution only depends on $|m_l|$. The $m_l \neq 0$ solutions are called the *associated Legendre polynomials*, which are given by

$$P_l^{|m_l|} = (1-\epsilon^2)^{|m_l|/2} \frac{d^{|m_l|} P_l^0}{d\epsilon^{|m_l|}}$$

The $|m_l|$ derivatives here apply to the polynomial which goes up to ϵ^l. This means that if $|m_l| > l$, then all the terms in the polynomial will disappear and we get zero. Hence, for a non-zero solution, we must restrict m_l to

$$|m_l| < l \qquad \text{i.e.} \qquad -l \leq m_l \leq l$$

which is exactly what we found in the previous lecture.

3 Spherical harmonics

The complete angular solutions are then

$$Y_{l\,m_l}(\theta, \phi) = A P_l^{|m_l|}(\cos\theta) e^{i m_l \phi}$$

The constant A is needed as the standard definitions of the (associated) Legendre polynomials do not give normalised wavefunctions. The functions $Y_{l\,m_l}$ are called the *spherical harmonics.* This is because they correspond to the fundamental and higher frequency wave-like excitations, not on a straight line like for a guitar string, but on the surface of a sphere.

The spherical harmonics have various useful properties. Firstly, as expected for any eigenstates of Hermitian operators, they are orthonormal. This means

$$\int_0^\pi \int_0^{2\pi} Y_{l\,m_l}^* Y_{l'\,m_l'} \sin\theta \, d\theta \, d\phi = \delta_{ll'} \delta_{m_l m_l'}$$

They also display the spherical isotropy of space when the probabilities for all the degenerate states are summed

$$\sum_{m_l=-l}^{m_l=+l} |Y_{l\,m_l}|^2 = \frac{2l+1}{4\pi}$$

and so is not a function of θ or ϕ. To check this for e.g. $l = 1$, then

$$|Y_{1\,1}|^2 + |Y_{1\,0}|^2 + |Y_{1\,-1}|^2 = \frac{3}{8\pi}\sin^2\theta + \frac{3}{4\pi}\cos^2\theta + \frac{3}{8\pi}\sin^2\theta = \frac{3}{4\pi}\sin^2\theta + \frac{3}{4\pi}\cos^2\theta = \frac{3}{4\pi}$$

Finally, in spherical polars, then the complete wavefunction is $\psi(r, \theta, \phi)$. We can write this as

$$\psi = f(r) Y_{l\,m_l}(\theta, \phi)$$

for any function $f(r)$. Then, since $\hat{L}^2$ only depends on θ and ϕ,

$$\hat{L}^2 \psi = f(r) \hat{L}^2 Y_{l\,m_l}(\theta, \phi) = f(r) l(l+1)\hbar^2 Y_{l\,m_l}(\theta, \phi) = l(l+1)\hbar^2 \psi$$

and so the total wavefunction is still an eigenstate of $\hat{L}^2$. The same clearly holds for $\hat{L}_z$ also. Hence, multiplying $Y_{l\,m_l}$ by any radial function keeps it an eigenstate of the angular momentum operators.

4 $\hat{L}_x$ and $\hat{L}_y$ eigenstates

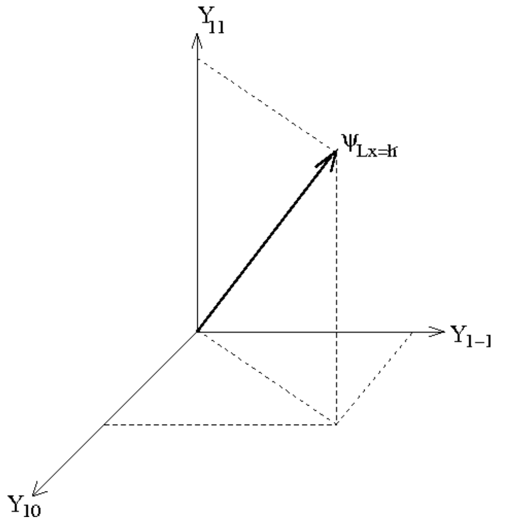
Clearly, it would be possible to derive similar results for $\hat{L}_x$ or $\hat{L}_y$ but it is a lot messier, again simply due to the definition of spherical polars. Also, as we showed, $\hat{L}_x$, $\hat{L}_y$ and $\hat{L}_z$ are not compatible so the eigenstates of e.g. $\hat{L}_x$ and $\hat{L}^2$ would *not* be spherical harmonics but would be combinations of these with the same l but different m_l. This is again messy so the choice is to always define z to be the axis along which we resolve the angular momentum vector and to have x and y as the components with uncertainty.

There is one exception, which is when $l = 0$ for which $L^2 = l(l+1)\hbar^2 = 0$. If the total angular momentum magnitude is zero, then clearly, all three components must be zero too. Hence, in this one case, we know L^2 and *all* three components at the same time. There is $2l + 1 = 1$ state for $l = 0$, which means this is the case with no degeneracy and hence it would not in fact be possible to mix several degenerate states for this case.

However, for $l > 0$, then it is only possible to make $\hat{L}_x$ or $\hat{L}_y$ eigenstates by mixing together the spherical harmonics. To give an example; for $l = 1$ then the component eigenvalues can be 0 or $\pm\hbar$. The z component eigenstate for $\hbar$ is clearly $Y_{1\,1}$. The state which gives the $\hat{L}_x$ eigenvalue $\hbar$ is a mixture of the three spherical harmonics with $l = 1$ and turns out to be

$$\frac{1}{2}\left(Y_{1\,1} + \sqrt{2}Y_{1\,0} + Y_{1\,-1}\right) = \sqrt{\frac{3}{8\pi}}\left(\cos\theta - i\sin\theta\sin\phi\right)$$

as can be shown by direct application of the $\hat{L}_x$ operator as given at the start of the lecture. This is clearly a mixture of the $\hat{L}_z$ eigenstates and so does not have a definite value for L_z, as expected. The three $l = 1$ spherical harmonics form a set of "axes", in terms of which any of the $\hat{L}_x$ or $\hat{L}_y$ states can be expressed.



The other equivalent states are given explicitly in the table below.

