

Chem. 113A:

Dr. Daniel Neuhauser

Notes based on handouts of Prof. Shimon Weiss

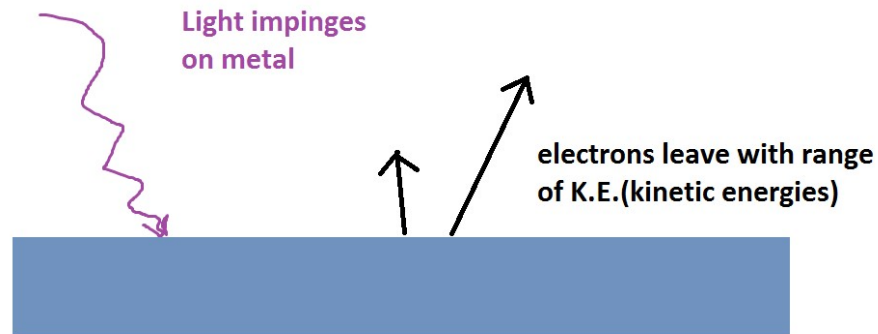
Chapter I. Classical Mechanics (CM) failures:

I.1 Classical atom not stable



Orbiting charge radiates, atom falls inside.

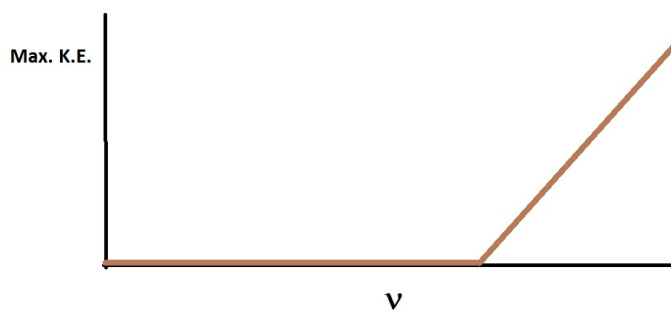
I.2 Photoelectric Effect



Classical mechanics predicts that max of kinetic energy determined only by I , the intensity of the light, and is independent of the frequency, ν .

But experiment says differently.

Experimentally:



i.e., no electrons beyond a threshold (which depends on the metal), and for frequencies above the threshold,

$$\text{Max. } K.E. = h\nu - W$$

where W is the work-function of the metal, i.e., how much energy we need to give an electron to be kicked out a metal.

Most importantly h , the slope of this curve, is universal, i.e., it is the same no matter which metal we use.

h is measured to be

$$h = 6.62 * 10^{-34} \text{ J} * \text{sec}$$

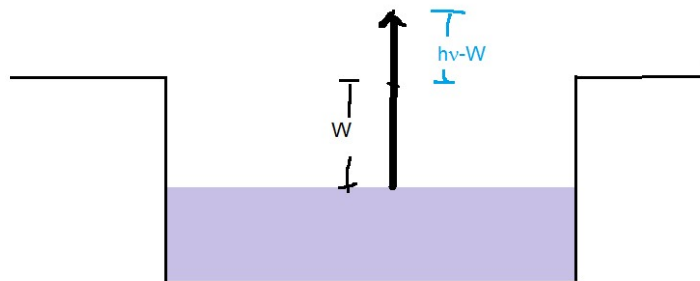
Einstein explained the photoelectric effect: light is made of photons. Each photon has energy

$$E_{\text{photon}} = h\nu$$

(Einstein based this on the work of Planck on Blackbody radiation, which we'll review later, i.e., we proceed non-historically).

If the photon energy is beyond the work function of the metal, than when a photon is absorbed it gives the electron enough energy to climb out of the metal, and the electron still be left with some energy.

Think of the electron as being stuck in a “well” – the photon needs to give it enough energy to climb out of the well, and if it does, the remaining energy (not spent on climbing out of the well) is available for the electron as kinetic energy.



Example: for sodium the work function is 2.28 eV, where eV, the energy that a single electron gets upon a change of voltage of 1V, equals

$$1\text{eV} = 1.60 * 10^{-19}\text{J}$$

So if we shine blue light with frequency of, say, $\nu = 6.70 * 10^{14}\text{Hz}$ on sodium, the photon energy is

$$E_{\text{photon}} = h\nu = 6.70 * 10^{14}\text{Hz} = 6.62 * 10^{-34} \text{ J} * \text{s} = 4.44 * 10^{-19}\text{J} = 2.77\text{eV}$$

(since $\text{Hz} = \frac{1}{\text{s}}$).

I.e., when this blue photon is absorbed, it will have enough energy to kick out an electron, and give it as much as

$$\text{K. E. max} = h\nu - W = 2.77\text{eV} - 2.28\text{eV} = 0.49\text{eV}$$

I.3 Another problem of CM: Blackbody radiation predicted to be infinitely strong...

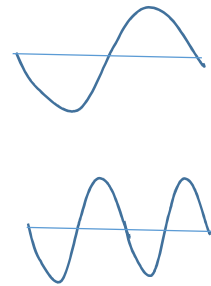
A blackbody: absorbs essentially all impinging radiation and emits it later.

Modes inside a black body

Inside a blackbody – many modes of radiation which contribute.

The modes are discretized – key property of standing waves (e.g., guitar harmonics). Incidentally, this discussion of modes will be very relevant for QM later on, e.g., for particle in a box.,

1D: For E.M. the allowed modes look like:



etc., where the end points are the end of the box, which has size L (i.e., the horizontal line's length is denoted by L)

The “wavelength” of this standing waves are, by inspection

$$\lambda_1 = L$$

$$\lambda_2 = \frac{L}{2}$$

And in general

$$\lambda_j = \frac{L}{j} \quad j = 1, 2, 3, \dots$$

Note that unlike guitar modes, there's no mode with $\lambda = \frac{L}{2}$.

We get the frequencies from recalling that frequency * wavelength = speed = c (here speed of light)

$$\nu\lambda = c$$

so

$$\nu_j = \frac{c}{\lambda_j} = \frac{c}{L} j$$

So the frequency is proportional to the index telling us what is the frequency

So in 1D, the density of modes, (i.e., the number of modes with frequency between $\nu, \nu + d\nu$ divided by the frequency range $d\nu$) denoted as $\rho_{1D}(\nu)$, will be constant

$$\rho_{1d}(\nu) = \frac{L}{c}$$

I.e., the same number of modes will be, for a given black body, between 100MHZ and 105MHZ and 100 GHZ and 100.005 GHZ (since the range of frequencies is the same in both cases, 5MHZ.). And of course, the bigger the black body, the more modes will be in it.

In 3D this “density of modes” will get an additional factor of ν^2 , it is an “area factor” – you saw it in 20B when you learned that the Maxwell speed distribution is (where $u = speed$)

$$f_{3d,Maxwell}(u) = constant * \mathbf{u^2} * \exp\left(-\frac{Mu^2}{2k_B T}\right)$$

The bolded u^2 in Maxwell’s distribution is due to the 3D (it will be absent if space was 1D), and similarly here for radiation $\rho_{3D}(\nu)$ is proportional to the squared frequency ν^2 ; the proportionality constant can be shown to be $8\pi L^3/c^3$, so

$$\rho_{3D}(\nu) = \frac{L^3}{c^3} * 8\pi\nu^2$$

Digression: classical heat capacities.

If you did not learn it, just accept that classically every degree of freedom has a heat capacity, which is essentially the derivative of the energy with

respect to temperature, of $\frac{R}{2}$. Classically this explains why the heat capacity of a Helium gas is $3R/2$, as in Helium, or Neon, or other noble gases, the only degrees of freedom are the overall translational motion of the atoms

You learnt that each degree of freedom has a heat capacity of $\frac{R}{2}$, but that was per mole; per one degree of freedom we will need to divide by the number of particles in a mole, i.e., the per-one-mode heat capacity classically is be $\frac{1}{2} \frac{R}{N_{Avogadro}} = \frac{k_B}{2}$, where Boltzmann's constant, is

$$k_B = \frac{R}{N_{Avogadro}} = \frac{8.31 \frac{\text{J}}{\text{K mol}}}{6.02 * 10^{23} \frac{\text{mol}}{\text{mol}}} = 1.38 * 10^{-23} \frac{\text{J}}{\text{K}}$$

End of digression.

Each radiation mode – same as Harmonic Oscillator

In electrodynamics, one can show that each radiation mode we talked about above is associated with an oscillating electric and magnetic field (i.e., “electromagnetic” field). Further, the intensity of that electromagnetic field (for each mode) behaves as if it was a single harmonic oscillator.

You learned in 20B and/or 110A (and especially 110B if you took it) that an harmonic oscillator has two degrees of freedom associated with it, the position and momentum. For radiation, the degrees of freedom are different, they are the strength of the electric field and the strength of the magnetic field, but mathematically they are analogous.

Here we have two degrees of freedom (electric and magnetic fields), so the heat capacity associated with a single radiation mode is k_B .

Therefore classically we will expect that each of the modes will have an average energy per mode of $k_B T$, where $k_B = \frac{R}{N_{Avogadro}}$ is Boltzmann's constant.

CM: problem with energy.

Now you see that problem. Since the density of modes $\rho_{3d}(\nu)$ grows with ν , and since the energy of each mode is classically expected to be the same ($k_B T$), then if we sum up all the energies of each mode we will get infinity (since the sum of all the energies will be the sum over the number of modes over all frequencies, which is infinite.)

QM: solution: discrete photons.

Planck offered a mathematical solution (later understood to be true physically)

$$E(\text{photon with frequency } \nu) = h\nu$$

So the photon energies are discrete. Therefore, defining:

$$P_0 = \text{Prob}(0 \text{ photons with } \nu)$$

we get from statistical mechanics that

$$\begin{aligned} \text{Prob}(1 \text{ photon with } \nu) &= \exp\left(-\frac{E(\text{one photon})}{k_B T}\right) P_0 \\ &= \exp\left(-\frac{h\nu}{k_B T}\right) P_0 \end{aligned}$$

(if you are not familiar with this form, recall that you have seen it in 20B as $\exp\left(-\frac{E}{RT}\right)$ where then E and R referred to per-mole quantities; here R is replaced by k_B since we talk about per mode or per particle quantities, not per-mole)

Similarly

$$\text{Prob}(2 \text{ photon with } \nu) = \exp\left(-\frac{2h\nu}{k_B T}\right) P_0$$

$$\text{Prob}(3 \text{ photon with } \nu) = \exp\left(-\frac{3h\nu}{k_B T}\right) P_0$$

And generally

$$\text{Prob}(n \text{ photon with } \nu) = \exp\left(-n \frac{h\nu}{k_B T}\right) P_0$$

Planck's photons solved the blackbody problem

Physically

Before showing the full expression mathematically, let's understand physically why the introduction of photons solved the problem.

Now if our photon frequency is large, $\nu \gg \frac{k_B T}{h}$, then the photon energy will be much larger than the thermal energy, $h\nu \gg k_B T$, so the Boltzmann factor for even having a single photon (at an energy ν) will be exponentially small

$$h\nu \gg k_B T \rightarrow \exp\left(-\frac{h\nu}{k_B T}\right) \sim 0.$$

Therefore, a black body can only emit radiation at frequencies up to a few times $k_B T$.

Mathematically:

$$\begin{aligned} \langle E_{tot}(\text{mode with } \nu) \rangle &= \text{Average } E \text{ in mode with frequency } \nu = \\ &= \frac{((\text{prob. } 0 \text{ photons}) * E(0 \text{ photons } \nu) + (\text{prob. } 1 \text{ photons}) * E(1 \text{ photons } \nu) + \dots)}{(\text{prob. } 0 \text{ photons}) + \text{prob. } 1 \text{ photon} + \text{prob. } 2 \text{ photons} + \dots)} \\ \langle E_{tot}(\text{mode with } \nu) \rangle &= \frac{\left(0 * P_0 + h\nu * P_0 * \exp\left(-\frac{h\nu}{kT}\right) + 2h\nu * P_0 * \exp\left(-2\frac{h\nu}{kT}\right) + \dots\right)}{P_0 + P_0 * \exp\left(-\frac{h\nu}{kT}\right) + P_0 * \exp\left(-2\frac{h\nu}{kT}\right) + \dots} \end{aligned}$$

$$\langle E_{tot}(\text{mode with } \nu) \rangle = h\nu \frac{\left(0 + 1 * \exp\left(-\frac{h\nu}{kT}\right) + 2 * \exp\left(-2\frac{h\nu}{kT}\right) + \dots\right)}{1 + \exp\left(-\frac{h\nu}{kT}\right) + \exp\left(-2\frac{h\nu}{kT}\right) + \dots}$$

and defining $s \equiv \frac{h\nu}{k_B T}$ we get:

$$\langle E_{tot}(\text{mode with } \nu) \rangle = h\nu \frac{0 + 1 * \exp(-s) + 2 * \exp(-2s) + \dots}{1 + \exp(-s) + \exp(-2s) + \dots}$$

And it can be shown mathematically that the ratio becomes

$$\langle E_{tot}(\text{mode with } \nu) \rangle = h\nu \frac{\exp(-s)}{1 - \exp(-s)} = \frac{h\nu \exp\left(-\frac{h\nu}{k_B T}\right)}{1 - \exp\left(-\frac{h\nu}{k_B T}\right)} = \frac{h\nu}{\exp\left(\frac{h\nu}{k_B T}\right) - 1}$$

Thus, we see mathematically that for high frequencies due to the $\exp\left(-\frac{h\nu}{k_B T}\right)$ factor in the numerator this expression will vanish, i.e., the energy that a black body has associated with high frequencies is damped since there are exponentially few occupied modes at high frequencies.

Final QM expression:

The final expression for the energy density that a blackbody has at each frequency ν is

(density of modes with ν) \times (average energy in each mode)

i.e.,

$$\text{Energy density}(\nu) = \rho_{3D}(\nu) * \langle E_{tot}(\text{mode with } \nu) \rangle$$

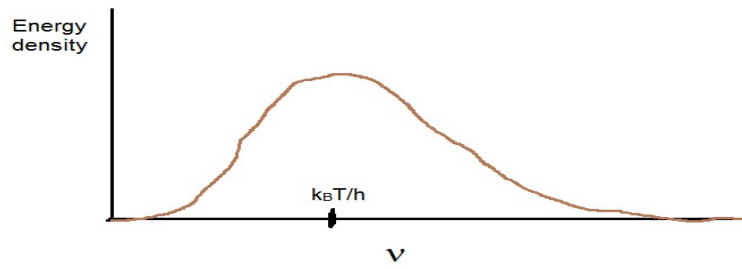
i.e.,

$$\text{Energy density}(\nu) = \frac{L^3}{c^3} 8\pi\nu^2 \times h\nu \frac{1}{\exp\left(\frac{h\nu}{k_B T}\right) - 1}$$

i.e.,

$$\text{Energy density}(\nu) = \frac{L^3}{c^3} 8\pi h \frac{\nu^3}{\exp\left(\frac{h\nu}{k_B T}\right) - 1}$$

Indeed, when we plot this expression we get exactly the extended bell-shape curve that is seen experimentally, and indeed it falls off significantly for $\nu > \frac{4k_B T}{h}$



At room temperature $\frac{4k_B T}{h}$ is at the IR so the earth radiates at the IR.

I.4 Another problem of CM: Heat Capacity of atomic solids

As we mentioned you learned (and you'll prove if you take 110B) that classically each degree of freedom gives $R/2$ to the heat. For example, C_v of a He gas is $3R/2$, as each atom has three degrees of freedom – moving in the x, y, or z direction.

So say we have cubic solids, for simplicity (Na Cl, e.g.)

We expect:

$$C_p \sim 3R = 6 \frac{R}{2}$$

This is because

- In a cubic solid each atom has around 6 bonds with its neighbors;
- Each bond is “half” belonging to each atom, i.e., there are 3 net bonds per atom.
- Each bond is like a harmonic oscillator (it is spring-like), i.e., has a kinetic energy and potential energy, i.e., 2 degrees of freedom.
- So therefore, per atom there are $3 \cdot 2 = 6$ degrees of freedom

And in reality (experimentally):

- At high T indeed $C_p \sim 3R$.
- But at low T, C_p decreases and vanishes (turns out that $C_p \sim T^2$ at low T).



Why isn't C_p constant?

Reason (Q.M.): At low T, the thermal energy is not sufficient to excite the bonds. Well to be more precise, one can excite at low T only the large scale motion of the atoms called “phonons”,

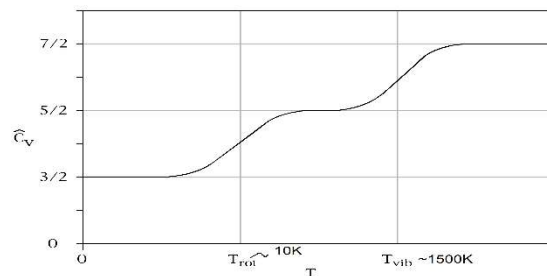
but at T gets lower the number of such phonons with frequencies ν that are not much larger than $\frac{k_B T}{h}$ gets smaller, so it gets difficult to excite such vibrations, i.e., the energy in vibrations diminishes at low T so the heat capacity decreases.