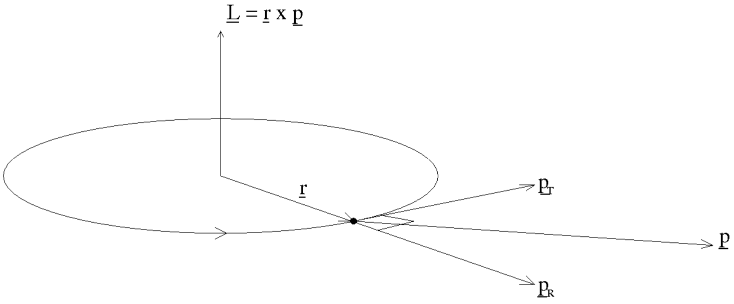
	x	y	z
$\hat{L}_i\phi = \hbar\phi$	$\frac{1}{2}\left(Y_{1\,1} + \sqrt{2}Y_{1\,0} + Y_{1\,-1}\right)$	$\frac{1}{2}\left(-iY_{1\,1} + \sqrt{2}Y_{1\,0} + iY_{1\,-1}\right)$	$Y_{1\,1}$
$\hat{L}_i\phi = 0\phi$	$\frac{1}{\sqrt{2}}\left(Y_{1\,1} - Y_{1\,-1}\right)$	$\frac{1}{\sqrt{2}}\left(Y_{1\,1} + Y_{1\,-1}\right)$	$Y_{1\,0}$
$\hat{L}_i\phi = -\hbar\phi$	$\frac{1}{2}\left(-Y_{1\,1} + \sqrt{2}Y_{1\,0} - Y_{1\,-1}\right)$	$\frac{1}{2}\left(iY_{1\,1} + \sqrt{2}Y_{1\,0} - iY_{1\,-1}\right)$	$Y_{1\,-1}$

Second Year Quantum Mechanics - Lecture 26
Central potentials

Paul Dauncey, 8 Dec 2011

1 Classical effective potential

The case of $V(\boldsymbol{r}) = V(r)$, i.e. V not being a function of θ and ϕ, is very common. For this case, consider a classical particle in an orbit due to a central potential.



We know its angular momentum is conserved. At any moment, the linear momentum $\boldsymbol{p}$ can be resolved along (p_R) and transverse to (p_T) the radial direction. Since the angular momentum

$$\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$$

then only the transverse component of $\boldsymbol{p}$ contributes to angular momentum. Explicitly, the magnitude of $\boldsymbol{L}$ is $|\boldsymbol{L}| = |\boldsymbol{r}||\boldsymbol{p}| \sin\alpha$ for an angle α between $\boldsymbol{r}$ and $\boldsymbol{p}$. Since $|\boldsymbol{p}_T| = |\boldsymbol{p}| \sin\alpha$, then the angular momentum magnitude is

$$|\boldsymbol{L}| = |\boldsymbol{r}||\boldsymbol{p}_T| \qquad \text{so} \qquad p_T^2 = \frac{L^2}{r^2}$$

The total energy of the particle is

$$E = \frac{p^2}{2m} + V(r) = \frac{p_R^2}{2m} + \frac{p_T^2}{2m} + V(r)$$

and so

$$E = \frac{p_R^2}{2m} + \frac{L^2}{2mr^2} + V(r)$$

Hence, when just considering the radial equation, i.e. solving for $r(t)$, then we need to consider an effective potential

$$V'(r) = V(r) + \frac{L^2}{2mr^2}$$

which is of course what we know gives the centrifugal force. The extra term is sometimes refered to as a potential barrier as it gets large for small r. Note that for non-zero angular momentum, then it becomes infinitely big for $r = 0$ so that the particle trajectory cannot go through the origin if $L > 0$. This is clearly also seen from considering the expression $L = rp_T$ above.

2 3D energy eigenstates

The TISE in 3D is

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r})\right] u(\boldsymbol{r}) = Eu(\boldsymbol{r})$$

In spherical polar coordinates, the Laplacian operator ∇^2 is quite complicated

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

However, the angular part of this may look familiar; it is exactly the same form as occurs in the angular momentum squared operator $\hat{L}^2$, where it is multiplied by $-\hbar^2$. Hence, we can write

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\left(-\frac{1}{\hbar^2}\hat{L}^2\right) = \nabla_R^2 - \frac{1}{r^2\hbar^2}\hat{L}^2$$

This means the total Hamiltonian is

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r}) = -\frac{\hbar^2}{2m}\nabla_R^2 + \frac{1}{2mr^2}\hat{L}^2 + V(\boldsymbol{r})$$

Hence, we get a term that looks identical to the potential barrier.

When is $\hat{L}^2$ conserved? Remembering that $\hat{L}^2$ only operates on θ and ϕ, then it will clearly commute with the radial part of the Hamiltonian and with the potential barrier term so

$$[\hat{H}, \hat{L}^2] = -\frac{\hbar^2}{2m}[\nabla_R^2, \hat{L}^2] + \frac{1}{2m}[\hat{L}^2/r^2, \hat{L}^2] + [V, \hat{L}^2] = [V, \hat{L}^2]$$

For a central potential, $V(\boldsymbol{r}) = V(r)$ and so is not a function of θ and ϕ so then

$$[\hat{H}, \hat{L}^2] = 0$$

i.e. in quantum mechanics, angular momentum is conserved for central potentials also. Note, the property that $\hat{L}^2$, and indeed $\hat{L}_z$, do not have any r dependence is crucial for this result.

3 The radial equation

We would normally try to solve the TISE by separation of variables, which is the same technique we used for the TDSE to separate it into a time part and the TISE. However, again, we can cut corners as we know the angular part is the same as the angular momentum which we already solved. We will look for solutions of the form

$$u(\boldsymbol{r}) = R(r)Y_{lm_l}(\theta, \phi)$$

where $Y_{lm_l}(\theta, \phi)$ are the eigenstates of $\hat{L}^2$ and $\hat{L}_z$. This means we have

$$\hat{L}^2 Y_{lm_l} = l(l+1)\hbar^2 Y_{lm_l}$$

Hence, we have

$$-\frac{\hbar^2}{2mr^2}Y_{lm_l}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \frac{1}{2mr^2}Rl(l+1)\hbar^2 Y_{lm_l} + V(r)RY_{lm_l} = ERY_{lm_l}$$

Dividing throughout by Y_{lm_l} then gives

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

which is an equation for the radial dependence of the energy eigenstate with all the angular dependence taken out. We can make this look very similar to the 1D TISE. Defining $\chi(r) = rR(r)$, then it is straightforward to show

$$\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) = r\frac{d^2\chi}{dr^2}$$

so the TISE can be written as

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

This looks just like a 1D TISE with an effective potential

$$V'(r) = V(r) + \frac{l(l+1)\hbar^2}{2mr^2}$$

which depends on l and hence corresponds to the classical form.