Interestingly, there is something you probably saw in 20A or 20B or 110A which is analogous: The fixed-volume heat capacity of light diatomic gases (e.g., H_2, O_2, N_2) is $5R/2$, and is due to the overall translational motion of the center of mass (3 degrees of freedom) and due to the rotation (2 degrees of freedom); the vibration is not active.

But at high T (say 1500K) the vibration of a diatom becomes active, i.e., as we input more heat one can populate significantly the higher energy vibrational states. I.e., at high temperature the probability to occupy the first excited or higher vibrational state, $\exp\left(-\frac{h\nu_{vib}}{k_B T}\right)$ is not zero anymore.

So at high T, $C_V \sim 7R/2$ for diatoms (7 since: 3 translations, 2 rotations, and 2 vibrations).

By the way, an analogous but opposite phenomena occurs at low temperatures – around 10K or lower; then the frequency associated with rotational excitations (see later, a few weeks from now) becomes larger than $\frac{k_B T}{h}$, so the diatom cannot rotate, and its heat capacity is only $3R/2$ due to the 3 translational motions



I. 5 Discreteness of spectra

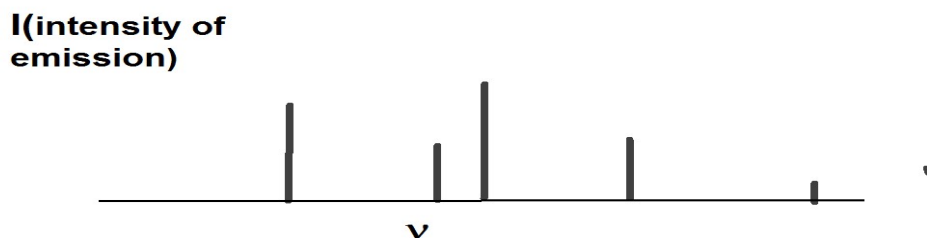
Another unexplained CM phenomena was the discreteness of spectra of atoms (and molecules too, though that's more complicated to see in real-life due to effects of rotation and vibrations – things are simpler for atoms).

Atoms can be excited (at that time, late 1800's)

- In space (emitted or absorbed light observed then in telescopes).
- In flames
- By an electrical discharge.

Then, when the atoms are excited, they emit light.

When one measures then the spectra of the emitted light, then lines are observed



C.M. : no explanation for discreteness of spectra.

Q.M.: couple with $E = h\nu$. I.e., absorption or emission of a photon is associated with a change of energy.

So If we assume (a strong assumption, soon explained) that electrons in atoms are in specific states or level (“orbitals” for us chemists) then emission just means that the electrons move from one state to another.

I.e., emission is associated with change of an electron energy, so the total energy is conserved: denoting the initial state (orbital) of the electron by 2,

and the final by 1, then $E_2 - E_1 = h\nu$

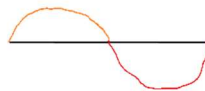


But where's discretization in nature ?

For waves – i.e., guitar modes (a single fundamental:



and harmonics (e.g., the first harmonics)



are discretized. So indeed electrons are waves, but before discussing that, let's discuss the first approach at quantifying electron motion – i.e., an approach which modifies classical mechanics, and gets fit to experiment for the simplest types of atoms: the Bohr model.

Chapter II. Towards QM

II. 1 The Bohr Model

Experimentally, turns out that the emission frequencies of **excited Hydrogen atoms** belong to a discrete series, which can be labeled by two integers, n, m where

$$n = 1, 2, 3, \dots$$

$$m < n$$

Such that

$$\nu_{nm} = \nu_0 \left(\frac{1}{m^2} - \frac{1}{n^2} \right)$$

(so, e.g., $\nu_{21} = \nu_0 * \left(1 - \frac{1}{2^2} \right) = \frac{3}{4} \nu_0$, etc.) where

$$\nu_0 = 3.29 * 10^{15} \text{ Hz}$$

(i.e., about 4-5 times the frequency of purple and blue light, which is $\sim 7 * 10^{14} \text{ Hz}$).

Bohr explained it as mentioned in the previous page, i.e., the emitted light is associated with transition from level n to level m :

$$\nu_{nm} = \frac{E_n - E_m}{h}$$

where the energy of each level is

$$E_n = -\frac{h\nu_0}{n^2}$$

i.e.,

$$E_1 = -h\nu_0 = 6.62 * 10^{-34} \text{ Jsec} * 3.29 * 10^{15} \text{ Hz} = 2.18 * 10^{-18} \text{ J} = 13.6 \text{ eV}$$

$$E_2 = -\frac{h\nu_0}{4} = -\frac{13.6 \text{ eV}}{4} = -3.40 \text{ eV}$$

etc.

Further, Bohr built a “theory” which explains it, the Bohr model, which says:

Bohr model assumptions:

- Electrons are in classical orbitals
- The only deviation from usual CM is that these classical orbitals have a quantized (i.e., discrete) angular momentum values:

$$|\mathbf{L}| = n\hbar = 0, \hbar, 2\hbar, 3\hbar, \dots$$

where

$$\hbar = \frac{h}{2\pi} = 1.05 \times 10^{-34} \text{Jsec}$$

and recall that the angular momentum is

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}$$

So: CM says (recall your physics classes) that \mathbf{L} is conserved; Bohr further says that its absolute values are quantized.

Implications of Bohr model for 1-elec systems (H, He⁺ Li⁺⁺, etc.).

The attraction to the nucleus for an electron that's distance r away is

$$F_{\text{attractive}} = b \frac{Ze^2}{r^2}$$

Where Z is the nuclear charge

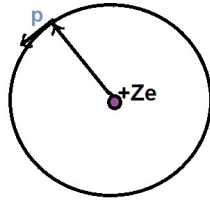
$$Z = 1 \text{ (for H)}; \quad Z = 2 \text{ (for He}^+); \quad Z = 3 \text{ (for Li}^{++})$$

and

$$b \equiv \frac{1}{4\pi\epsilon_0}$$

is part of the MKS units convention.

Let's look only at spherical orbits (though the model works also for elliptical ones).



Then r and p are perpendicular, so the cross product value is simply their product:

$$|L| = |r \times p| = rp$$

Couple with the Bohr's model assumption

$$|L| = n\hbar$$

to get

$$p = \frac{n\hbar}{r}$$

So the velocity, which is the momentum over the mass, becomes

$$v = \frac{p}{M} = \frac{n\hbar}{Mr}$$

Now just equate the centrifugal and attractive force on the electron (the centrifugal force pushes it away, so to speak, and the attractive Coulomb force pulls it to the nucleus):

$$F_{centrifugal} = F_{Coulomb}$$

i.e.,

$$\frac{Mv^2}{r} = b \frac{Ze^2}{r^2}$$

But we just derived $v = \frac{n\hbar}{Mr}$ so plug it in to get

$$\frac{M}{r} \left(\frac{n\hbar}{Mr} \right)^2 = b \frac{Ze^2}{r^2}$$

i.e.,

$$\frac{n^2 \hbar^2}{Mr^3} = b \frac{Ze^2}{r^2}$$

i.e.,

$$r = \frac{n^2}{Z} \left(\frac{\hbar^2}{bMe^2} \right)$$

i.e.,

$$r = \frac{n^2}{Z} a_0$$

where we introduced

$$a_0 = \frac{\hbar^2}{bMe^2} = 4\pi\epsilon_0 \frac{\hbar^2}{Me^2} = 0.529 \text{ Angstrom} = 0.529 * 10^{-1} \text{ m}$$

a_0 is called “bohr radius” and is a length unit.

Note:

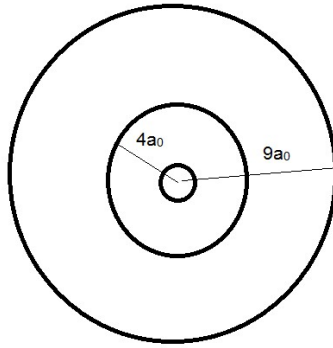
- The distance to the nucleus becomes smaller as the charge increases. This is indeed true for the deepest core states, which mostly feel only the attraction of the nucleus, and are “tighter” (i.e., more compact) as the charge increases.
- The distance from the nucleus increases quickly

$$r(n = 1) = \frac{a_0}{Z}$$

$$r(n = 2) = 4 \frac{a_0}{Z}$$

$$r(n = 3) = 9 \frac{a_0}{Z}$$

This is true for excited states of hydrogen (or of He^+ , etc.), they become very extended in gases; this is known in QM as “Rydberg states”



Finally, the most important part – the Bohr model prediction for energies

First, consider in general systems with Columbic interactions (regardless of the Bohr model). Then

$$E = E(\text{Kinetic}) + E(\text{attraction to nucleus})$$

Where

$$E(\text{kinetic}) = \frac{1}{2}Mv^2$$

$$E(\text{attraction}) = -\frac{bZe^2}{r}$$

From the equation on the force equality

$$\frac{Mv^2}{r^2} = b \frac{Ze^2}{r^2}$$

We get (by multiplying by r)

$$\frac{Mv^2}{r} = b \frac{Ze^2}{r}$$

i.e.

$$2E(\text{kinetic}) = -E(\text{attraction})$$

Or

$$E = E(\text{Kinetic}) + E(\text{attraction}) = E(\text{kinetic}) - 2E(\text{kinetic})$$

i.e.,

$$E = -E(\text{kinetic}) = \frac{1}{2}E(\text{attraction})$$

So far this statement, labeled the “Virial Theorem”, is general for any system with electric interactions (and is true, btw, also in QM).

Now plug in the Bohr model prediction $r = \frac{n^2 a_0}{Z}$ to get:

$$E = \frac{1}{2}E(\text{attraction}) = -\frac{1}{2} \frac{bZe^2}{r} = -\frac{1}{2} \frac{bZe^2}{\frac{n^2 a_0}{Z}} = -\frac{\frac{1}{2} b e^2}{a_0} \frac{Z^2}{n^2}$$

i.e.,

$$E = -R_H \frac{Z^2}{n^2}$$

where we introduced the Rydberg constant

$$R_H = \frac{1}{2} \frac{b e^2}{a_0} = \frac{1}{8\pi\epsilon_0} \frac{e^2}{a_0}$$

You can (and should) calculate that:

$$R_H = 2.18 * 10^{-18} J = 13.6 \text{ eV}$$

i.e., we get perfect agreement with experiment, since that’s exactly what experiment found.

Summary of Bohr model:

- Fits experiment amazingly for hydrogen (and hydrogen like atoms).
- The underlying assumption, $|L| = 0, \hbar, 2\hbar, \dots$, quantization of the angular momenta, seems very elegant, and is indeed related to quantum mechanical rules.
- Does not work for more complicated atoms with many electrons or for molecules – we need the better theory (QM, derived a few years later).

II.2 Wave nature of electrons

1880's:

Waves were thought to be:

- Extended
- Can interfere. E.g., as you have seen in 20A and organic chemistry, when discussing destructive combination of A.O.(atomic orbitals) to make an M.O. (molecular orbital); we'll talk about interference in classical electricity below.
- Can be quantized – e.g., fundamental and harmonics as discussed earlier

And Particles were thought to be very different

- Point like
- Cannot interfere.

By 1900-1920 however:

- Problems with E.M. waves (black body, photoelectric effect, heat capacity of solids) → waves need to have particle property, $E = h\nu$
- But problems with discreteness of energies of electron energy remained; Bohr model tells us the electron energy levels are discrete but why?

So de Broglie (1923)

Particles → are waves!

Reason: de Broglie explained: For relativistic photons:

$$p = \frac{E}{c} \quad (\text{for photons, or any particle with } v = c)$$

(Note the factor of 2 difference from nonrelativistic mechanics, where

$$\text{Non – relativistic particles: } p = Mv, E = \frac{1}{2}Mv^2 = \frac{1}{2}pv \rightarrow p = \frac{2E}{v})$$

The usual wave relation, coming from $c = \text{velocity} = \text{frequency} * \text{wavelength}$, gives

$$c = v\lambda$$

i.e.,

$$v = \frac{c}{\lambda}$$

Combine then with Planck's relation for photons to give

$$E = hv = \frac{hc}{\lambda}$$

i.e.,

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

Or conversely

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

De Broglie noticed that this relation does not involve the velocity of light. So he speculated that

- This relation could be general, not just for light i.e., any particle is also a wave with a wavelength $\lambda = \frac{h}{p}$
- Note that this will give (as we'll see soon) naturally the discreteness of the orbitals, i.e., if particles are waves then they can be in discrete states, like the guitar harmonics we mentioned, and this will solve the problem of why an electron does not fall into the nucleus – since it is an extended wave, not a particle!
- But what's the equation this wave fulfils? This was answered by 1925/1927 by Schrödinger and separately by Heisenberg, as we'll discuss.

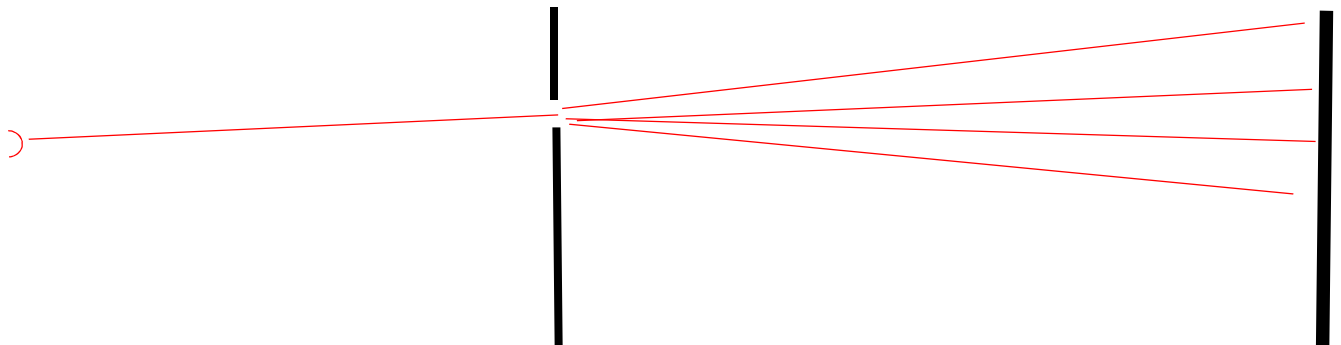
But before, we need to talk about : double slit, and interference.

II. 3 Interference and Double hole experiment: For light and for electrons.

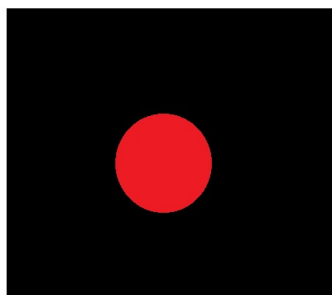
Consider a classical electromagnetic wave. For example, light that impinges on a screen. Say the screen (think of it as the blackboard) defines the y - z plane and the distance to the screen is the x -direction.

Say one of you lights a candle, and between the candle and the screen (the blackboard) we put a blocking wall with a single hole. Then on the screen we'll see a blob, due to the light that passes the hole.

Looking along an x - z cut, we'll see:

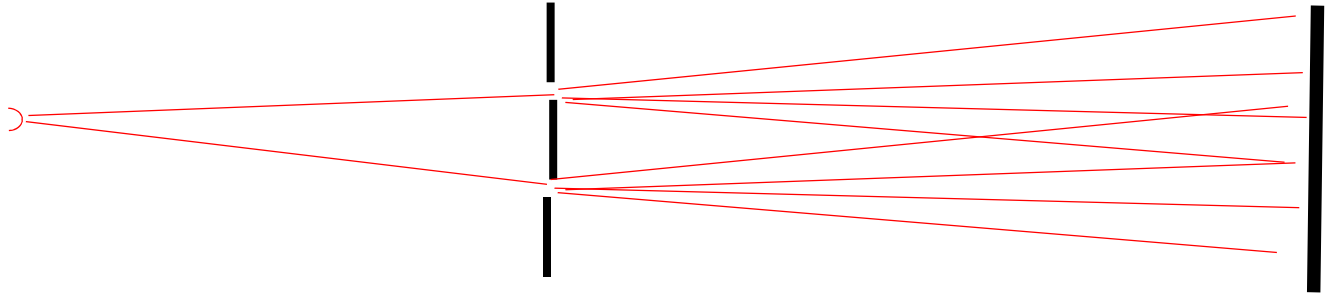


And the intensity of the light that's distributed on the right blackboard, will look like a blob; i.e., drawing the (y - z) screen, we'll see something like

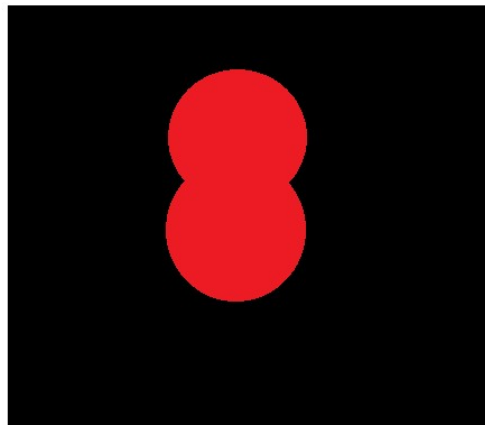


(obviously this is an overly simplified picture, we would have expected that the light beam will gradually fade away from the center of the blob, but I am not drawing too fancily here).