However, the critical thing is that all the (complicated) solution for the angular parts is *independent* of $V(r)$, under the assumption that the potential is central in the first place, of course. In general, we will need a third quantum number, usually called n, to identify the different solutions of the radial equation. Because this equation contains l through the potential barrier term, then the energy eigenstates and eigenvalues will in general depend on this value also. Hence, given a potential, we can then solve the above equation (at least in principle) for E_{nl} and $R_{nl}(r)$. Note, the l dependence is because we are changing the effective potential; given a value of l, then the solution only needs one quantum number, namely n, as expected.

Note, the energy, and radial solution, do *not* depend on m_l; why? A central potential is isotropic, i.e. independent of angle. For a given $\hat{L}^2$, then the different m_l values correspond to rotating the angular momentum vector around in space. The energy and radial distribution cannot depend on the direction of the angular momentum if the potential is central, so the fact that there is no m_l dependence is understood.

The total solution is

$$u_{nlm_l}(\boldsymbol{r}) = R_{nl}(r)Y_{lm_l}(\theta,\phi)$$

Hence, as expected, we find we need three quantum numbers to describe the complete solution in 3D.

4 Hydrogen atom

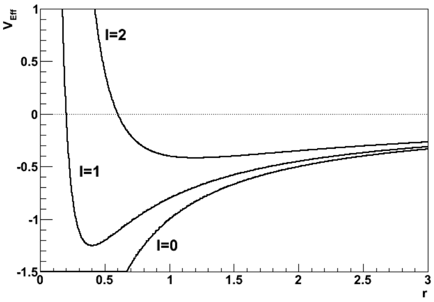
To proceed further, we need to pick a system and so have an explicit form for $V(r)$. For example, the hydrogen atom has

$$V(r) = -\frac{e^2}{4\pi\epsilon_0 r}$$

In fact, a very similar potential holds for many "hydrogen-like" systems, meaning any system where a single electron orbits around a charged nucleus. For a nucleus with Z protons, the potential is

$$V(r) = -\frac{Ze^2}{4\pi\epsilon_0 r}$$

and so a general Z factor is often included in the solutions. Including the potential barrier, then the effective potentials are as shown below.



Note $1/r^2$ will always beat $-1/r$ for small enough r, so for non-zero angular momentum, there is a large (effective) potential barrier for small r which goes infinite at the origin. Hence, all non-zero angular momentum wavefunctions will have their wavefunction go to zero at $r = 0$. This makes sense; no particle which can be found at the origin can have non-zero angular momentum.

The hydrogen-like radial functions are often expressed in terms of the *Bohr radius*

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{me^2} = 5.3 \times 10^{-11}\,\mathrm{m}$$

They generally have the form

$$R_{nl}(r) = f_{nl}(r)\exp\left(-\frac{Zr}{na_0}\right)$$

where n is an integer $n > l$ and is called the principal quantum number, l is the angular momentum quantum number and $f_{nl}(r)$ is a polynomial with terms up to r^{n-1}. Note, the solutions require $n > l$, which is usually more usefully considered as $l < n$.

The energy eigenvalues themselves are

$$E_n = -\frac{me^4}{2(4\pi\epsilon_0)^2\hbar^2n^2} = -\frac{e^2}{4\pi\epsilon_0 a_0}\frac{1}{2n^2} = -\frac{13.6}{n^2}\,\mathrm{eV}$$

It turns out that for the particular case of the Coulomb potential, the energy eigenvalues are independent of l and only depend on n; this is surprising given what we found before and is *not* a general result for central potentials. It is a so-called "accidental degeneracy". Because of this, there is generally a high degree of denegeracy for each energy state. Specifically, for $n = 1$ radial solutions, we can only combine these with $l = 0$ solutions so for this state only, there is no degeneracy. For $n = 2$, we can have $l = 0$ or $l = 1$, and so on. Remembering that $|m_l| < l$, then the first few states can have

n	l	$2l + 1$	Degeneracy
1	0	1	1
2	0	1	4
	1	3	
3	0	1	9
	1	3	
	2	5	

Hence, the totals are one state for $n = 1$, four states for $n = 2$, nine states for $n = 3$, etc, and generally there are n^2 degenerate states for energy E_n.

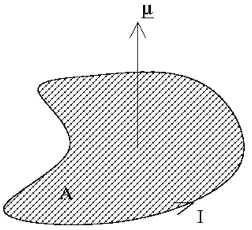
Second Year Quantum Mechanics - Lecture 27
Measuring angular momentum

Paul Dauncey, 9 Dec 2011

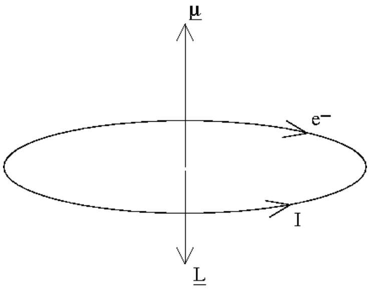
1 Classical magnetic dipoles

The obvious question is whether the angular momentum can be measured e.g. for an electron in an atom. It turns out that this is actually quite straightforward using a magnetic field.

A circuit with a current generates a magnetic field which for small circuits gives a magnetic dipole field. For a circuit with current I enclosing area A, then the magnitude of the magnetic dipole is $\mu = IA$.



We have seen that an electron in a classical circular orbit has angular momentum. However, it can also be considered as a circulating current, so its orbit generates a magnetic dipole. Clearly, the angular momentum and dipole moment arise from the same motion so they will be related.



The angular momentum $\boldsymbol{L}$ of the electron points out of the plane of the orbit with a magnitude

$$L = m_e \omega r^2$$

where ω is the angular frequency of the orbital rotation. The equivalent current (which is opposite to the motion of the electron, because of its negative charge) going round the loop of the circular orbit is

$$I = \frac{e}{\text{Period}} = e\nu = \frac{e\omega}{2\pi}$$

The magnetic dipole moment is the current times the area and has a direction opposite to $\boldsymbol{L}$ as the charge (and hence current) is negative. Its magnitude is

$$\mu = I\pi r^2 = \frac{e\omega}{2\pi}\pi\frac{L}{m_e\omega} = \frac{e}{2m_e}L$$

Vectorially

$$\boldsymbol{\mu} = -\frac{e}{2m_e}\boldsymbol{L}$$

This result in fact holds for any distributed system of mass and charge if the masses and charges have the same relative density distribution.