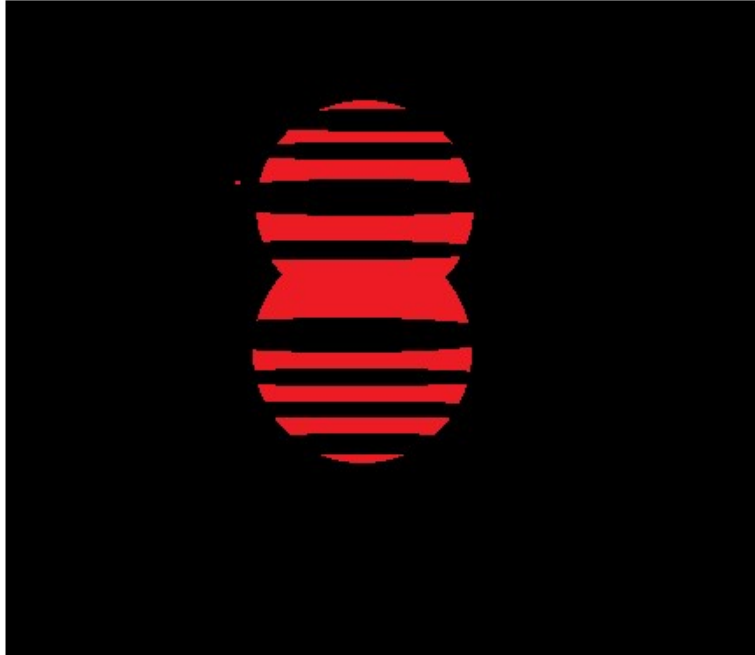
Next, open another hole in the intermediate blocking wall, i.e.,



IF light was made of purely particles, without any wave properties, we would have expected to see on the screen then an overlapping pattern, i.e., something like



In 1801 Thomas Young did this experiment using light (with slits, not holes, but same principle) and instead of finding this pattern above he found something which we call today interference pattern:



This proves that light is NOT made simply of particles (of course, the photoelectric effect proves that light is not a simple wave either, but that's for later in this section).

Why are the patterns? The intensity of the light at each point in the screen is proportional to the squared absolute electric field at that point.

$$I(\mathbf{r}) = |E(\mathbf{r})|^2$$

where \mathbf{r} is the 3-d position of the point on the screen.

(Truly the electric field has also direction, called "polarization", but we ignore it in this discussion).

Complex numbers:

To understand the interference, it turns out it is much easier to use complex electric field. Don't worry about the fact that electric fields have to be real, as they are measurable, for light this is a mathematical trick which we could have avoided at the price of a lot more math (for electrons the complex nature of the wavefunctions is essential, but that's later).

What we quickly need to know about complex numbers is that, first, a complex number, C , can be written as

$$C = a + ib$$

where a and b are real, and formally the complex "i" fulfils $i^2 = -1$.

We can define the conjugate of a complex number

$$C^* = a - ib$$

We can also define the absolute squared of a complex number, which is a real number:

$$|C| \equiv \sqrt{a^2 + b^2}$$

and fulfills

$$CC^* = |C|^2$$

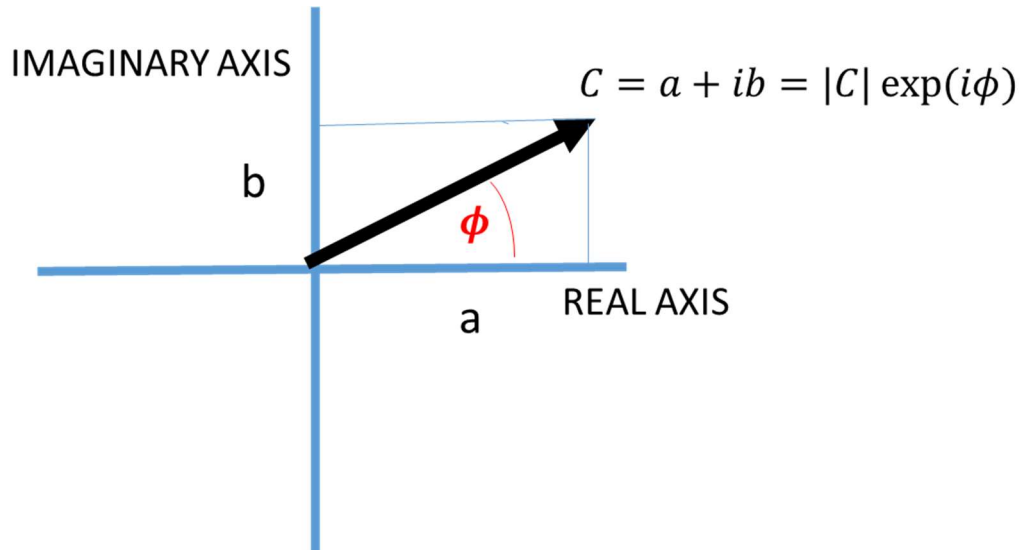
Proof:

$$\begin{aligned} CC^* &= (a + ib)(a - ib) \\ CC^* &= a^2 + iba - iab + ib(-ib) \\ CC^* &= a^2 - i^2b^2 \\ CC^* &= a^2 + b^2, \quad Q.E.D. \end{aligned}$$

Finally, we note that any complex number can be written as its absolute value times a phase; this is called a "polar representation"

$$C = |C| \exp(i\phi)$$

where ϕ is the angle between a, b , see below:



We can also write

$$C^* = (|C| \exp(i\phi))^* = |C| \exp(-i\phi),$$

i.e., the phase of the conjugate number is minus the phase of the original number, and the absolute value is the same. In particular, if the absolute value is one, we just get

$$(\exp(i\phi))^* = \exp(-i\phi).$$

This is consistent with another property of complex numbers,

$$\exp(i\phi) = \cos(\phi) + i \sin(\phi).$$

Multiplying complex numbers is very easy if we use this “polar” representation

$$C_1 C_2 = |C_1| \exp(i\phi_1) |C_2| \exp(i\phi_2) = |C_1| |C_2| \exp(i(\phi_1 + \phi_2))$$

Transmission through one hole mathematically:

So with complex numbers, we can say that if there is only one hole (e.g., we cover the other one), at a point \mathbf{r}_1 , then the electric field at any point \mathbf{r} on the screen is

$$E(\mathbf{r}) = |E_1(\mathbf{r})| \exp(i \phi_1(\mathbf{r}))$$

where

- $|E_1(\mathbf{r})|$ is a (real) blob-function, which looks like the red blobs we drew (in the y-z directions), i.e., it dies away from the center of the blob,
- the phase $\phi_1(\mathbf{r})$ which can be shown to be related to the distance of the hole from the screen ($|r - r_1|$) and to the wavelength λ :

$$\phi_1 \simeq \frac{2\pi}{\lambda} (|r - r_1|)$$

So with a single hole the intensity of the light will be

$$I(\mathbf{r}) = |E(\mathbf{r})|^2$$

i.e., for a single hole we will just see a blob, no surprise here.

Two holes: mathematically

More interestingly is what happens when we use now **two holes**. Now the electric field at a point \mathbf{r} will be **the sum of the electric fields coming from each hole**

$$E(\mathbf{r}) = E_1(\mathbf{r}) + E_2(\mathbf{r})$$

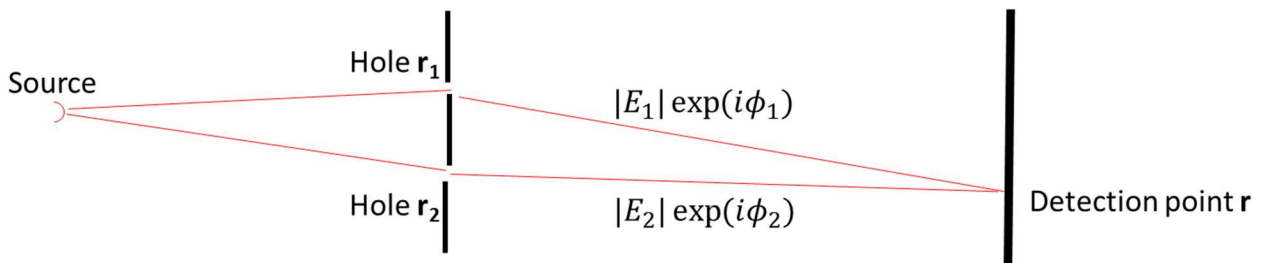
where, as before,

$$E_1 = |E_1(\mathbf{r})| \exp(i\phi_1(\mathbf{r}))$$

and analogously

$$E_2 = |E_2(\mathbf{r})| \exp(i\phi_2(\mathbf{r}))$$

See the figure, where I now pick a specific point and show the electric fields contributing from both holes”



so

$$|E(\mathbf{r})|^2 = |E_1 + E_2|^2$$

i.e.,

$$|E(\mathbf{r})|^2 = E(\mathbf{r})E^*(\mathbf{r}) = (E_1 + E_2)(E_1^* + E_2^*)$$

$$|E(\mathbf{r})|^2 = E_1E_1^* + E_1E_2^* + E_2E_1^* + E_2E_2^*$$

The first and last terms are just the intensities due to each hole separately, i.e., $E_1E_1^* = |E_1|^2$, etc., so

$$|E(\mathbf{r})|^2 = |E_1|^2 + E_1E_2^* + E_2E_1^* + |E_2|^2$$

The two terms in the middle are the interference terms;

$$E_1E_2^* + E_2E_1^* = |E_1| \exp(i\phi_1) |E_2| \exp(-i\phi_2) + |E_2| \exp(i\phi_2) |E_1| \exp(-i\phi_1)$$

$$E_1E_2^* + E_2E_1^* = |E_1||E_2|(\exp(i\phi_1) \exp(-i\phi_2) + \exp(i\phi_2) \exp(-i\phi_1))$$

so

$$E_1E_2^* + E_2E_1^* = |E_1||E_2|(\exp(i(\phi_1 - \phi_2)) + \exp(-i(\phi_1 - \phi_2)))$$

and using

$$\exp(i\phi) = \cos(\phi) + i \sin(\phi)$$

we get

$$\exp(i\phi) + \exp(-i\phi) = 2\cos(\phi)$$

so that the interference terms become

$$E_1E_2^* + E_2E_1^* = 2|E_1||E_2| \cos(\phi_1 - \phi_2)$$

and therefore the total intensity is

$$|E(\mathbf{r})|^2 = |E_1|^2 + |E_2|^2 + 2|E_1||E_2| \cos(\phi_1 - \phi_2)$$

Therefore when $\cos(\phi_1 - \phi_2) = -1$ there will be destructive interference, so the intensity will be $|E_1|^2 + |E_2|^2 - 2|E_1||E_2| = (|E_1| - |E_2|)^2$!

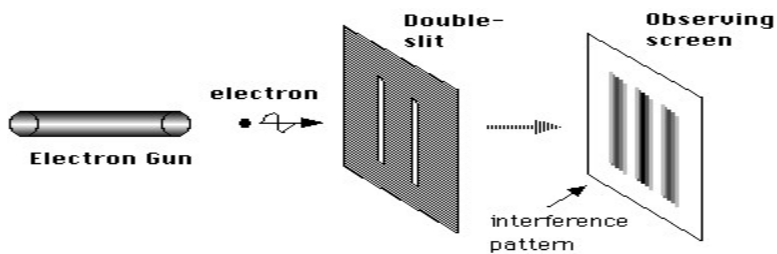
And if $|E_1| = |E_2|$ at the point, then we'll have completely destructive interference, $I = 0$! i.e., no light at that point (where the phases are opposite of each other, and the intensities are the same) when both holes are open. These points (or regions more generally) of destructive interference are denoted by black stripes in the interference pattern above.

BTW, this is completely analogous to destructive interference in Σ^* or Π^* bonds.

Reconciling the wave and photon picture of light

How to reconcile? For that we need to do the slit experiment in the limit of a small number of photons.

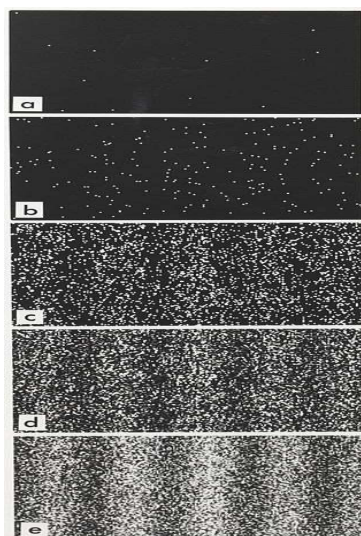
Below there's a true experimental example, actually for electrons (from Wikipedia), for the full slit experiment – but the same thing will happen for photons, although the physical length scale of the experiment will be different of course.



Below are the patterns, as function of time. Initially we see “particles” – electrons here, or photons if we were to do an the Yong slit experiment.

Later, as more as more “hits” are recorded, we start seeing the patterns.

So the continuous wave pattern are really a disguise for a pattern developed by many “hits”



This is like a branch of Impressionism, called “Pointillism”– where a pattern is made by many dots! See the picture by Seurat below.



Relation to electrons

Experiments on beams of electrons have shown, as we just saw, that electrons are indeed also a wave. A wave of what? Well, the wave is a “wavefunction” of an electron. Seems circular...

But the important thing is that we understand that the probability of an electron to be in a certain place, is analogous to the intensity of a light in a certain place $|E(\mathbf{r})|^2$;

So the probability of electron in be at a point \mathbf{r} is similarly $|\psi(\mathbf{r})|^2$ where ψ is a complex function of position, known as the wavefunction.

What we learn from the slit experiment, is that if an event (e.g., an electron or a photon hitting the screen at a point \mathbf{r}) can occur by one of several possibilities (e.g., the possibility that the light passes through slit 1 or through slit 2; or that the electron passes through slit 1 or through slit 2), then:

The total electric field(for light) or **total wavefunction, i.e., wave-amplitude** (for electrons or other “particles”) is

$$E = E_1 + E_2 \quad (\text{light})$$
$$\psi = \psi_1 + \psi_2 \quad (\text{electrons})$$

So in the latter case, the probability is

$$P = |\psi_1 + \psi_2|^2$$

And there's interference between the different pathways.

Note that if we were to measure where does the particle (photon or electron) goes through, e.g., if we were to cover the 2nd hole and measure how much goes through the first, the interference will be lost and we would have got no pattern, i.e., $P = |\psi_1|^2$.

Summary: double slits

Four Lessons for Quantum Chemistry from the Double-Slit (or Double hole) experiments:

- Particle wave duality: particles (e.g., electrons) also have wave properties.
- Superposition principle: an electron is in a superposition of two exclusive states (passing slit #1 & slit #2).
- Effect of measurement/observation: observation collapses the wavefunction to one of the states (passing slit #1 or slit #2), with equal probability in this specific case.
- Interpretation of wavefunction: probability of finding the particle is the squared of the absolute value of the wavefunction.

III. Waves, Operators and the Schrödinger equation

From Young's experiment and its equivalent on electrons: Central to our formulation of QM is the **probability amplitude** (or "**wavefunction**").

From the double-slit experiment, we have seen that the wavefunction gives rise to interference, just as wave amplitudes add to give interference.

As such, we seek a mathematical description of particle wavefunctions.

This ultimately leads to the **Schrödinger wave equation**, the solution of which produces the wavefunctions for any given quantum mechanical system.

Before we get to the Schrödinger wave equation, though, it is pertinent to review a bit what we know about waves and the mathematics thereof.

Wave Properties

- There are transverse monochromatic waves (TEM = Transverse Electric-Magnetic): disturbances whose **displacement oscillates in a direction perpendicular to k** , the direction of motion of the wave (water waves, EM waves).
- There are also longitudinal waves: a disturbance whose displacement oscillates along the direction of propagation of (sound).
- Amplitude of a wave = maximum displacement.
- The wavelength λ is the distance between two successive amplitude maxima at a given instant (units: μm , nm, Å).
- The velocity v = the distance a given wave crest moves per unit time [units: m/sec]. (**Actually, it is more complicated than that...there's something called group velocity and something called phase velocity...but we won't discuss those here**).
- The frequency ν : number of times per second a wave amplitude at a given point in space passes through the maximum value [unit: Hz].

See:

http://www.youtube.com/watch?v=CswoSQC_NX0&feature=endscreen&NR=1

<http://www.youtube.com/watch?v=UHCse1jJAto&feature=endscreen&NR=1>

<http://www.youtube.com/watch?v=aguCWnbRETU&feature=related>

Wave Properties

The radial frequency $\omega = 2\pi\nu$ [radians/sec].

- The period T = time required to complete one cycle of the oscillation [sec]
- The phase = number of radians accumulated since some arbitrary zero time [radians]: denoted by ϕ .
- **Principle of superposition: displacements add up** (for same position r and time t).

Henceforth: 1D (later we'll generalize to 3D).

The simplest **wave equation** for a (classical) wave $f(x, t)$ is

$$\frac{\partial^2 f}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 f}{\partial t^2} = 0$$

What is $f(x, t)$? Could be the density (for sound waves), the electric field (for light propagating in a homogenous media), etc. For this discussion assume v is constant.