A magnetic dipole not only generates a magnetic field, but also interacts with an external magnetic field. There are two cases. If the field is *uniform*, i.e. $\boldsymbol{B}$ is the same everywhere in space, then the dipole feels a torque

$$\boldsymbol{\tau} = \boldsymbol{\mu} \times \boldsymbol{B}$$

This can be described in terms of a potential energy term given by

$$V_\mu = -\boldsymbol{\mu}.\boldsymbol{B}$$

i.e. the lowest energy is when the dipole is aligned with the magnetic field. In a *non-uniform* magnetic field, besides the above torque, there will also be a net force on the dipole, which can be found from the above potential

$$\boldsymbol{F} = -\boldsymbol{\nabla} V_\mu = \boldsymbol{\nabla}(\boldsymbol{\mu}.\boldsymbol{B})$$

Both these effects can be used experimentally to measure $\boldsymbol{\mu}$ and hence deduce $\boldsymbol{L}$.

2 The Zeeman effect

The effect of a uniform magnetic field is easily seen by measuring the spectra of atoms with and without magnetic fields. The energy shifts can be observed and hence $\boldsymbol{\mu}.\boldsymbol{B}$ (and so in principle $\boldsymbol{L}.\boldsymbol{B}$) can be determined.

To do this, we need to know the effect of the change to the potential on the atomic energy eigenvalues. This is where we can use perturbation theory; as long as the change to the energy is small (which is always possible with a small enough magnetic field) then we can approximate the energy shift by the first order perturbation result. The extra part of the Hamiltonian operator is

$$\hat{H}' = -\hat{\boldsymbol{\mu}}.\boldsymbol{B}$$

Defining the z axis to point along the magnetic field, then this is

$$\hat{H}' = \frac{e}{2m_e}\hat{\boldsymbol{L}}.\boldsymbol{B} = \frac{eB}{2m_e}\hat{L}_z$$

Since, as we have seen, angular momentum for atoms is of order $\hbar$, then the usual unit used in the *Bohr magneton* μ_B

$$\mu_B = \frac{e\hbar}{2m_e}$$

so that

$$\hat{H}' = \frac{\mu_B B}{\hbar}\hat{L}_z$$

The first order perturbation result is

$$\Delta E = \langle H' \rangle = \frac{\mu_B B}{\hbar}\langle L_z \rangle$$

If the particle is in an L_z eigenstate, then this is simple

$$\Delta E = \frac{\mu_B B}{\hbar} m_l \hbar = \mu_B B m_l$$

Hence, the different $\hat{L}_z$ eigenstates will split into different energy levels depending on their m_l values. The magnetic field breaks the spherical symmetry, and hence the degeneracy, so the energy now depends on the orientation and hence on m_l. We therefore expect to see each set of degenerate L^2 states split into the $2l+1$ separate L_z states, in energy steps of $\mu_B B$.

This effect for hydrogen (or other) atoms is called the Zeeman effect, after the person who pioneered this technique. The ground state of hydrogen has only $l=0$ so we expect no shift to the energy levels there. The first excited state has $l=0$ or 1 and so we would expect the $l=0$ and $l=1$, $m_l=0$ states to not be shifted, while the $l=1$, $m_l=\pm1$ states will be moved up or down in energy, respectively. However, this was *not* observed; the ground state actually split into two energies, moving $\pm\mu_B B$ while the first excited state split into four energies. This was known as the *anomalous Zeeman effect.*

3 The Stern-Gerlach experiment

Using a non-uniform magnetic field, then a force can be applied to an atom depending on the angular momentum orientation. For a non-uniformity (to a good approximation) only in the z direction, then the magnitude of the force is

$$F \approx \mu_z \frac{\partial B}{\partial z} = -\mu_B \frac{\partial B}{\partial z} m_l$$

and the force will be in the z direction also. This force will therefore deflect a beam of atoms depending on the value of m_l. Classically, for a fixed μ magnitude, we would expect a continuous range of μ_z values as the atoms would generally have random orientations to the z axis and hence a continuous range of deflections. However, the experiment first done by Stern and Gerlach showed only particular deflections were found; this is a direct manifestation of the quantisation of angular momentum. Again, we would expect hydrogen atoms not to be deflected in their ground state, as $l=0$, and the first excited state to be split, with $m_l=\pm1$ being deflected in different directions while the two $m_l=0$ states are not deflected at all. However, it was found that the ground state gave two lines, corresponding to the deflections expected for $m_l=\pm1$. The initial conclusion was that they were somehow seeing the $m_l=\pm1$ states and the $m_l=0$ states were "missing" in some undefined way.

4 Spin

Both the anomalous Zeeman effect and the Stern-Gerlach experiments were only properly understood when Uhlenbeck and Goudsmit proposed that there is some extra angular momentum in the atoms which does not arise from its overall motion. They called this *spin* and the concept is that the electron has some "intrinsic" angular momentum as if it was continously spinning on its axis. This is analogous to the total angular momentum of the Moon being due both to its orbit around the Earth and also to its spinning on its axis; these are separate sources of angular momentum. In the electron case, the magnitude of the angular momentum can never be changed; it is a basic property of the electron (and many other particles). Its direction can be changed by rotating the electron spin around in space, but the length is a constant of nature.

We have so far discussed the angular momentum due to the orbit and so this is called "orbital angular momentum". We now need to also consider this "spin angular momentum". A spinning electron might also form a magnetic dipole but as it is a point-like particle, it is not at all clear how the dipole is related to its spin angular momentum. For orbital angular momentum, we found

$$\boldsymbol{\mu}_L = -\frac{e}{2m_e}\boldsymbol{L}$$

Because the proportionality constant for spin is not obvious, then we write

$$\boldsymbol{\mu}_S = -\frac{ge}{2m_e}\boldsymbol{S}$$

where g is called the “gyromagnetic ratio”. It turns out that $g = 2$; this can be explained in relativistic quantum mechanics. This means we need to take

$$\boldsymbol{\mu}_S = -\frac{e}{m_e}\boldsymbol{S}$$

Hence it turns out that the spin angular momentum is twice as good as making a magnetic dipole moment as the orbital angular momentum. (In fact, in practise the factor g is not exactly 2 but has some relativistic corrections of around 0.1%, which we will ignore here.)

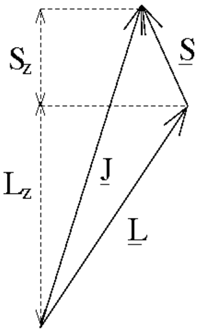
5 Addition of angular momentum

If there are both orbital and spin angular momentum, then adding these is complicated, but the easy case is when the electron is in the hydrogen ground state for which the orbital angular momentum has $l = 0$. Then, the only magnetic dipole is due to the spin. Both the anomalous Zeeman effect and the Stern-Gerlach results indicate that there are two possible orientations of the spin angular momentum. This is not possible for orbital angular momentum; the number of m_l values is always $2l + 1$, i.e. odd for integer l. To get only two values would mean that $2s + 1 = 2$ so $s = 1/2$, i.e. half-integer values. We found these are allowed by the ladder algebra but excluded for orbital angular momentum due to the angular boundary conditions. However, for a spinning point-like particle, no such angular boundary conditions are relevant, so we can use these extra solutions. The electron is therefore said to have “spin-half” with s (the equivalent of l) always being 1/2 and m_s (the equivalent of m_l) being $m_s = \pm 1/2$.

The addition of non-zero orbital and spin angular momenta will be covered in other courses but the basic result is that the total angular momentum

$$\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$$

satisfies the same operator relations as both $\boldsymbol{L}$ and $\boldsymbol{S}$ and so has the same structure for its eigenvalues. Note that for states with well-defined L_z and S_z, then the total z component $J_z = L_z + S_z$ is also definite.



Since the spin can either add to or subtract from the orbital angular momentum (roughly, whether $\boldsymbol{L}$ and $\boldsymbol{S}$ are parallel or anti-parallel), then the values of j (the equivalent of l) are $j = l + 1/2$ and $j = l - 1/2$. For any j, there is a $2j + 1$ degeneracy corresponding to the possible m_j values, just as for l and m_l. For example, for an electron in an $l = 1$ state, the possible total angular momentum eigenvalues are $j = 3/2$ or $j = 1/2$. The first has four possible m_j values, namely 3/2, 1/2, $-1/2$ and $-3/2$, while the second has two, namely 1/2 and $-1/2$.

Second Year Quantum Mechanics - Lecture 28
Spin eigenvalues and eigenstates

Paul Dauncey, 13 Dec 2011

1 Introduction

We will assume the electron (and many other fundamental particles) have an "intrinsic" angular momentum, which we will call spin, due to Uhlenbeck and Goudsmit. Because two states were found when splitting a beam in a magnetic field, then they postulated the value of s which allows two values of m_s is $s = 1/2$, for which $m_s = \pm 1/2$. This value of the total angular momentum is intrinsic to the electron and never takes any other value; it always has spin 1/2 and can never be "excited" to higher spin states.

This is not due to an angular wavefunction, i.e. a function of θ and ϕ, since we know the electron is a point particle from scattering experiments. As there is no spatial wavefunction, then we can drop the requirement of angular boundary conditions and hence allow the value of 1/2. However, we will need to go to a different way of writing the states, i.e. a different "representation". For this we need to use matrices.

2 Spin in QM

A reminder of the "normal" angular momentum (called "orbital" angular momentum to distinguish it from "spin" angular momentum). We used Postulate 3 to express $\hat{L}_i$ in terms of $\hat{r}_i$ and $\hat{p}_i$ for which we know the commutation relations. We then deduced the commutation relations for the $\hat{L}_i$ from these, e.g. $[\hat{L}_x, \hat{L}_y] = i\hbar \hat{L}_z$.

For spin, we cannot construct $\hat{S}_i$ from spatial and momentum operators as the spin is not due to any spatial coordinates. However, we will need to add it to orbital angular momentum and so we want it to act like orbital angular momentum in order for the sum to make sense. Hence we postulate the commutator rules for spin operators to be the same as for $\hat{L}_i$, e.g. $[\hat{S}_x, \hat{S}_y] = i\hbar \hat{S}_z$. We know we can deduce the eigenvalues directly using the commutators via the ladder operators, and so we know the spin must also satisfy

$$\hat{S}^2 \chi = s(s+1)\hbar^2 \chi, \qquad \hat{S}_z \chi = m_s \hbar \chi$$

The restriction to only a single s value, i.e. $s = 1/2$, and only two values of m_s, i.e. $m_s = \pm 1/2$, means the operators cannot be the same as the orbital angular momentum ones. In fact, we cannot make any operators like our previous operators with a restricted number of eigenstates. For all the previous cases ($\hat{r}_i$, $\hat{p}_i$, $\hat{E}$, $\hat{L}_i$), then there were an infinite number of eigenvalues and eigenstates. You can think of this as basically because the wavefunctions were functions of x, and x has an infinite number of possible values. To get a finite number of states, we need a "wavefunction-equivalent", χ, which depends on a variable with a finite number of values, i.e. $\chi(i)$ for a finite set of i. In fact, for only two possible m_s values, we need something which only takes two values, so we need something like $\chi(1)$ and $\chi(2)$. This effectively means the function has only two resulting values $\chi_1 = \chi(1)$ and $\chi_2 = \chi(2)$. In fact, rather than considering this as some function with a limited range on the input variable, it is easiest to write it as a two-component vector

$$\chi = \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix}$$

as we know how to handle vectors mathematically. We will also need "operator-equivalents" which take a vector and make another vector. We know of such a quantity mathematically already; this is simply a matrix, which here must be 2 × 2

$$\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix}$$

Hence, for spin, we have to use, not operators and functions, but matrices (for $\hat{S}$) and vectors (for χ). Hence, handling spin is often simpler than handling normal operators because there are not infinite numbers of eigenvalues and hence eigenstates. Allowing for these differences, then spin acts just like other angular momentum.