The solutions of the wave equation could be written in terms of **propagating Normal modes**, $f_{k\omega}$ i.e., solutions that depend on specific wave vector (k) and frequency (ω) – (henceforth when I write frequency I mean angular frequency, $\omega = 2\pi\nu$)

$$f_{k\omega}(x, t) = \exp\left(i2\pi\left(\frac{x}{\lambda} - \nu t\right)\right)$$

And writing that $\omega \equiv 2\pi\nu$, $k \equiv \frac{2\pi}{\lambda}$, we get:

$$f_{k\omega}(x, t) = \exp(i(kx - \omega t))$$

Note that I could at this stage just look at the real part, i.e., $\cos(i(kx - \omega t))$ but we need to start practicing the complex solutions.

First, verify that these normal modes are indeed solving the wave equation

$$\frac{\partial^2 \exp(i(kx - \omega t))}{\partial x^2} = -k^2 \exp(i(kx - \omega t))$$

$$\frac{\partial^2 \exp(i(kx - \omega t))}{\partial t^2} = -\omega^2 \exp(i(kx - \omega t))$$

So

$$\frac{\partial^2 f_{k\omega}(x, t)}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 f_{k\omega}(x, t)}{\partial t^2} = \left(-k^2 + \frac{\omega^2}{v^2} \right) f_{k\omega}(x, t)$$

So in order to have a solution, we need to have

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

You actually know this relation—you learned that $\omega = 2\pi\nu$, that $k = \frac{2\pi}{\lambda}$ and that $v = \nu\lambda$, which is exactly equal!

Adding up solutions.

The key is the principle of superposition, which will lead us eventually to the Schrödinger equation!. Specifically:

Stationary Solutions-Warmup

Let's warm up by adding two counter propagating wave. One with $k > 0$, and the other with $-k$, i.e., define a wave

$$g(x, t) \equiv \exp(i(kx - \omega t)) + \exp(i(-kx - \omega t))$$

i.e.,

$$g(x, t) = (\exp(ikx) + \exp(-ikx)) \exp(-i\omega t)$$

i.e.,

$$g(x, t) = 2 \cos(kx) \exp(-i\omega t)$$

So we got that g is a stationary solution times an oscillating time-dependent factor!

$$g(x, t) = g_{stationary}(x) * \exp(-i\omega t), \quad g_{stationary}(x) = 2\cos(kx)$$

This is exactly what happens in, e.g., guitar waves where the stationary harmonics go up and down. Of course, in reality the solution of course is the real part of g , i.e.,

$$g_{true}(x, t) = g_{stationary}(x) * \cos(\omega t).$$

Adding up waves, and operators

The following will be **painful**, but once we understand it we can see the reason behind the specific form of Schrödinger's equation. (The usual alternative, just giving it outright without justification, is not educational in my opinion).

First, consider what happens when a wave is made of several (from 2 to infinitely many) components, with different ω, k .

For example,

$$a(x, t) \equiv 5 \exp(i(k_1x - \omega_1t)) + 8 \exp(i(k_2x - \omega_2t))$$

Where the "5" and "8" are arbitrary coefficients, and each component fulfils the wave relation, so that $\omega_1 = vk_1, \omega_2 = vk_2$

Well, if we were to pick some values for k_1 and k_2 (and therefore to ω_1 and ω_2) and plot the real and imaginary part of the wave, we will generally get a strange function which is non periodic. But it could represent an acceptable wave.

(That's what happens, essentially, when we play a guitar or any other instrument – we excite the fundamentals and harmonics of the different cords simultaneously, and get a wave of sound which is non periodic, i.e., we get music!)

Now, comes the question – how could we talk about "the" frequency of such waves, or more modestly, what will we mean by multiplying the wave by its "frequency".

Well, there is no single frequency to this wave $a(x, t)$, since it is made from 2 components with different frequencies. But we could still define a frequency **operator**.

Recall from linear algebra that **operators are things which convert functions to functions**.

We will denote the frequency operator by $\hat{\omega}$ — where the hat reminds us that it is an operator, not a number. This frequency operator, when acting on our wave, should produce another function, labeled here $b(x, t)$

$$b = \hat{\omega}a$$

What should $b(x, t)$ look like? Well, almost intuitively, the action of the operator of frequency on our wave, which I remind you is:

$$a(x, t) \equiv 5 \exp(i(k_1x - \omega_1t)) + 8 \exp(i(k_2x - \omega_2t))$$

should be

$$b(x, t) \equiv 5\omega_1 \exp(i(k_1x - \omega_1t)) + 8\omega_2 \exp(i(k_2x - \omega_2t))$$

(where I highlighted the frequency factors), i.e., each component multiplied by its own frequency. Makes sense.

Well, amazingly, we can find an elegant representation of this operator – it is just:

$$\hat{\omega} = i \frac{\partial}{\partial t}$$

You may not like it – how come we converted the concept of “frequency” into “a complex time-dependent derivative” – but let’s see indeed that it works. Apply the time-dependent partial derivative to our wave:

$$\hat{\omega}a(x, t) = i \frac{\partial}{\partial t} a(x, t)$$

But since

$$i \frac{\partial}{\partial t} \exp(-i\omega t) = i(-i)\omega \exp(-i\omega t) = \omega \exp(-i\omega t) \quad \text{since } i(-i) = 1$$

we get that

$$\hat{\omega}a(x, t) = i \frac{\partial}{\partial t} (5 \exp(i(k_1x - \omega_1t)) + 8 \exp(i(k_2x - \omega_2t)))$$

$$\hat{\omega}a(x, t) = \omega_1 5 \exp(i(k_1x - \omega_1t)) + 8\omega_2 \exp(i(k_2x - \omega_2t)) \quad ! \text{It worked!!}$$

The beauty is that now we can take any wave and find the action of the frequency on it. We don't need to even write it anymore as a combination of normal modes!

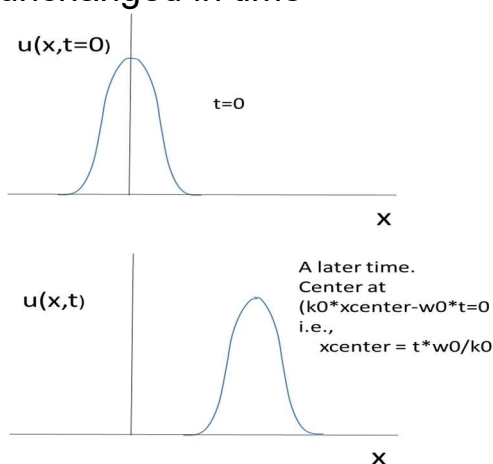
EXAMPLE – given a Gaussian wave, i.e., a Bell-Shaped wave of the form, e.g.,

$$u(x, t) = \exp\left(-\frac{1}{2}(k_0x - \omega_0t)^2\right)$$

- Draw this packet for two different times
- Find the action of the frequency operator on it this wave
- Find the action of the squared frequency operator on this wave

Answer:

- Drawing: u is a “wavepacket” – and it moves with its shape unchanged in time



- Apply the derivative

$$\begin{aligned}\hat{\omega}u &= \frac{i\partial}{\partial t}u = -i\frac{1}{2}2 * (k_0x - \omega_0t)(-\omega_0) \exp\left(-\frac{1}{2}(k_0x - \omega_0t)^2\right) \\ &= -i\omega_0(k_0x - \omega_0t)u(x, t)\end{aligned}$$

- Apply the derivative twice:

$$\hat{\omega}^2u = \hat{\omega}\hat{\omega}u = \frac{i\partial}{\partial t}\frac{i\partial}{\partial t}u = -\frac{\partial^2}{\partial t^2}u = -\frac{\partial^2}{\partial t^2}\exp\left(-\frac{1}{2}(k_0x - \omega_0t)^2\right)$$

A tedious 2nd derivative calculation gives then:

$$-\frac{\partial^2}{\partial t^2}u = \omega_0^2((k_0x - \omega_0t)^2 - 1) \exp\left(-\frac{1}{2}(k_0x - \omega_0t)^2\right)$$

i.e.,

$$\hat{\omega}^2u = \omega_0^2((k_0x - \omega_0t)^2 - 1)u$$

Strange indeed...positive for some x's & t's,, negative for others ...but that's what the math gives.

END OF EXAMPLE.

Similarly we can find the wavevector operator:

$$\hat{k} = -\frac{i\partial}{\partial x}$$

And I leave it up to you to verify that

$$\hat{k}a(x, t) = 5\mathbf{k}_1 \exp(i(k_1x - \omega_1t)) + 8\mathbf{k}_2 \exp(i(k_2x - \omega_2t))$$

as it should.

So now you see that we can **rewrite the classical wave equation,**

$$\frac{\partial^2 f}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 f}{\partial t^2}$$

Instead as

$$\left(\hat{k}^2 - \frac{1}{v^2}\hat{\omega}^2\right)f = 0$$

It is in a sense as we have $\left(\hat{k}^2 - \frac{1}{v^2}\hat{\omega}^2\right)=0$, but not really always, only when acting on physical waves – i.e., which fulfil this equation.

Schrödinger's equation

These equations above were for classical waves. Now let's ask follow Schrödinger (at least simplistically) and ask **what should be a similar equation that waves of electrons should fulfil.**

Well, one thing we learned from de-Broglie is that the momentum is related to the wavelength:

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

so the wavevector k , which I remind you is simplistically

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

is related to the momentum, i.e.,

$$p = \hbar k$$

So the first ingredient will be to **replace the momentum by a momentum operator, i.e.,**

$$\hat{p} = \hbar \hat{k}$$

i.e., the momentum becomes the derivative operator:

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

Similarly, Schrödinger took Planck's relation

$$E = h\nu$$

i.e.,

$$E = \hbar\omega$$

(henceforth E denotes **energy, not electrical field anymore**), and turned it into an operator relation, i.e.,

$$\hat{E} = \hbar \hat{\omega}$$

i.e.,

$$\hat{E} = i\hbar \frac{\partial}{\partial t}$$

Now we know for classical particles without a potential that

$$\frac{p^2}{2M} = E$$

(where M is the mass). So the last ingredient is to write the same equation for waves:

$$\frac{\hat{p}^2}{2M} \Psi = \hat{E} \Psi$$

where

$$\Psi(x, t)$$

is the “wavefunction” representing the electron (or the system more generally).

But

$$\hat{p}^2 = \left(-i\hbar \frac{\partial}{\partial x}\right)^2$$

i.e.,

$$\hat{p}^2 = -\hbar^2 \frac{\partial^2}{\partial x^2}$$

so

$$-\frac{\hbar^2}{2M} \frac{\partial^2}{\partial x^2} \Psi = i\hbar \frac{\partial}{\partial t} \Psi$$

This is the Schrödinger equation for a free particle in 1D. Now let's extend it to include a potential, and to 3D.

First, include a potential. Recall that the conservation of energy in classical mechanics adds a potential very simply:

$$\frac{p^2}{2M} + V(x) = E$$

So here we do the same,

i.e., extend the operator equation above ($\frac{\hat{p}^2}{2M}\Psi = \hat{E}\Psi$) to also include the potential:

$$\left(\frac{\hat{p}^2}{2M} + V(x)\right)\Psi(x, t) = \hat{E}\Psi(x, t)$$

i.e.,

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

This is the complete time-dependent Schrödinger equation in 1D!

Extension to 3D:

In 3D, the wavefunction is now $\psi(\mathbf{r}, t)$, where $\mathbf{r} = (x, y, z)$. We can get immediately the 3D Schrodinger equation by recalling that, classically, in 3D the conservation of energy becomes

$$\frac{\mathbf{p} \cdot \mathbf{p}}{2M} + V(\mathbf{r}) = E$$

where

$$\mathbf{p} \cdot \mathbf{p} \equiv p_x^2 + p_y^2 + p_z^2$$

So quantally, we need to replace the 2nd derivative w.r.t. one variable, $\frac{\partial^2}{\partial x^2}$, by the sum of the 2nd derivative w.r.t. three variables:

$$\mathbf{p} \cdot \mathbf{p} = -\hbar^2 \Delta$$

where we introduced the Laplacian:

$$\Delta \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

Therefore, the Schrödinger equation becomes

$$-\frac{\hbar^2}{2M}\Delta\Psi(\mathbf{r},t) + V(\mathbf{r})\Psi(\mathbf{r},t) = i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r},t)$$

The rest of the quarter will be devoted essentially to solving (and extending) this equation in many circumstances!

Time-dependent vs. Time-Independent Schrödinger equation

The time-dependent Schrödinger equation which we derived is an initial-value equation in time. This generally means that it is of the form

$$\frac{\partial}{\partial t}\Psi(\mathbf{r},t) = \text{something which depends on } \Psi(\mathbf{r},t) \text{ at } t \text{ only.}$$

Therefore, if we know the wavefunction at time t , then we can find the wavefunction at a later time, $t + dt$ (and then repeat the procedure to $t + 2dt, t + 3dt$, etc., i.e., to all future times).

Here, specifically, we show it by realizing that

$$\frac{\partial\Psi(x,t)}{\partial t} \simeq \frac{\Psi(x,t+dt) - \Psi(x,t)}{dt}$$

So plugging to the Schrödinger equation we get

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

i.e.,

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

Note the difference from classical mechanics(CM). In CM we need to know just two things about a particle initially, its position and momentum (or velocity), from which we find how it behaves at later times. In QM we need to know the full initial wavefunction at the initial time, $\Psi(x, t_{start})$.

The time-independent Schrödinger equation: the Hamiltonian and stationary states.

When we talked about propagating modes, $\exp(i(kx - \omega t))$ we have shown that combination of two such propagating modes can lead to a stationary mode, $g(x, t) = g_{stationary}(x) * \exp(-i\omega t)$.

A similar thing happens in QM. But first:

The Hamiltonian

To facilitate, let's define the "Hamiltonian" operator, which for us, at present, is just the sum of the kinetic energy and potential energy operators. We'll write in 1D but our definitions are true in 3D too

$$H \equiv \frac{\hat{p}^2}{2M} + V$$

i.e.,

$$H \equiv -\frac{\hbar^2}{2M} \frac{\partial^2}{\partial x^2} + V(x)$$

i.e.,

$$H\Psi(x, t) \equiv -\frac{\hbar^2}{2M} \frac{\partial^2 \Psi(x, t)}{\partial x^2} + V(x)\Psi(x, t)$$

Note that H is an operator but for convenience and historical reasons we don't give "hats" to it. Henceforth almost everything we'll deal with will be an operator (except for the wavefunctions and for numbers) so to simplify we don't put "hats" on the operators.

Also, H is NOT the enthalpy – no relation (some faint relation actually, which we can ignore). For us, at present, let's think of H as the energy operator.

In terms of H , the Schrödinger equation is simply:

$$i\hbar \frac{\partial \Psi(x, t)}{\partial t} = H\Psi(x, t)$$

Now let's look for product-solutions, i.e.,

$$\Psi(x, t) = \text{something}(t) * \text{something}(x)$$

It is easy to guess that the function in time will be simple, actually of the same form we had before for normal mode, $\exp(-i\omega t)$ (since we are really just looking for normal modes of the Schrödinger's equation). So write down:

$$\Psi(x, t) = \exp(-i\omega t) \psi(x)$$

Let's see what properties this $\psi(x)$ has to fulfil.

Plugging in to $i\hbar \frac{\partial \Psi(x, t)}{\partial t} = H\Psi(x, t)$, we get

$$i\hbar \frac{\partial}{\partial t} \exp(-i\omega t) \psi(x) = H \exp(-i\omega t) \psi(x)$$

i.e.,

$$\hbar\omega \exp(-i\omega t) \psi(x) = \exp(-i\omega t) H\psi(x)$$

Where the RHS is true since H does not involve any derivative in time, so $H \exp(-i\omega t) = \exp(-i\omega t) H$.

Dividing by $\exp(-i\omega t)$, and exchanging the RHS and the LHS, we get the **time-independent Schrödinger equation**:

$$H\psi = E\psi$$

where

$$E = \hbar\omega$$

To conclude: if we can find a solution of the highlighted **time-independent Schrödinger equation**, then a time-dependent solution will follow as

$$\Psi(x, t) = \exp\left(-\frac{iE}{\hbar} t\right) \psi(x)$$

(where the phase factor is simply $\exp(-i\omega t)$).

In chemistry, already in 20A, we talk about molecular orbitals – those are exactly the solution of the time-independent Schrödinger equation for electrons in atoms.

Most of the class will be devoted to finding such orbitals for different cases.

As a side-note, we will later see that the energies of such orbitals are bounded from below for the Coulomb potential – which will be indeed a proof the quantum mechanics solves the first presented challenge that CM faced, i.e., the electron will not damp to the origin.

Part III: Summary.

To summarize this perhaps the most important chapter of the course, we:

- Talked about classical waves
- We understood the wavevector and frequency of waves of the form $\cos(kx-wt)$ etc.
- We saw that in order to “apply”/multiply a general combination of waves by their “frequency” we can replace the frequency by an operator, $\frac{i\partial}{\partial t}$, and similarly the wavevector becomes $\hat{k} = -\frac{i\partial}{\partial x}$
- This, together with $p = \hbar k = -\frac{i\hbar\partial}{\partial x}$, $E = \hbar\omega$ and the classical $H = E$ where $H \equiv \frac{\hat{p}^2}{2M} + V$ led to the time-dependent Schroedinger equation $i\hbar \frac{\partial\Psi}{\partial t} = H\Psi$.
- We understood the Schrödinger equation is an initial value equation.
- We derived then back the time-independent Schrödinger equation, $H\psi = E\psi$, $\Psi = \exp\left(-\frac{iE}{\hbar}t\right)\psi$

IV. Eigenfunctions, Eigenvalues, Operators, Observables.

The time-independent S.E. is also an example of an “eigenvalue equation”. Such an equation always has the form:

$$A f = a f$$

i.e.,

Operator* function = number * same-function

In an eigenvalue equation an operator acts on a function, (known as an eigenfunction) in such a way that the result is that same function multiplied by a constant - the eigenvalue.