Note, the spin is really *in addition* to the other parts of the wavefunction and so the total wavefunction must contain both. This will be explained in the next lecture but in general, instead of $\psi(x, y, z)$, we need to use

$$\psi(x, y, z, w) = \begin{bmatrix} \psi_1(x, y, z) \\ \psi_2(x, y, z) \end{bmatrix}$$

where $w = 1$ or 2. Here for simplicity, we will ignore all the spatial dependence of the wavefunction and just look at the spin vector part.

3 Spin z component

As you should know from solving for eigenvalues with matrices, to get two eigenvalues, i.e. $m_s = \pm 1/2$ means we need to have a two-component vector for χ and so a 2×2 matrix for the $\hat{S}_i$. As usual we will look for states with well-defined $\hat{S}_z$ values and the simplest approach is with a matrix for $\hat{S}_z$ which is diagonal. Since the eigenvalues are $\pm\hbar/2$, then this matrix can be taken as

$$\hat{S}_z = \begin{pmatrix} \hbar/2 & 0 \\ 0 & -\hbar/2 \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Hence, the $\hat{S}_z$ eigenvalue equation is

$$\frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} = \lambda \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix}$$

It is easy to see that $\chi_1 = 1$, $\chi_2 = 0$ will give the eigenvalue $\lambda = \hbar/2$, while $\chi_1 = 0$, $\chi_2 = 1$ will give the eigenvalue $\lambda = -\hbar/2$. Hence, the eigenvectors of $\hat{S}_z$ are

$$\alpha_z = \chi_+ = \uparrow = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{spin up} \tag{1}$$

$$\beta_z = \chi_- = \downarrow = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \text{spin down} \tag{2}$$

where the different forms show the various different conventions used. Note, the term "eigenstate" is used to mean either an eigenfunction (of a differential operator) or an eigenvector (of a matrix operator) or a combination of these, as for the total wavefunction above. The Dirac notation $|\psi\rangle$ is used to stand for either type, which is another reason why it is widely used.

4 Eigenstates as a CONS

We need to check these states obey our usual rules, i.e. that the eigenstates form a CONS. Firstly the equivalent of normalisation; usually we do

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

i.e. it requires multiplying the eigenstate by its complex conjugate and integrating to give a single number. For the two-component eigenvector, this would become

$$\sum_{w=1}^{2} \chi^*(w)\chi(w) = \chi_1^*\chi_1 + \chi_2^*\chi_2 = \begin{pmatrix} \chi_1^* & \chi_2^* \end{pmatrix} \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} = \chi^\dagger \chi$$

where the † means take the complex conjugate and transpose the vector. For functions, the transpose is not explicitly specified but is implicit in the definition of function multiplication. The combined operation of complex conjugate and transpose is called the *Hermitian conjugate.* We can therefore check the normalisation for e.g. α_z using

$$\alpha_z^\dagger \alpha_z = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1$$

Similarly, β_z is also normalised.

For orthogonality, we previously had

$$\int \phi_n^*(x)\phi_m(x)\, dx = 0$$

for $n \neq m$. In terms of vectors, we now need to do e.g.

$$\alpha_z^\dagger \beta_z = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0$$

and so the states are indeed orthogonal.

For completeness, then we need to be able to express any state in terms of the eigenstates. Previously, this meant that for any function $\psi(x)$

$$\psi(x) = \sum_n a_n \phi_n(x)$$

In the case of vectors, for any vector

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = a_1\alpha_z + a_2\beta_z$$

and so it is trivial to see the two eigenstates form a complete set.

5 Other spin operators

We will expect to find raising and lowering operators, which must have the properties when operating on α_z of

$$\hat{S}_+ \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0, \qquad \hat{S}_- \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \hbar \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and similarly for β_z. (The $\hbar$ here is to keep the dimensions the same as for the orbital angular momentum raising and lowering operators.) These must also be 2×2 matrices and by inspection, they are simply

$$\hat{S}_+ = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \hat{S}_- = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

We previously found that $\hat{L}_\pm = \hat{L}_x \pm iL_y$ and in fact deduced the properties of the raising and lowering operators from these. Here, we can work backwards now we have $\hat{S}_+$ and $\hat{S}_-$ and use these to find $\hat{S}_x$ and $\hat{S}_y$. Using $\hat{S}_\pm = \hat{S}_x \pm i S_y$, these give

$$\hat{S}_x = \frac{1}{2}\left(\hat{S}_- + \hat{S}_+\right), \qquad \hat{S}_y = \frac{1}{2}\left(i\hat{S}_- - i\hat{S}_+\right)$$

which are therefore

$$\hat{S}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \hat{S}_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

We can define three dimensionless matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

which are the "Pauli spin matrices" and then write

$$\hat{S}_x = \frac{\hbar}{2}\sigma_x, \qquad \hat{S}_y = \frac{\hbar}{2}\sigma_y, \qquad \hat{S}_z = \frac{\hbar}{2}\sigma_z$$

These matrices are used for many applications, not just spin. Any two-level quantum system uses them and they are even useful in classical calculations. They have a number of important properties, one of which is that for any of the three

$$\sigma_i^2 = I$$

where I is the 2×2 unit matrix. They also have a property that for any of them, they are equal to their Hermitian conjugate $\sigma_i^\dagger = \sigma_i$, i.e. transposing them and taking a complex conjugate gives the same matrix back. This is the equivalent of requiring that the operators are Hermitian and hence ensures the $\hat{S}_i$ give real eigenvalues and complete orthonormal eigenstates.

6 Spin magnitude

We will also need the $\hat{S}^2$ operator. This is defined, as we would expect

$$\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2 = \frac{\hbar^2}{4}I + \frac{\hbar^2}{4}I + \frac{\hbar^2}{4}I = \frac{3\hbar^2}{4}I$$

Hence, the operator is a particularly simple matrix, being the unit matrix. Hence, *any* two-component vector η is an eigenvector of this matrix, with an eigenvalue given by

$$\frac{3\hbar^2}{4}I\eta = \frac{3\hbar^2}{4}\eta = \lambda\eta$$

i.e. $\lambda = 3\hbar^2/4$. If we write this as $s(s+1)\hbar^2$ then it is clear we need $s = 1/2$. Hence, the spin quantum number for any state is always $s = 1/2$ and never changes.