In the time-independent S.E. we have:

$$H \psi = E \psi$$

Ψ is said to be an eigenfunction of H with an eigenvalue E .

Now, eigenvalues play a central role in Q.M. in that the eigenvalues of a given operator are the only values that can be observed in measurements of the observable associated with that operator.

E.g. for any given system $H \psi = E \psi$ is, in general, satisfied by numerous (often infinite) pairs of eigenfunctions and eigenvalues $(e_j, \psi_j, j = 1, \dots)$ such that: $H \psi_j = e_j \psi_j$

Such eigenvalue equations apply to the operators of all observables associated with a given system

SIDE NOTE: For such a system, if a measurement is made of the energy of the system, the only possible observable values are the e_j . This is a little tricky to formulate and understand though but all it means is that if the system is in a mixture of states the act of measurement forces it to choose one state with a specific energy; let's leave for now and get back to it in 10 pages. **END OF SIDE NOTE.**

For example, the possible measured values of the x -component of the linear momentum of a particle are the eigenvalues that satisfy the equation:

$$\hat{p}_x \phi_n = a_n \phi_n$$

$$-i\hbar \frac{\partial}{\partial x} \phi_n = a_n \phi_n$$

So

$$\phi_n(x) = \exp\left(i \frac{a_n}{\hbar} x\right)$$

For a general system, the eigenstates of $H\psi_n = e_n\psi_n$ are known as the “stationary-state wave functions” of the system. From these wavefunctions we can determine all the other properties associated with the stationary states of the system.

Hermitian Operators

Let’s look at an operator A .

We’ll only consider **linear Operator** (for meaning, consider your linear Algebra course).

Now, if A corresponds to an observable, its eigenvalues are necessarily **real numbers** and Linear algebra tells us (and we’ll show) that A must be what is called an **Hermitian Operator**.

An Hermitian Operator is defined as one which

$$\int f^*(Ag)dx = \left(\int g^*(Af)dx \right)^*$$

I.e., except for the complex conjugates, the idea is that we can apply A on a function $g(x)$ and then multiply the resulting function (Ag) which is a function of position, by (the complex conjugate) of another function $f^*(x)$ and integrate over all space; or transpose the functions (apply A on f ,

multiply by g^* and then integrate); and the results are complex conjugates of each other (if A is an Hermitian operator).

Since what we observe in experiments are always real quantities, observations in quantum mechanics are intimately tied to Hermitian operators.

So we'll spend a few sections on Hermitian operators

Examples of Hermitian Operators

Let's see some Hermitian (linear) operators.

(i) Derivative: anti-Hermitian. Is the derivative operator Hermitian? No... it is "anti-Hermitian"!

Proof: assume that we deal with functions which go to zero near infinity (a little wishy-washy for this operator, but never mind for now). Then we can integrate by parts so

$$\int g^*(x) \frac{d}{dx} f(x) dx = - \int f(x) \frac{d}{dx} g^*(x) dx$$

since we can interchange the conjugate, and since $(f^*)^* = f$, we can write the RHS as

$$- \int f(x) \frac{d}{dx} g^*(x) dx = - \left(\int f^*(x) \frac{d}{dx} g(x) dx \right)^*$$

So therefore:

$$\int g^*(x) \frac{d}{dx} f(x) dx = - \left(\int f^*(x) \frac{d}{dx} g(x) dx \right)^*$$

So we get this "minus" sign relative to a Hermitian operator. So the derivative operator is anti-Hermitian.

(i'): $i \frac{d}{dx}$ is Hermitian: To make the derivative operator into Hermitian all we need is to multiply the equations above by i .

$$i \int g^*(x) i \frac{d}{dx} f(x) dx = -i \left(\int f^*(x) \frac{d}{dx} g(x) dx \right)^*$$

And since $i^* = -i$ we can insert it into the parentheses on the RHS (with a sign change) to get

$$\int g^*(x) i \frac{d}{dx} f(x) dx = \left(\int f^*(x) i \frac{d}{dx} g(x) dx \right)^*$$

which proves indeed that $i \frac{d}{dx}$ is Hermitian. This is good, as it means that $\hat{p} = -i\hbar \frac{d}{dx}$ is Hermitian, so it will have (see below) real eigenvalues.

Note also that the second-derivative operator is Hermitian, as you can prove.

(ii) The operator of multiplying by x (or x^2, x^3 , any **real** local function of x , such as the potential $V(x)$) -- is Hermitian. The proof is trivial:

$$\begin{aligned} \int g^*(Vf) dx &= \int g^*(x)V(x)f(x) dx = \int f^*(x)V(x)g(x) dx \\ &= \left(\int f^*(x)V(x)g(x) dx \right)^* \quad Q.E.D. \end{aligned}$$

BTW, note that adding two Hermitian operators gives a Hermitian operator. So for example an operator like

$$a = x - i \frac{d}{dx}$$

will be Hermitian. We'll have a lot to talk about this "raising operator" later when we discuss harmonic oscillators.

Hermitian Operators: Real Eigenvalues

Let's see that indeed **Hermitian operators have only real eigenvalues** (it is also the opposite, but let's not worry about the opposite, i.e., that if a linear operator has only real eigenvalues it is Hermitian).

The proof is simple. Take an eigenfunction, $\phi_j(x)$ of A . We know that

$$A\phi_j(x) = a_j\phi_j(x)$$

Define $g(x) = f(j) = \phi_j(x)$. The definition above of an Hermitian operator is $\int f^*(Ag)dx = (\int g^*(Af)dx)^*$ so it gives, once we insert $g(x) = f(j) = \phi_j(x)$, that

$$\int \phi_j^*(A\phi_j)dx = \left(\int \phi_j^*(A\phi_j)dx \right)^*$$

But $A\phi_j = a_j\phi_j$, so

$$\int \phi_j^*(a_j\phi_j)dx = \left(\int \phi_j^*(a_j\phi_j)dx \right)^*$$

Now a_j is a number so we can take it out of the integral, so

$$a_j \int \phi_j^*(x)\phi_j(x)dx = \left(a_j \int \phi_j^*(x)\phi_j(x)dx \right)^*$$

But $\phi_j^*(x)\phi_j(x) = |\phi_j(x)|^2$ and is real, so

$$a_j \int |\phi_j(x)|^2 dx = \left(a_j \int |\phi_j(x)|^2 dx \right)^* = a_j^* \left(\int |\phi_j(x)|^2 dx \right)^*$$

Further, since $|\phi_j(x)|^2$ is real, $\int |\phi_j(x)|^2 dx$ is real so it is equal to its complex conjugate; so

$$a_j \int |\phi_j(x)|^2 dx = a_j^* \int |\phi_j(x)|^2 dx$$

Further, $\int |\phi_j(x)|^2 dx$ cannot be zero unless ϕ_j is trivially zero, so we can divide both sides by it so

$$a_j = a_j^*$$

Therefore, a_j is equal to its complex conjugate, so it must be a real number – Q.E.D.

Next:

Orthogonality of eigenfunctions of Hermitian Operators

We'll now prove the **eigenfunctions of Hermitian operators which have different eigenvalues are orthogonal**:

$$\int \phi_j(x)\phi_n(x)dx = 0 \quad \text{if } a_j \neq a_n$$

In our 20A language, this means that **molecular orbitals** (i.e., eigenfunctions of the Hamiltonian governing a single electron motion) **with different energies** (eigenvalues of the Hamiltonians) **are orthogonal**.

Proof: similar technique to the proof that the a_j are real. Start again with

$$A\phi_j(x) = a_j\phi_j(x)$$

Now multiply on the right with $\phi_n^*(x)$ and integrate

$$\int \phi_n^*(x)A\phi_j(x)dx = \int \phi_n^*(x)a_j\phi_j(x) dx = a_j \int \phi_n^*(x)\phi_j(x) dx$$

Reverse the role of j, n and rewrite the equation above then as

$$\int \phi_j^*(x)A\phi_n(x)dx = a_n \int \phi_j^*(x)\phi_n(x) dx$$

But since A is Hermitian the LHS of the “red” and “magenta” parts are related

$$\int \phi_n^*(x)A\phi_j(x)dx = \left(\int \phi_j^*(x)A\phi_n(x)dx \right)^*$$

And therefore so are the RHS:

$$a_j \int \phi_n^*(x)\phi_j(x) dx = \left(a_n \int \phi_j^*(x)\phi_n(x) dx \right)^*$$

i.e. (no more colors...)

$$a_j \int \phi_n^*(x)\phi_j(x) dx = a_n^* \left(\int \phi_j^*(x)\phi_n(x) dx \right)^*$$

i.e., since $a_n^* = a_n$ as the eigenvalues are real, and since $\left(\int \phi_j^*(x)\phi_n(x) dx \right)^* = \int \phi_n^*(x)\phi_j(x) dx$:

$$a_j \int \phi_n^*(x)\phi_j(x) dx = a_n \int \phi_n^*(x)\phi_j(x) dx$$

i.e.,

$$(a_j - a_n) \int \phi_n^*(x) \phi_j(x) dx = 0$$

It follows that either the eigenvalues are equal ($a_j = a_n$), or, if they are not equal, their overlap is zero, so $\int \phi_n^*(x) \phi_j(x) dx = 0$. But this is exactly the definition of orthogonal orbitals!

BTW, this theorem says nothing about orthogonality of eigenfunctions with the same eigenvalues; such eigenfunctions are called “degenerate” eigenfunctions. For example, you know that the hydrogen atomic 2p orbitals for hydrogen (2s, 2px, 2py, 2pz) have all the same energy.

It turns out we can take degenerate eigenfunctions and make orthogonal combinations of them. (E.g., an SP combination of 2s and 2pz orbitals on the hydrogen atom).

So the end result is that in practice the eigenfunctions of all Hermitian operators we’ll deal with will be orthogonal, whether automatically (when the eigenvalues are different) or made to be so (if the eigenvalues are the same).

Normalizing:

We can “always” scale an eigenfunction to be normalized, i.e.,

$$\int \phi_n^*(x) \phi_n(x) dx = \int |\phi_n(x)|^2 dx = 1$$

(actually, there are problems if we talk about scattered electrons which have an infinitely extended wavefunction, e.g., conduction-electrons in metals, but let’s not worry about this here).

We can therefore summarize that in general the eigenstates are (or can be made to be) orthonormal

$$\int \phi_n^* \phi_j dx = \delta_{jn}$$

where we did not write explicitly $\phi_j(x)$, abbreviating instead ϕ_j . Also, δ_{jn} is the Kronecker delta

$$\delta_{jn} = 1 \text{ if } j = n, \quad 0 \text{ otherwise.}$$

Finally: given any Hermitian operator, it can be shown (we won't) that its eigenfunctions are a complete set, i.e. given a general wavefunction ψ we can write

$$\psi(x) = \sum_j c_j \phi_j(x)$$

and the expansion coefficients c_j are easily extracted for orthogonal eigenstates; for example to get $c_{n=7}$, multiply by $\phi_7(x)$ and integrate to get

$$\begin{aligned} \int \phi_7(x)\psi(x) dx &= \int \phi_7(x) \sum_j c_j \phi_j(x) dx = \sum_j c_j \int \phi_7(x)\phi_j(x) dx \\ &= \sum_j c_j \delta_{7,j} = (c_1 * 0 + c_2 * 0 + \dots + c_6 * 0 + c_7 * 1 + c_8 * 0 + \dots) \\ &= c_7 \end{aligned}$$

i.e., $c_7 = \int \phi_7(x)\psi(x) dx$, i.e., in general

$$c_n = \int \phi_n(x)\psi(x) dx$$

Analogy to classical 3D vectors

This is **completely analogous to saying that if we have a single-3D vector**, $\mathbf{v} = (v_x, v_y, v_z)$, then we can get its components by dot-product with the axis unit vector, e.g.,

$$v_y = \hat{\mathbf{y}} \cdot \mathbf{v}$$

i.e., c_n in QM should be thought of as expansion coefficients in some infinitely big space of eigenfunctions.

Space dimensionality

NOTE ON DIMENSION: In the discussion of Hermitian operators we used "x". But the discussion is completely general. x could refer to

- 1D,
- or for a single particle in the real world it will be 3D

- or even more interestingly, when we consider e.g., a molecule with 50 electrons then x is really $3 * 50 = 150$ dimensional i.e., " x " refers really to $x_1, y_1, z_1, x_2, y_2, \dots, x_{150}, y_{150}, z_{150}$.
- And there's always spin, as we'll see later, which will double the dimensionality further...

QM postulates and Hermitian operators:

- (I) The state of a Q.M.'al system is completely specified by a wave function.

The probability that a particle will be observed at a time in a spatial interval between $\left[x - \frac{dx}{2}, x + \frac{dx}{2} \right]$ (or $[x, x + dx]$) is given by:

$|\Psi(x, t)|^2 dx$ (which is the same as $\Psi^*(x, t)\Psi(x, t)dx$.)

In 3D, the equivalent probability is $|\Psi(\mathbf{r}, t)|^2 d^3r$ where $d^3r \equiv dx * dy * dz$, and $\mathbf{r} = (x, y, z)$.

The wavefunction of a Q.M.'al system must obey the time dependent S.E.:

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

Where H is the Hamiltonian, or total energy operator, associated with the system.

- (II) For every measurable property of a system in classical mechanics such as position, momentum, and energy, and even some that are not known in classical mechanics such as "spin", there exists a corresponding operator in quantum mechanics.

An experiment in the laboratory to measure a value for such an observable is simulated in the theory by operating on the wavefunction of the system with the corresponding operator.

- (III) Properties of Q.M.'al operators:

The eigenvalues of the QM operators corresponding to experimental observables are real

The eigenfunctions of the QM operators corresponding to experimental observables are orthogonal to each other

The normalized eigenfunctions of the QM operators corresponding to experimental observables form an orthonormal basis

- (IV) In any single measurement of the observable corresponding to operator, the only values that will ever be measured are the eigenvalues of the operator.

This is a fancy way of saying that each time we measure a quantity, we force the system to be an eigenstate of that quantity. In average, it will find itself in some probability of being in some value, another probability of another eigenvalue, etc.; but in each measurement it will only be in one eigenstate with one eigenvalue.

For example, when we measure the momentum, we force the system to “choose” its momentum – sometime it will have one value, sometimes another, but each time we measure, as soon as a measurement of A is made, the system’s wavefunction changes. We force it to be in an eigenstate of the momentum operator (or another operator, all depends on our measuring device).

Thus, as soon as a measurement of an operator A is made, the system’s wavefunction changes. In fact, it converts by the very act of the measurement to the eigenfunction of corresponding to the value a_j obtained in the measurement:

$$\Psi_{old} = \sum_j b_j \phi_j \rightarrow \Psi_{new} = \phi_k$$

where the specific state “k” chosen is random (1 or 2 or 3 or 4...etc.) and the **probability P_k for the system to land in a**

specific state k is proportional to $|b_k|^2$ This will be made clearer in the next postulate.

- (V) If a system is in a state characterized by the wavefunction Ψ , and the value of the observable associated with operator is measured once each on many identically prepared systems in this state, then the average value (the expectation value) of these measurements is given by:

$$\langle A \rangle = \frac{\int \Psi^* A \Psi dx}{\int \Psi^* \Psi dx}$$

(where dx in 1D, or d^3r in 3D, etc.).

Note that Ψ should be normalized to 1, i.e., by the 1st postulate $\int \Psi^* \Psi dx = 1$ (since the integral of the probability that the system is somewhere is 1), so then of course

$$\langle A \rangle = \int \Psi^* A \Psi dx$$

Specifically, Quantum particles, in general, give inconsistent answers (measurement outcomes):

- (i) What are the possible outcomes? Discrete or continuous?
⇒ determined by the observable or its corresponding operator.
- (ii) What are the probabilities?
⇒ determined by the wavefunction
- (iii) What is the average (expectation value)?
⇒ algebraic calculation

Expectation values: details

The next few pages will deal with expectation values.

Let us examine this by expanding Ψ in terms of the complete, orthonormal set of eigenfunctions ϕ_j of A (so $A\phi_j = a_j\phi_j$)

$$\Psi = \sum_j b_j \phi_j \quad (\text{completeness})$$

Normalization: the normalization=1 condition is

$$\begin{aligned}1 &= \int \Psi^* \Psi dx = \int \left(\sum_j b_j \phi_j \right)^* \sum_l b_l \phi_l dx \\1 &= \sum_j \sum_l b_j^* b_l \int \phi_j^* \phi_l dx \\1 &= \sum_j \sum_l b_j^* b_l \delta_{jl}\end{aligned}$$

But we can remove the δ_{jl} and the summation over l , as for a given j the only value of l contributing is $l = j$ due to the delta-function overlap, $\int \phi_j^* \phi_l dx = \delta_{jl}$. Therefore

$$1 = \sum_j b_j^* b_j = \sum_j |b_j|^2$$

is the condition on the overlap.

BTW, note that this is completely analogous to saying that if a 5D vector $v = (v_1, v_2, \dots, v_5)$ is normalized, $v^* \cdot v = 1$, then $\sum_{j=1, \dots, 5} |v_j|^2 = 1$, i.e., $|v_1|^2 + |v_2|^2 + \dots + |v_5|^2 = 1$

Expectation value: Similarly, the expectation value of A is defined in a shortcut as:

$$\langle A \rangle \equiv \int \Psi^* A \Psi dx$$

So based on the expansion of Ψ we get

$$\langle A \rangle = \int \left(\sum_j b_j \phi_j \right)^* A \sum_l b_l \phi_l dx$$

i.e.,

$$\langle A \rangle = \sum_j \sum_l b_j^* b_l \int \phi_j^* A \phi_l dx$$

But we chose a basis of eigenfunctions of A , so $A\phi_j = a_j\phi_j$

$$\int \phi_j^* A\phi_l dx = a_j \int \phi_j^* \phi_l dx = a_j \delta_{jl}$$

So

$$\langle A \rangle = \sum_j \sum_l b_j^* b_l a_j \delta_{jl}$$

i.e.

$$\langle A \rangle = \sum_j |b_j|^2 a_j$$

In words, the average (expectation) value of A is given by a weighted sum over all the eigenvalues (possible measured values) of A , the weights being determined by the contributions of the eigenfunctions of A , as measured by the expansion coefficients b_j , to the system's wavefunction, Ψ .

Now, recall that in general the average of any observable quantity can be written as:

$$\langle x \rangle = \sum_j x_j P(x_j)$$

where $P(x_j)$ is the probability of observing x_j when making a measurement of x

Comparing this with the above expression for

$$\langle A \rangle = \sum_j P(a_j) a_j = \sum_j |b_j|^2 a_j$$

we see that $P(a_j) = |b_j|^2$

The probability of measuring a_j is given by the square of the coefficient of in the expansion of Ψ . i.e., b_j is a probability amplitude.

Measurements – Conclusion

Note the probabilistic nature of Q.M. Generally, in any given measurement of A it is possible to observe numerous values. (Only when $\Psi = \phi_j$ i.e., a single eigenstate of A can we know definitively what will be observed) .

Generally, the best we can do is determine the **probability** that a particular eigenvalue of A will be measured.

On the other hand we can determine $\langle A \rangle$ definitively . That is because the average value of A is obtained by $N \rightarrow \infty$ separate measurements on identical systems.

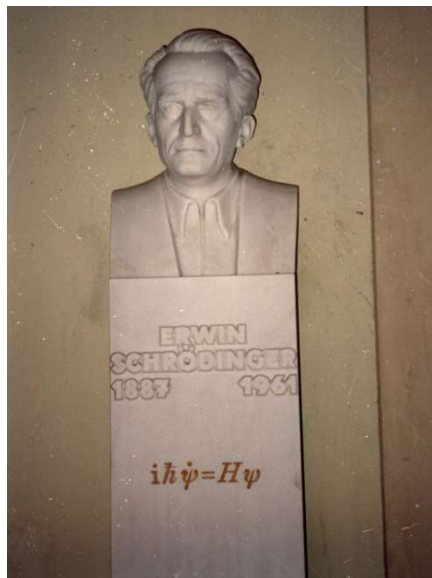
Summary: Quantum Mechanics at Work --Two Steps.

Step #1: Solve the Schrodinger equation to obtain the spatial distribution + time evolution of the quantum wavefunction

Step #2: Use Max Born's statistical interpretation (Postulate II-V above) to interpret the resulting quantum wavefunction

See also:

<https://www.youtube.com/watch?v=uq1h6ig61yl&playnext=1&list=PLFCF1FC91BEF2D365>



Henceforth: let's solve simplified and real systems!

Time-Dependent Wavepackets and Norm Conservation.

We talked a lot about the interpretation of a wavefunction as a probability amplitude, i.e.,

$$P(x, t) = |\Psi(x, t)|^2$$

and

$$\int P(x, t) dx = 1$$

i.e., the probability for the particle (or particles in the case of a multi-particle wavefunction) to be somewhere in space is 1.

We actually need to do a consistency check, to prove that the Schrödinger equation “respects” the conservation of the norm. Let’s do that in two different ways. The first is more general and based directly on the Schrödinger equation, while the 2nd will be based on expansion in eigenstates

1st Norm conservation Proof: From the Schrödinger wavefunction directly:

Let’s define the “total probability” (which we want to show is constant, does not change in time)

$$Q(t) \equiv \int P(x, t) dx = \int |\Psi(x, t)|^2 dx = \int \Psi^*(x, t) \Psi(x, t) dx$$

Let’s differentiate $Q(t)$ to show that its derivative (in time) is zero, i.e., it does not change. (Our proof will make some assumptions on the form of H so we could prove things faster, i.e., our result will be general but the proof will not be):

$$\begin{aligned} \frac{d}{dt} Q(t) &= \int \frac{\partial}{\partial t} (\Psi^*(x, t) \Psi(x, t)) dx \\ \frac{d}{dt} Q(t) &= \int \frac{\partial \Psi^*(x, t)}{\partial t} \Psi(x, t) dx + \int \Psi^*(x, t) \frac{\partial \Psi(x, t)}{\partial t} dx \end{aligned}$$

But

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

And therefore when we conjugate (presuming the Hamiltonian is real, there are some cases it isn't but we won't worry about them now):

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

(where we used the fact that when we multiply two complex numbers, the conjugate of the product is the product of the conjugate – and here one of the complex numbers is $-i$ so its conjugate is i).

Inserting these two “blue” equations into the expression for $\frac{d}{dt} Q(t)$ gives:

$$\frac{d}{dt} Q(t) = -\frac{i}{\hbar} \int \Psi(x, t) (H\Psi^*(x, t)) dx + \frac{i}{\hbar} \int \Psi^*(x, t) (H\Psi(x, t)) dx$$

and changing the order gives

$$\frac{d}{dt} Q(t) = \frac{i}{\hbar} \left(\int \Psi^*(x, t) H\Psi(x, t) dx - \int \Psi(x, t) H\Psi^*(x, t) dx \right)$$

I claim that the term in parentheses is zero. The Hamiltonians we are usually dealing with are either a local potential or a second derivative. For the local potential this is simple:

$$\int \Psi^*(x, t) V(x) \Psi(x, t) dx - \int \Psi(x, t) V(x) \Psi^*(x, t) dx = 0$$

For the 2nd derivative (kinetic) term, $-\frac{\hbar^2}{2M} \frac{\partial^2}{\partial x^2}$ this is slightly more complicated, but is still true

$$\int \Psi^*(x, t) \frac{\partial^2}{\partial x^2} \Psi(x, t) dx - \int \left(\frac{\partial^2}{\partial x^2} \Psi^*(x, t) \right) \Psi(x, t) dx = 0$$

as you can show by integration by parts.

Therefore, to conclude, the “green” equation above is zero, so $Q(t)$ is conserved, so $\int P(x, t) dx$ is conserved, so if initially it is 1 (as it should be) it remains 1.

2nd Norm conservation Proof: Expansion in eigenstates

To warm up on the other proof, let's prove first that if we expand the wavefunction and the sum of the abs-squared expansion coefficients is 1, the wavefunction is normalized:

$$\Psi(x, t) = \sum_j a_j(t) \phi_j(x), \quad \sum_j |a_j(t)|^2 = 1 \quad \rightarrow \quad \int \Psi^*(x, t) \Psi(x, t) dx = 1$$

To exemplify even that “warmup” proof, let's take a specific case. For example, say we prepared (experimentalists could do that now) the initial wavefunction to be an equal combination of the 1st and 2nd eigenstates

$$\Psi(x, t = 0) = \frac{1}{\sqrt{2}} \phi_1(x) + \frac{1}{\sqrt{2}} \phi_2(x)$$

(which is a specific case of the general proof, with $a_1 = a_2 = \frac{1}{\sqrt{2}}$, $a_3 = a_4 = \dots = 0$).

First, let's see that this combination is normalized, like a wavefunction should be:

$$\begin{aligned} \int \Psi(x, t = 0)^* \Psi(x, t = 0) dx &= \\ \int \left(\frac{1}{\sqrt{2}} \phi_1(x) + \frac{1}{\sqrt{2}} \phi_2(x) \right)^* \left(\frac{1}{\sqrt{2}} \phi_1(x) + \frac{1}{\sqrt{2}} \phi_2(x) \right) dx &= \\ = \frac{1}{2} \int (\phi_1(x) + \phi_2(x))^* (\phi_1(x) + \phi_2(x)) dx &= \\ = \frac{1}{2} \left(\int \phi_1^*(x) \phi_1(x) dx + \int \phi_1^*(x) \phi_2(x) dx + \int \phi_2^*(x) \phi_1(x) dx \right. & \\ \left. + \int \phi_2^*(x) \phi_2(x) dx \right) = \frac{1}{2} (1 + 0 + 0 + 1) = 1 \quad Q.E.D. & \end{aligned}$$

Now we can prove our assumption (“green” equation above) for the general case. Say

$$\Psi(x, t) = \sum_j a_j \phi_j(x)$$

and $\sum_j |a_j(t)|^2 = 1$. Then:

$$\begin{aligned} \int \Psi(x,t)^* \Psi(x,t) dx &= \int \left(\sum_j a_j(t) \phi_j(x) \right)^* \left(\sum_l a_l(t) \phi_l(x) \right) dx \\ &= \sum_{jl} a_j(t) a_l^*(t) \int \phi_j(x) \phi_l(x) dx = \sum_{jl} a_j^*(t) a_l(t) \delta_{jl} = \sum_j a_j^*(t) a_j(t) \\ &= \sum_j |a_j(t)|^2 = 1 \quad Q.E.D. \end{aligned}$$

So that proves the “green” equation above.

Now note, that if we choose states we use to be the eigenstates (a different word for eigenfunctions) of H then

$$a_j(t) = \exp\left(-\frac{iE_j t}{\hbar}\right) a_j(0)$$

i.e.,

$$\Psi(x,t) = \sum_t a_j(0) \exp\left(-\frac{iE_j t}{\hbar}\right) \phi_j(x)$$

i.e., the amplitude of each wavefunction is multiplied by a time-dependent phase $\exp(-iE_j t)$. The proof of this “yellow” equation (expansion of the time-dependent wavefunction in terms of eigenstates) is not difficult, we just apply the Schrödinger equation:

$$\begin{aligned} i\hbar \frac{\partial \Psi}{\partial t} &= i\hbar \frac{\partial}{\partial t} \sum_t a_j(0) \exp\left(-\frac{iE_j t}{\hbar}\right) \phi_j(x) \\ i\hbar \frac{\partial \Psi}{\partial t} &= \sum_t a_j(0) \left(-i\hbar \frac{iE_j}{\hbar}\right) \exp\left(-\frac{iE_j t}{\hbar}\right) \phi_j(x) \end{aligned}$$

$$i\hbar \frac{\partial \Psi}{\partial t} = \sum_t a_j(0) E_j \exp\left(-\frac{iE_j t}{\hbar}\right) \phi_j(x)$$

And on the other hand,

$$H\Psi = H \sum_t a_j(0) \exp\left(-\frac{iE_j t}{\hbar}\right) \phi_j(x)$$

$$H\Psi = \sum_t a_j(0) \exp\left(-\frac{iE_j t}{\hbar}\right) H\phi_j(x)$$

$$H\Psi = \sum_t a_j(0) \exp\left(-\frac{iE_j t}{\hbar}\right) E_j \phi_j(x)$$

So the two new “blue” equations are equal so indeed $H\Psi = \frac{i\hbar\partial\Psi}{\partial t}$. Therefore indeed $a_j(t) = \exp\left(-\frac{iE_j t}{\hbar}\right) a_j(0)$, and therefore (since the absolute value of $\exp\left(-\frac{iE_j t}{\hbar}\right)$ is one):

$$|a_j(t)| = |a_j(0)|$$

Therefore, if $\sum_j |a_j|^2 = 1$ initially, it will be 1 at all times, and then, based on the “green” equation above, the norm of the wavefunction will be conserved.

Now we can finally get to the simplest (yet quite rich) world of piecewise constant potentials:

Constant potential & piecewise-constant potentials.

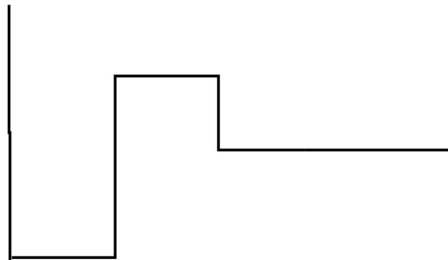
Introduction: relevance to chemistry

Piecewise-constant potentials are both simple and contain a lot of the physics and chemistry of many relevant problems.

For example, a bowl-shaped potential:



Describes the potential that electrons feel in a metal – a low value inside and a higher potential outside. Similarly, a potential with a finite barrier,



Describes pre-dissociation where a particle (electron or proton or a molecule) needs to “tunnel” through a barrier to leave.

We’ll therefore devote some time to such potentials, as the math is either trivially or easily solvable and there’s a lot of science.

Let’s warmup with the simplest math description.

In a region where the potential is constant, $V(x) = V = \text{const.}$, the solution of the time-independent Schrödinger equation

$$H\psi = E\psi$$

i.e., of:

$$-\frac{\hbar^2}{2M} \frac{\partial^2 \psi(x)}{\partial x^2} + V\psi(x) = E\psi(x)$$

depends on whether $V < E$ or $V > E$.

Oscillating region

For $V > E$ we can write ψ as an oscillating combination of sin and cos or of $\exp(ikx)$ and $\exp(-ikx)$

$$\psi(x) = a \cos kx + b \sin kx = c \exp(ikx) + d \exp(-ikx) \quad (\text{for } V < E)$$

where a, b, c, d are constants, and I leave it up to you to find "c" and "d" here in terms of "a" and "b", and

$$k = \frac{p}{\hbar}$$

and "p" is the "local momentum":

$$p = \sqrt{2M(E - V)}.$$

Proof: take, e.g., $\psi(x) = \cos(kx)$

$$-\frac{\hbar^2}{2M} \frac{\partial^2 \psi(x)}{\partial x^2} = -\frac{\hbar^2}{2M} \frac{\partial^2 \cos(kx)}{\partial x^2} = \frac{\hbar^2 k^2}{2M} \cos(kx) = (E - V)\psi(x) \quad Q.E.D.$$

Tunneling region

For $V > E$ the w.f. is in the so-called tunneling region:

$$\psi(x) = \alpha \exp(-\kappa x) + \beta \exp(\kappa x) \quad V > E$$

$$\kappa = \frac{\sqrt{2M(E - V)}}{\hbar}$$

Boundary conditions:

The discussion above makes it seem as if there's always a solution to the Schrödinger equation. But in reality some energies are allowed, some are not – this is a consequence of the boundary conditions. Rather than detail it mathematically, we'll consider some cases and details will emerge as we go along.

EXAMPLES:

I. Flat potential (free particle):

If $V(x) = V = \text{const}$ everywhere, then the solution must be (we use the exp functions, not cos and sin):

$$\psi(x) = c \exp(ikx) + d \exp(-ikx)$$

NOTE THAT the $\exp(ikx)$ wave is “rightward propagating”, while the $\exp(-ikx)$ wave is “leftward propagating”; to see it lets multiply by the time-dependent phase factor:

$$\Psi(x, t) = \exp\left(-it \frac{E}{\hbar}\right) \psi(x) = c \exp(i(kx - \omega t)) + d \exp(-i(kx + \omega t)),$$

$$\omega \equiv \frac{E}{\hbar}$$

Note that the $kx - \omega t$ factor indicates that indeed the wave is rightward propagating (when the phase is, for example, 0, then when t increases x increases).

Such a solution is feasible for ALL energies above V , $E > V$.

Uncertainty principle example:

Note that these wavefunctions are “extended”, and have large amplitude everywhere.

This is an example of the uncertainty principle (which we do not devote a lot of time to, but you should be aware of it):

$$\delta x * \delta p > \frac{\hbar}{4}$$

where δx is the uncertainty in position, etc. for δp .

Energy must be higher than the potential at least somewhere

BTW, if the potential is flat everywhere, then a tunneling solution is not possible since $\alpha \exp(-\kappa x) + \beta \exp(\kappa x)$ will explode at both positive and negative x ($\exp(\kappa x)$ explodes at positive x , and $\exp(-\kappa x)$ explodes at negative x).

This is an example of a principle which I won't prove, that the energy needs to be higher than the potential in at least some region of space.