7 Eigenvalues for $\hat{S}_x$ and $\hat{S}_y$

By symmetry, we would expect all three axes to be equivalent so we had better find the same eigenvalues for $\hat{S}_x$ and $\hat{S}_y$ as we found for $\hat{S}_z$. Taking $\hat{S}_x$, then

$$\frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \lambda \begin{pmatrix} a \\ b \end{pmatrix}$$

This could be solved by the usual methods for finding eigenvalues, i.e. writing as

$$\begin{pmatrix} -\lambda & \hbar/2 \\ \hbar/2 & -\lambda \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = 0$$

and then requiring the determinant of this matrix to be zero so it cannot be inverted. However, we can use the properties of the Pauli spin matrices to find the eigenvalues easily. Multiplying by $\hat{S}_x$ again, then we get

$$\frac{\hbar^2}{4}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}a\\ b\end{pmatrix} = \lambda\frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}a\\ b\end{pmatrix} = \lambda^2\begin{pmatrix}a\\ b\end{pmatrix}$$

But as stated above, the square of any of the Pauli matrices is I so this is

$$\frac{\hbar^2}{4}\begin{pmatrix}a\\ b\end{pmatrix} = \lambda^2\begin{pmatrix}a\\ b\end{pmatrix}$$

for which is it clear that the eigenvalues are $\lambda = \pm\hbar/2$. This method works for any of the three spin components and so they all have the same eigenvalues. This is clearly required due to symmetry. To find the eigenstates then this this needs to be done explicitly so for $\hat{S}_x$, we have

$$\frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}a\\ b\end{pmatrix} = \pm\frac{\hbar}{2}\begin{pmatrix}a\\ b\end{pmatrix}$$

meaning $b = \pm a$. Hence, normalising

$$\alpha_x = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ 1\end{pmatrix}, \qquad \beta_x = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -1\end{pmatrix}$$

Similarly, for $\hat{S}_y$

$$\frac{\hbar}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\begin{pmatrix}a\\ b\end{pmatrix} = \pm\frac{\hbar}{2}\begin{pmatrix}a\\ b\end{pmatrix}$$

means $b = \pm ia$, so

$$\alpha_y = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\end{pmatrix}, \qquad \beta_y = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -i\end{pmatrix}$$

Second Year Quantum Mechanics - Lecture 29
Spin in magnetic fields

Paul Dauncey, 15 Dec 2011

1 Introduction

We have seen that we need to consider a new two-component vector "wavefunction" to accomodate spin into the formalism of quantum mechanics. This has 2 × 2 matrices acting as the operators which act on the vector. We now need to show how this fits into the structure we had previously.

2 The wavefunction with spin

Previously, the Hamiltonian was

$$\hat{H}(\hat{\boldsymbol{r}}, \hat{\boldsymbol{p}})$$

However, with spin then in general the Hamiltonian will be

$$\hat{H}(\hat{\boldsymbol{r}}, \hat{\boldsymbol{p}}, \hat{\boldsymbol{S}})$$

Given this, then the overall eigenstates, including both spatial and spin components, will generally be

$$\psi = \begin{bmatrix} \psi_1(\boldsymbol{r}, t) \\ \psi_2(\boldsymbol{r}, t) \end{bmatrix} = \psi_1(\boldsymbol{r}, t) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \psi_2(\boldsymbol{r}, t) \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

so in general we have two coupled differential equations to solve, one for ψ_1 and one for ψ_2.

However, in some systems, the spatial and spin parts of the Hamiltonian only occur in separate terms

$$\hat{H} = \hat{H}_r(\hat{\boldsymbol{r}}, \hat{\boldsymbol{p}}) + \hat{H}_s(\hat{\boldsymbol{S}})$$

If this is the case, then the overall wavefunction separates into two parts

$$\psi = \psi_r(\boldsymbol{r}, t)\chi_s(t)$$

where

$$\chi_s(t) = \begin{bmatrix} \chi_1(t) \\ \chi_2(t) \end{bmatrix}$$

Note $\chi_i \neq \chi_i(\boldsymbol{r})$. If this holds, the TDSE then becomes

$$\begin{aligned} i\hbar\frac{\partial\psi}{\partial t} &= \hat{H}\psi \\ i\hbar\frac{\partial}{\partial t}(\psi_r\chi_s) = i\hbar\chi_s\frac{\partial\psi_r}{\partial t} + i\hbar\psi_r\frac{d\chi_s}{dt} &= (\hat{H}_r + \hat{H}_s)\psi = \chi_s\hat{H}_r\psi_r + \psi_r\hat{H}_s\chi_s \end{aligned}$$

It is clear this is satisfied if the two separate parts of the wavefunction have their own separate TDSEs

$$\begin{aligned} i\hbar\frac{\partial\psi_r}{\partial t} &= \hat{H}_r\psi_r \\ i\hbar\frac{d\chi_s}{dt} &= \hat{H}_s\chi_s \end{aligned}$$

as then the pairs of terms in the above equation cancel. This is the equivalent of the separation of variables we have seen several times. The first of these is just the TDSE we have been solving previously in this course; we now see there is another part as well. In terms of energy eigenstates, then the above TDSEs give two TISEs

$$\hat{H}_r\psi_r = E_r\psi_r, \qquad \hat{H}_s\chi_s = E_s\chi_s$$

so

$$\hat{H}\psi = \chi_s\hat{H}_r\psi_r + \psi_r\hat{H}_s\chi_s = E_r\chi_s\psi_r + E_s\psi_r\chi_s = (E_r + E_s)\psi_r\chi_s = (E_r + E_s)\psi$$

Hence, the total energy is $E = E_r + E_s$ as would be expected.

To clarify the form of $\hat{H}_s$, then the time dependence of the spin part of the wavefunction can be written more explicitly as

$$i\hbar\frac{d}{dt}\begin{pmatrix}\chi_1\\ \chi_2\end{pmatrix} = \hat{H}_s\begin{pmatrix}\chi_1\\ \chi_2\end{pmatrix}$$

so it is clear that $\hat{H}_s$ has to be a 2×2 matrix.

3 Spin in a uniform magnetic field

We saw in a previous lecture that the potential energy of a dipole in a magnetic field is

$$V_\mu = -\boldsymbol{\mu}.\boldsymbol{B}$$

and, remembering that the spin makes a dipole twice the naive value, then an electron has a magnetic dipole due to its spin which is given by

$$\boldsymbol{\mu}_S = -\frac{e}{m_e}\hat{\boldsymbol{S}}$$

If the magnetic field is uniform, then $\boldsymbol{B} \neq \boldsymbol{B}(\boldsymbol{r})$ and so the potential due to the dipole is separate from any other energy terms. This means this is a case where the Hamiltonian separates into spin and spatial parts, where the Hamiltonian spin term is

$$\hat{H}_s = -\boldsymbol{\mu}_S.\boldsymbol{B} = \frac{e}{m_e}\boldsymbol{S}.\boldsymbol{B}$$

Defining the z axis to be along the magnetic field direction, then this is

$$\hat{H}_s = \frac{eB}{m_e}\hat{S}_z = \frac{e\hbar B}{2m_e}\sigma_z = \mu_B B\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

in terms of the Bohr magneton μ_B again. Hence, it is clear than the spin-energy eigenstates will also be $\hat{S}_z$ eigenstates, which we found in the last lecture. Hence, there will be two energy eigenstates

$$u_1 = \begin{pmatrix}1\\ 0\end{pmatrix}, \qquad u_2 = \begin{pmatrix}0\\ 1\end{pmatrix}$$

The energies of these eigenstates are easily found. For the first

$$\hat{H}_s\begin{pmatrix}1\\ 0\end{pmatrix} = \mu_B B\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix} = \mu_B B\begin{pmatrix}1\\ 0\end{pmatrix}$$

which therefore has an eigenvalue of $E_1 = \mu_B B$, while for the second

$$\hat{H}_s\begin{pmatrix}0\\ 1\end{pmatrix} = \mu_B B\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}0\\ 1\end{pmatrix} = \mu_B B\begin{pmatrix}0\\ -1\end{pmatrix} = -\mu_B B\begin{pmatrix}0\\ 1\end{pmatrix}$$

which therefore has an eigenvalue of $E_2 = -\mu_B B$. Note, the higher energy state is the one with the spin aligned with the magnetic field, as this has the dipole moment anti-aligned.

The coupling of the spin to the magnetic field is used in MRI where transitions between the two energy levels are detected. Using the nuclear spin (rather than the electron spin) gives lower frequencies as the energy difference goes as $1/m$. These are usually in the radio range which can be easily detected.

4 Larmor precession

Larmor precession concerns the motion of the spin vector in a magnetic field. We solved above for the eigenstates, which are stationary states and so have no time dependence for observables. However, when the system is not in an energy eigenstate then there is in general some visible motion. For this case, we can also tackle this problem directly as we can explicitly solve the TDSE. Writing

$$i\hbar\frac{d}{dt}\begin{pmatrix}\chi_1\\ \chi_2\end{pmatrix} = \hat{H}_s\begin{pmatrix}\chi_1\\ \chi_2\end{pmatrix} = \mu_B B\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}\chi_1\\ \chi_2\end{pmatrix} = \mu_B B\begin{pmatrix}\chi_1\\ -\chi_2\end{pmatrix}$$

which means

$$i\hbar\frac{d\chi_1}{dt} = \mu_B B\chi_1 = E_1\chi_1, \qquad i\hbar\frac{d\chi_2}{dt} = -\mu_B B\chi_2 = E_2\chi_2$$

These can be solved to give

$$\chi_1 = a_1 e^{-iE_1 t/\hbar} = a_1 e^{-i\mu_B Bt/\hbar}, \qquad \chi_2 = a_2 e^{-iE_2 t/\hbar} = a_2 e^{i\mu_B Bt/\hbar}$$

where a_1 and a_2 are constants of integration and physically correspond to the values of χ_1 and χ_2 at $t = 0$, i.e. the initial conditions. The spin vector is then given by

$$\chi_s = \begin{pmatrix}\chi_1\\ \chi_2\end{pmatrix} = \begin{pmatrix}a_1 e^{-i\mu_B Bt/\hbar}\\ a_2 e^{i\mu_B Bt/\hbar}\end{pmatrix} = a_1\begin{pmatrix}1\\ 0\end{pmatrix}e^{-i\mu_B Bt/\hbar} + a_2\begin{pmatrix}0\\ 1\end{pmatrix}e^{i\mu_B Bt/\hbar} = a_1 u_1 e^{-iE_1 t/\hbar} + a_2 u_2 e^{-iE_2 t/\hbar}$$

Hence, it has the same form as we have seen previously; i.e. the general solution is an arbitrary sum of the energy eigenstates, with their separate time dependences

$$\begin{aligned}\psi(x, t = 0) &= \sum_n a_n u_n\\ \psi(x, t) &= \sum_n a_n u_n e^{-iE_n t/\hbar}\end{aligned}$$

We can now find the expectation values of the three spin components. This is most easily done if a_1 and a_2 are assumed to be real. The expectation value would usually be found using

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

As shown in the last lecture, the equivalent of the integral for the spin vectors involves the Hermitian conjugate and so the spin expectation values are given by

$$\langle S_i\rangle = \chi_s^\dagger \hat{S}_i \chi_s = \frac{\hbar}{2}\chi_s^\dagger \sigma_i \chi_s$$

Putting in the three matrices gives

$$\begin{aligned}\langle S_x\rangle &= \hbar a_1\, a_2 \cos(2\mu_B Bt/\hbar)\\ \langle S_y\rangle &= -\hbar a_1\, a_2 \sin(2\mu_B Bt/\hbar)\\ \langle S_z\rangle &= \frac{\hbar}{2}(a_1^2 - a_2^2)\end{aligned}$$

Hence, the expectation value of S_z is constant with time, as would be expected since the energy eigenstates are also $\hat{S}_z$ eigenstates. However, the S_x and S_y expectation values *precess* around the magnetic field direction with an angular frequency $2\mu_B B/\hbar = eB/m_e$. This is analogous to a top precessing in a gravitational field and is called "Larmor precession". This is an example of Ehrenfest's theorem where the expectation values follow the classical equations.

Note, if $a_1 = 0$ or $a_2 = 0$, then the expectation values $\langle S_x\rangle$ and $\langle S_y\rangle$ are zero so there is again no motion. In this case, since the other a_i must be 1 to give the correct probability, then $\langle S_z\rangle = \hbar/2$ or $-\hbar/2$. This corresponds to the energy eigenstate case, as with one of the $a_i = 0$ and the other 1, then the state is simply an energy (and hence $\hat{S}_z$) eigenstate.