(Classically, it must be of course higher than the potential everywhere the particle is; quantum mechanically the energy can be lower the potential somewhere, but not everywhere – see discussion below of semi-infinite well and infinite-well with a barrier inside)

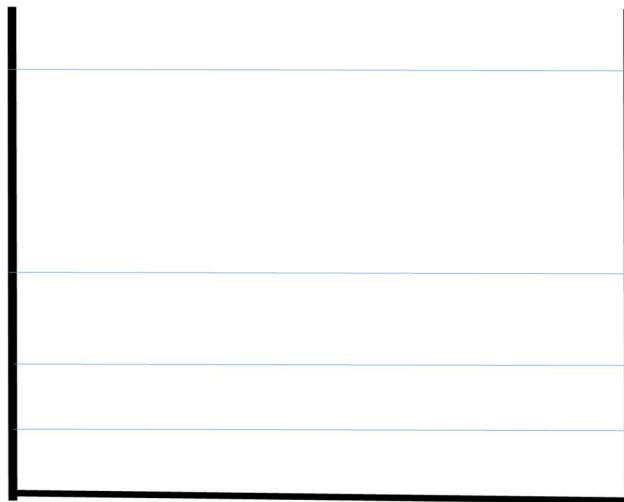
II. Infinite well

The opposite case to an everywhere-flat potential is an infinite well,

$$V(x) = V_1 \quad \text{for } 0 < x < L$$
$$V(x) = \infty \quad \text{for } x < 0 \text{ or } L < x$$

Where V_1 is some constant.

Such a potential (in black) and the associated eigenvalues we'll derive (blue) is shown below:



We know from our initial discussion that in the flat region (between 0 and L) the wavefunction must be some combination of $\sin(kx)$ and $\cos(kx)$ for some k 's.

But here we have also boundary conditions. The wavefunction cannot live (i.e., is zero) in the regions of infinite potential, and it can be shown that it is always continuous, therefore,

$$\psi(x = 0) = \psi(x = L) = 0$$

The condition that ψ is zero at the origin implies that it cannot have a “ $\cos(kx)$ ” term, i.e., it is of the form $\sin(kx)$. The condition that it vanishes at $x = L$ implies therefore that the k values are discretized:

$$k_n = \frac{\pi}{L}, \frac{2\pi}{L}, \frac{3\pi}{L}, \dots = \frac{\pi}{L}n \quad n = 1, 2, \dots$$

since $\sin(k_n * L) = \sin\left(\frac{\pi}{L} * n * L\right) = \sin(n\pi) = 0$.

This is, BTW, exactly the same for waves in a violin or guitar, etc.

Thus, the wavefunctions are proportional to $\sin(k_n x)$:

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin(k_n x)$$

The proportionality constant $\sqrt{\frac{2}{L}}$ ensures that the wavefunctions are normalized:

$$\int |\psi_n(x)|^2 dx = \frac{2}{L} \int_0^L \sin^2(k_n x) dx$$

But

$$\sin^2 y = \frac{1 - \cos(2y)}{2}$$

So

$$\int |\psi_n(x)|^2 dx = \frac{2}{L} * \frac{1}{2} * \left(\int_0^L dx - \int_0^L \cos(2k_n x) dx \right) = \frac{1}{L} \left(L + \frac{1}{2k_n} \sin(2k_n L) \right)$$

But $\sin(2k_n L) = \sin(2n\pi) = 0$ so

$$\int |\psi_n(x)|^2 dx = 1$$

BTW, we don't even have to prove that the eigenstates are orthogonal ($\int \phi_n(x)\phi_m(x)dx = 0$ if $n \neq m$), that comes automatically from the fact that these functions are eigenstates of H .

What about the eigenvalues? Plugging to

$$H\psi_n = E_n\psi_n$$

It is easy to see (we've actually seen it earlier) that

$$E_n = \frac{\hbar^2}{2M} k_n^2 + V_1$$

i.e., in this case

$$E_n = \frac{\hbar^2\pi^2}{2ML^2} n^2 + V_1 = \gamma + V_1, \quad 4\gamma + V_1, \quad 9\gamma + V_1 \dots$$

$$\gamma \equiv \frac{\hbar^2\pi^2}{2ML^2}$$

Example: Say we approximate very roughly the potential in a semiconductor cluster as being flat, of size typically 5eV (i.e., the potential is about -5eV in the cluster). Such clusters can be made at different sizes. Say we take a diameter of $L = 2\text{nm} = 20 \text{ Angstrom} \sim 38 \text{ bohr}$, Of course, such cluster is 3D and we solved for 1D problems but let's dare and apply what we learned.

It turns out that for M we should not take the electron mass since the mass is "modified" due to the interaction of an electron with all the other ones. The correct mass to take is the "effective mass", which for Si is about $0.19m_e$. Then:

$$\gamma = \frac{\hbar^2\pi^2}{2 * 0.19m_e * (40 \text{ bohr})^2}$$

Side note: Atomic units:

The good news is that if we use atomic units, i.e., units where the fundamental energy unit is 1 Hartree (=27.2114 eV), the fundamental length is 1 bohr and the fundamental mass is m_e (the mass of the electron), then in these units $\hbar = 1$. (BTW, the fundamental time unit in atomic unit is $\frac{\hbar}{\text{Hartree}} = 2.4 * 10^{-17} \text{ s} = 0.024 \text{ fs}$)

Therefore, we can calculate γ in atomic units, so

$$\gamma = 1 * \frac{\pi^2}{2 * 0.19 * 1 * (40^2)} = 0.018$$

This γ has energy units, i.e.,

$$\begin{aligned} \gamma &= 0.018 \text{ Hartree} \\ &= 0.018 * 27 \text{ eV} \sim 0.48 \text{ eV} \sim 0.5 \text{ eV} \end{aligned}$$

So we'll expect that the energy levels will be about

$$E = V + 0.5 \text{ eV} * n^2, \quad n = 1, 2, 3 \dots$$

i.e., ($V = -5 \text{ eV}$):

$$\begin{aligned} E_1 &= -4.5 \text{ eV}, \\ E_2 &= -5 + 4 * 0.5 = -3 \text{ eV}, \\ E_3 &= -5 + 9 * 0.5 \text{ eV} = -0.5 \text{ eV}, \quad \text{etc} \end{aligned}$$

Obviously, already by the 3rd level we really cannot expect these energy levels to be accurate since the quantum dot is not an infinite well but is a finite well, of depth 5eV, as mentioned (close to the picture in page 63), but for the lowest two levels it should be more or less OK (at most in the right ballpark, since we used 1D rather than 3D).

So for example if we de-excite an electron from $n = 2$ to $n = 1$ we expect that the energy difference will be 1.5eV. That means that the photon will have an optical frequency, and a red color (or near IR).¹

Our whole derivation was of course not completely quantitative, i.e., we are probably wrong by a factor of ~ 2 , since a quantum dot is 3D; but it does teach us that quantum dots with different diameters will emit at different wavelengths, i.e., will have different colors.

¹ This is because red photons, with wavelength of 700nm, have a frequency of $\nu = \frac{c}{\lambda} = \frac{3 * 10^8 \frac{\text{m}}{\text{s}}}{700 * 10^{-9} \text{m}} = 4.3 * 10^{14} \text{ Hz}$ and therefore, since $h = 4.13 * 10^{-15} \text{ eV} * \text{ Hz}$, an energy of $h\nu = 1.8 \text{ eV}$; purple photons with $\lambda = 400 \text{ nm}$ will have an energy that 7/4 times higher, i.e., 3eV.

Such dots actually find much use now in LED-replacement type TV's and much of the achievements in that field are based on an initial work by Prof. Shimon Weiss (from our department!) and his colleagues in the late 1990's and early 2000's.

END OF EXAMPLE.

Wavepackets in an infinite well and recurrences.

Interesting things happen when we consider functions that are originally a combination of eigenstates. Such states are called "Wavepackets"

Recall that we proved in the earlier chapter that the norm of such wavepackets is conserved. But the wavepackets themselves will change in time, sloshing back and forth.

Let's start in a combination

$$\Psi(x, t = 0) = \frac{1}{\sqrt{2}}\phi_1 + \frac{1}{\sqrt{2}}\phi_2$$

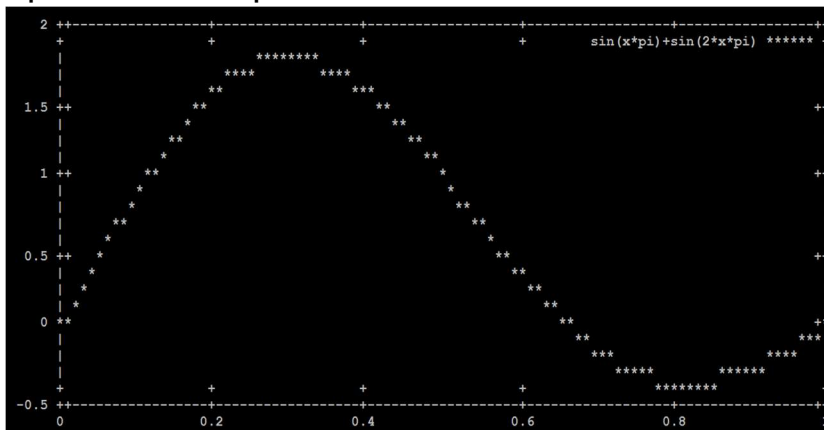
and let's assume for simplicity that the well-range is 1, i.e., $L = 1$ (simplifies the math below). Then:

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

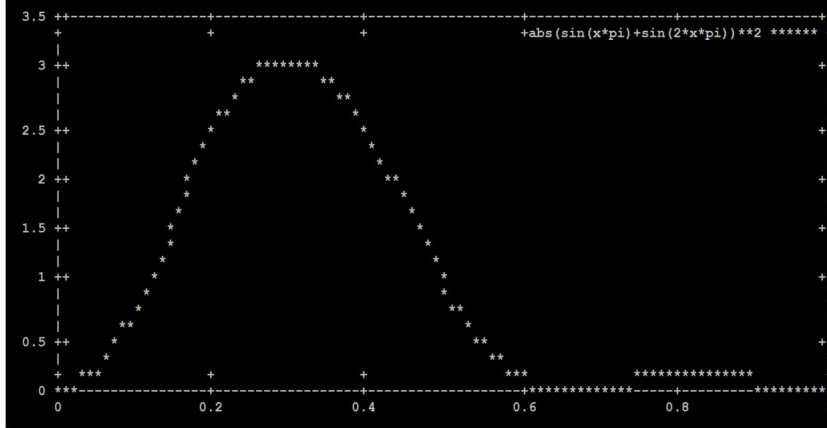
i.e.,

$$\Psi(x, t = 0) = \sin(\pi x) + \sin(2\pi x)$$

I plot this wavepacket below:



You notice that it is mostly on one side (left), as is even more evident when we plot the density associated with it, $|\Psi(x, t = 0)|^2$



Now let's see what happens at a time $t = \hbar \frac{\pi}{2E_1}$. Then,

$$\exp\left(-\frac{iE_1 t}{\hbar}\right) = \exp\left(-\frac{i\hbar \frac{\pi}{2E_1} E_1}{\hbar}\right) = \exp\left(-\frac{i\pi}{2}\right) = -i$$

While, since in our case $E_2 = 4E_1$

$$\exp\left(-\frac{iE_2 t}{\hbar}\right) = \exp\left(-4\frac{iE_1 t}{\hbar}\right) = \exp(-2i\pi) = 1.$$

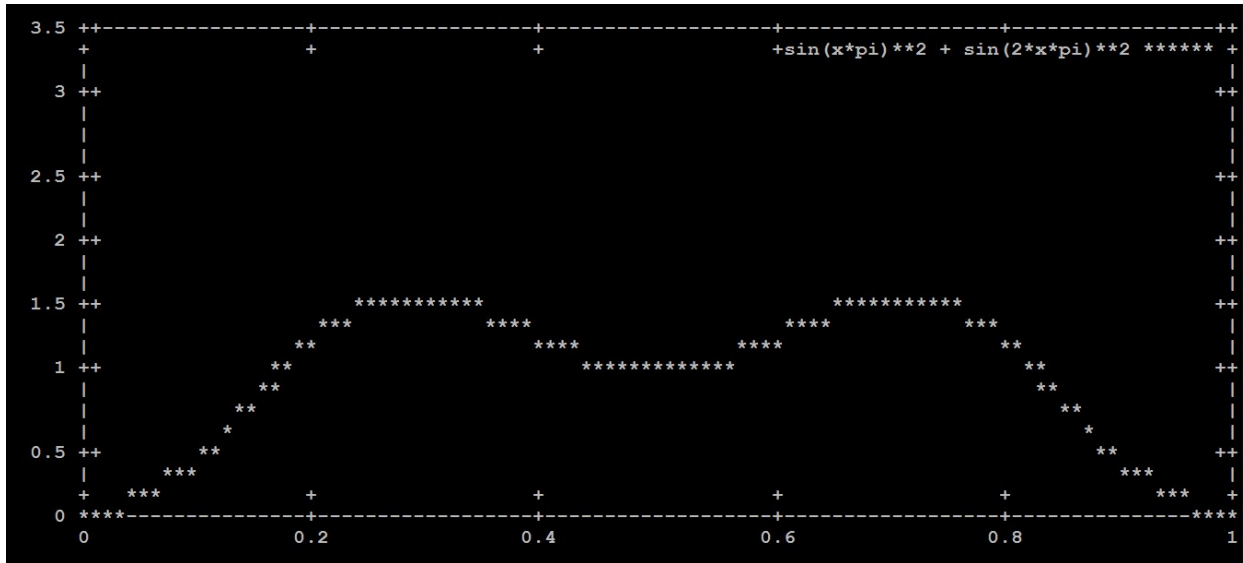
So at that time

$$\Psi\left(x, t = \hbar \frac{\pi}{2E_1}\right) = -i \sin(\pi x) + \sin(2\pi x)$$

So the real part of the wavefunction is essentially (up to $\sqrt{2}$) $\phi_2(x)$ and the imaginary part is $-\phi_1(x)$. The density of that wavefunction is then

$$\left|\Psi\left(x, t = \hbar \frac{\pi}{2E_1}\right)\right|^2 = \sin^2(\pi x) + \sin^2(2\pi x)$$

And is plotted below; it is symmetric, the wavefunction “spread” to the whole well equally on the left and right:



Finally, at $t = \hbar \frac{\pi}{E_1}$ (i.e., twice the previous time), the wavefunction sloshed all the way to the right, i.e., is a mirror image of the initial wavefunction, since

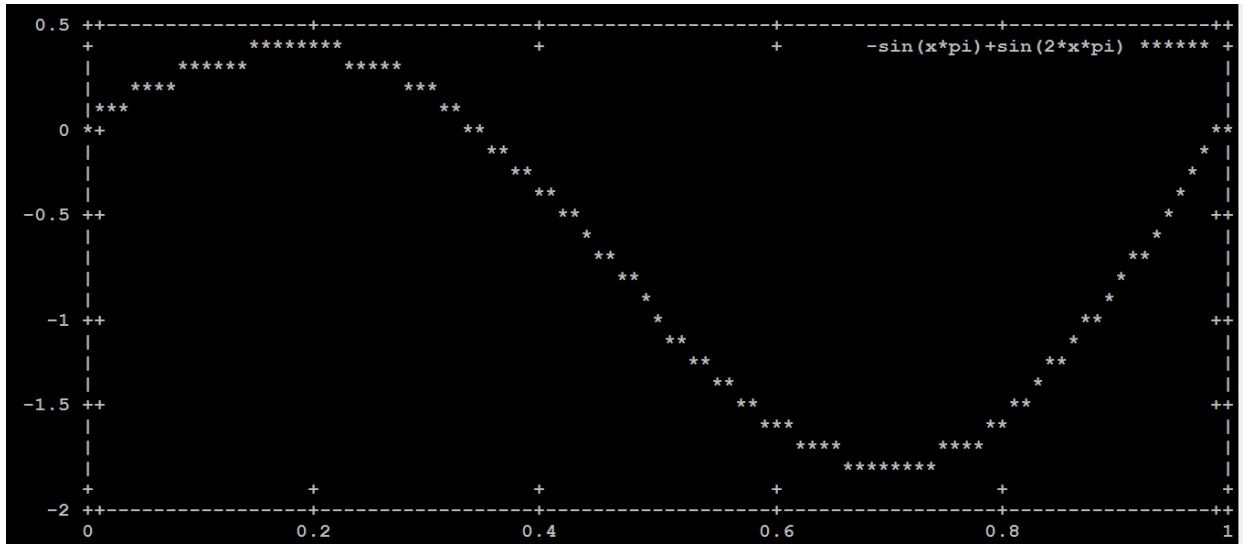
$$\exp\left(-\frac{iE_1 t}{\hbar}\right) = \exp(-i\pi) = -1$$

$$\exp\left(-\frac{iE_2 t}{\hbar}\right) = \exp(-4\pi i) = 1$$

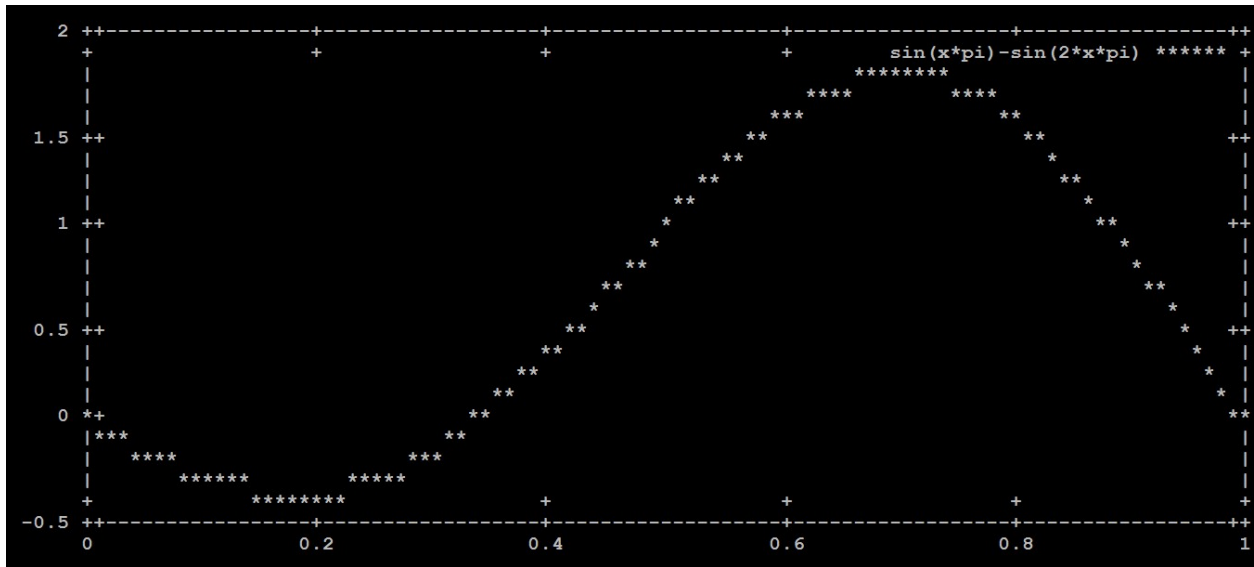
So

$$\Psi\left(x, t = \hbar \frac{\pi}{E_1}\right) = -\sin(\pi x) + \sin(2\pi x)$$

i.e., the wavefunction looks like:



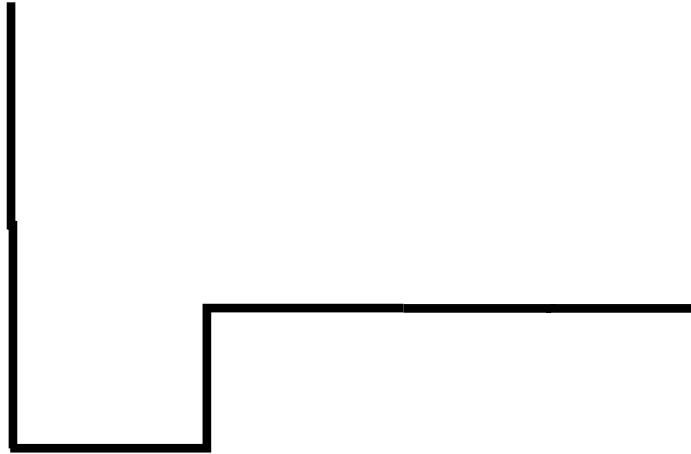
And if we plot its negative, you'll see that it is the same wavefunction initially, except for the mirror image:



In conclusion: this simple example has shown us how a general wavepacket sloshes back and forth in time!

HALF INFINITE WELL:

The last case we'll consider is the half-infinite well



i.e.,

$$V(x) = \begin{cases} \infty & x < 0 \\ V_1 & 0 < x < L \\ V_2 & L < x \end{cases}$$

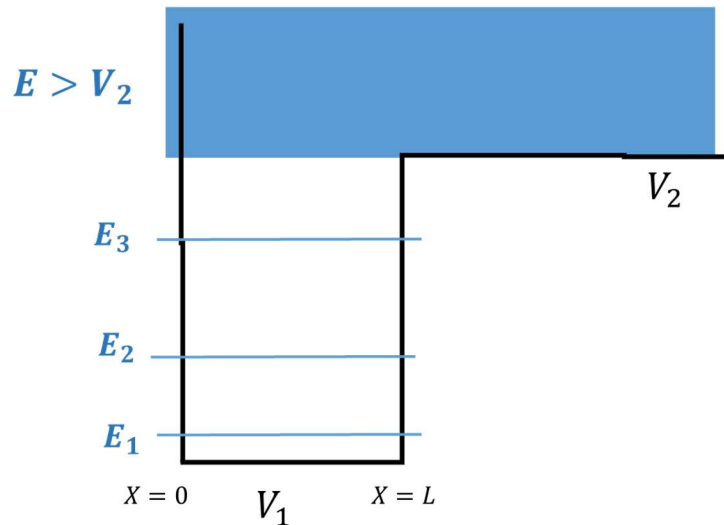
(where of course $V_1 < V_2$)

We know that again $\psi(x = 0) = 0$ since the w.f. is zero in the infinite potential region. Therefore, for the well region

$$\psi(x) = a \sin(k_1 x) \quad 0 < x < L$$

where a is some constant, and $k_1 = \frac{p_1}{\hbar}$, $p_1 = \sqrt{2M(E - V_1)}$

What about the region to the right of $x = L$? For that we need to separate two cases (see figure below)



(i) $E < V_2$ (and of course always $E > V_1$, since we learned that the energy must be higher the potential at least somewhere)

Then, we know that

$$\psi(x) = \beta \exp(-\kappa_2 x) \quad x > L$$

$$\kappa_2 = \frac{\sqrt{2M(V_2 - E)}}{\hbar}$$

How to determine E? The idea is quite simple. It turns out that even though the potential is discontinuous at $x = L$, the wavefunction and its derivative are both continuous at $x = L$ (formally, the derivative is continuous since the potential is not infinite on either side; this is in contrast to $x = 0$ where the potential is infinite on the left so the derivative is not continuous at $x = 0$).

The wavefunction continuity at $x = L$ means that

$$a \sin(k_1 x) = b \exp(-\kappa_2 x) \quad \text{for } x = L$$

i.e.,

$$a \sin(k_1 L) = b \exp(-\kappa_2 L)$$

And the continuity of the derivative means that

$$\frac{d}{dx} a \sin(k_1 x) = \frac{d}{dx} b \exp(-\kappa_2 x) \quad \text{for } x = L$$

i.e.,

$$a k_1 \cos(k_1 L) = -\kappa_2 \exp(-\kappa_2 L)$$

By dividing both “blue” equations by each other (LHS by LHS, and RHS by RHS) we get:

$$k_1 \cot(k_1 L) = -\kappa_2$$

Note that for a given system, i.e., for a given M, L, V_1 and V_2 , the only unknown in both sides of the equation is the energy E (once we know the energy we can determine k_1 and κ_2) so in principle we can just plot the LHS as a function of energy and the RHS too; then:

The values of energy where there will agreement, i.e., the eigenvalues, will be those where the LHS and RHS of the yellow equation match!

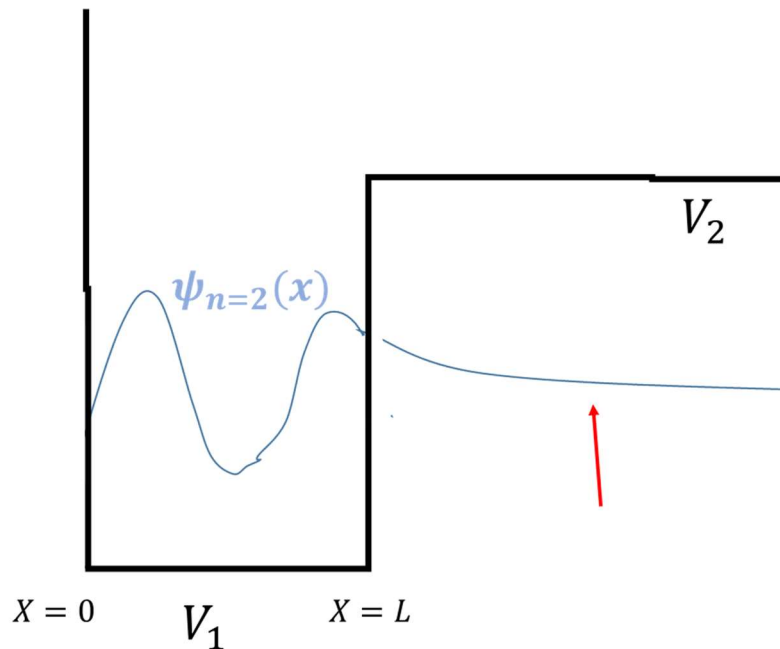
Exercise (at home); consider an electron ($M = M_{electron} = 1$ in atomic units) in a well of depth

$$V_1 = -8.68eV \sim -0.32 \text{ Hartree},$$

and $V_2 = 0$, and a well length of $L = 5.3 \text{ Angstrom} \approx 10 \text{ bohr}$. Plot (e.g., using excel or Wolfram Alpha online) the RHS and left hand side of the yellow equation above and use them to determine the numerical values of the eigenvalues that are below V_2 , i.e., below 0. How many such eigenvalues are there? Of course use atomic units in your calculation to simplify them.

Physical relevance.

The case we just considered is physically kind of similar to the scanning tunneling microscope. In the microscope, you can imagine that we put a probe on the x-axis far to the right of L , i.e., where the wavefunction is deeply tunneling. A probe of such wavefunction is in the figure below



Where the red arrow points to a position where we measure the probability to find an electron. (In practice we will measure the current of the electrons, when we inject them from the left, but the idea is similar).

The probability to find an electron at a point “x” where the arrow is, i.e., to the right of $x = L$, is

$$P(x) = |\psi(x)|^2 = |b|^2 \exp(-2x\kappa_2)$$

Let’s see how much the probability falls when we increase x by 1 angstrom, i.e., by 1.9 bohr. Since $P(x)$ falls exponentially with x , then clearly P will fall by

$$\begin{aligned} & \frac{P(x + 1\text{Angstrom})}{P(x)} \\ &= \frac{|b|^2 \exp(-2(x + 1.9 \text{ bohr})\kappa_2)}{|b|^2 \exp(-2x\kappa_2)} \\ &= \exp(-2 * 1.9\text{bohr} * \kappa_2) \end{aligned}$$

Now if we assume that $V_2 - E \sim 4\text{eV} \sim 0.15$ Hartree (a typical number for how much the energy of the state is lower than the outside potential); also, for this purpose, use the mass of an electron, i.e., ignore the effects that other electrons have on its mass – which is OK since we consider how does the electron tunnel out in a region where there are few other electrons. Then

$$\kappa_2 = \sqrt{2M(V_2 - E)} = \sqrt{2 * m_e * 0.15 \text{Hartree}}$$

Again, in atomic units Hartree is 1 and $m_e = 1$, so

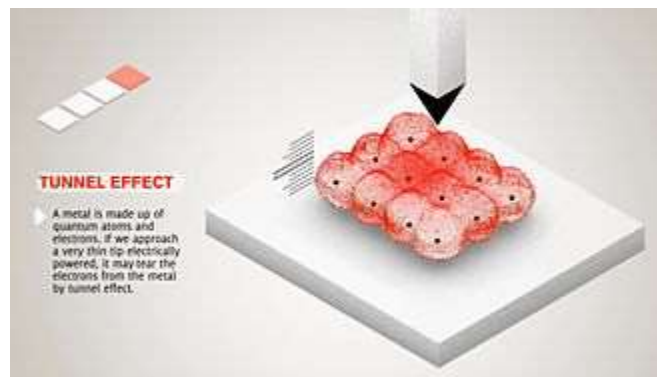
$$\kappa_2 = \sqrt{2 * 0.15} = \sqrt{0.3} \sim 0.55$$

Thus (all units fall out when we use atomic units consistently)

$$\frac{P(x + 1 \text{Angstrom})}{P(x)} \sim \exp(-2 * 1.9 * 0.55) = 0.12$$

Thus, the probability falls off steeply with distance.

This is the basis of the STM (scanning tunneling microscope): (see Wikipedia picture below):



A feedback loop is set to point a needle at a fixed distance (typically 5 angstrom or so) above a surface:

When the needle is not at the right distance, there will be too much current (if it is closer than the desired 5 angstrom) or too little (if it is too far).

The needle position is recorded exactly, which means that we know exactly the height of the surface.

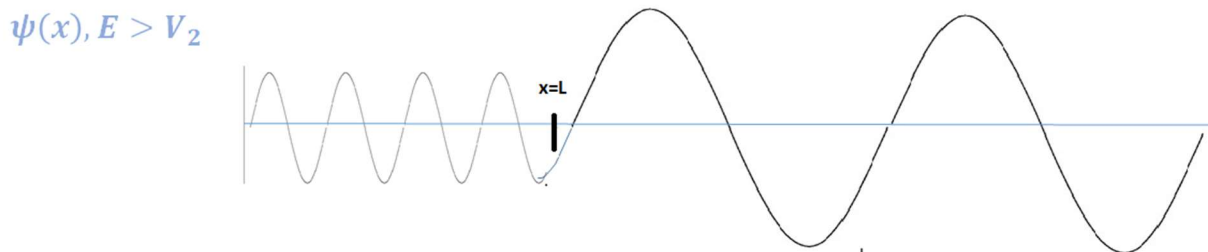
BTW, the idea for STM floated around for many years, but nobody took it seriously since nobody thought it could work. But in the 80's Gerd Binnig and Heinrich Rohrer (at IBM Zürich) tried and made it work, and the STM (and its many analogs) became the main tool for nanostudies henceforth.

Now we turn to the other case:

(ii) $E > V_2$

In that case the wavefunction on the right is $b \sin(k_2x) + c \cos(k_2x)$ where you should be clear by now what's k_2 . Now every energy is allowed (that's the blue strip in the figure, extending to all energies). The coefficients a, b, c are again related by the continuity of the wavefunction and its derivative, but we won't get into that. BTW, such case is called "scattering", and there are interesting ramifications of the existence of the well.

Below we show one example – you see that the w.f. oscillates much in the left region (where the kinetic energy is higher since the potential is deeper) and is less oscillating and of higher amplitude in the right region. The w.f. and its derivative are continuous along the border point (at $x=L$)



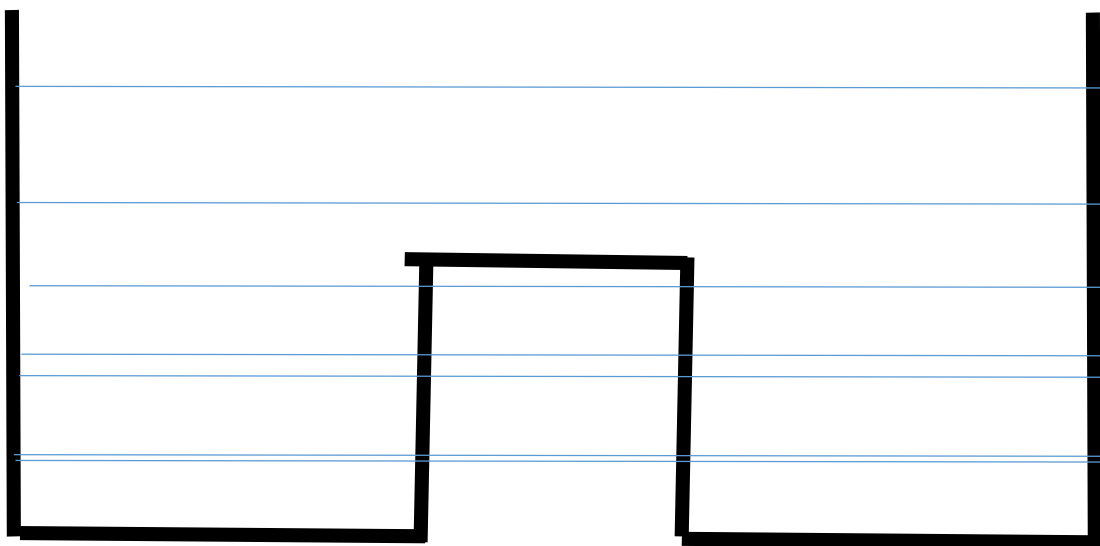
By the way, the physical reason that the w.f. has a lower amplitude in the region where it oscillates more is something called "conservation of flux". We won't go into that, but it is the same as a freeway motion – in a region where the traffic is faster the density of cars is smaller.

BTW the same thing is responsible for the bigger devastation from earthquake in sandy lands (like Santa-Monica) than in mountainous regions – the earthquake wave is much faster in the granite of the mountains than in the sandy regions, so it has a lower amplitude in the mountains (less devastation). Luckily, much of UCLA is on bedrock...

OTHER EXAMPLES:

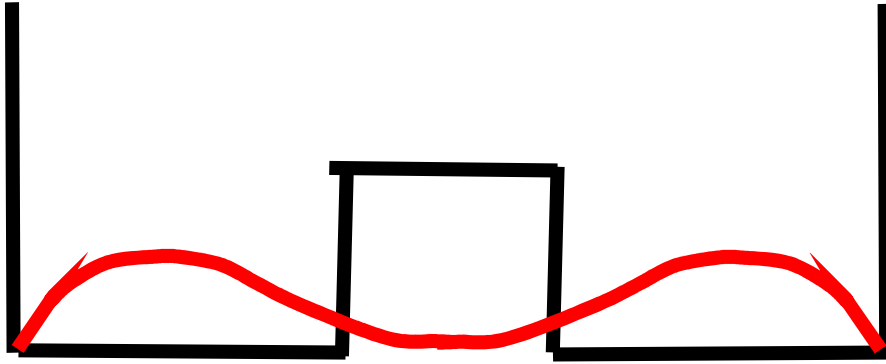
There are so many other examples and aspects of the 1D piecewise-constant potentials that we could have spent the whole course on them. I therefore just, for conclusion, quickly draw two examples without too much discussion.

Splitting in Finite Barrier,



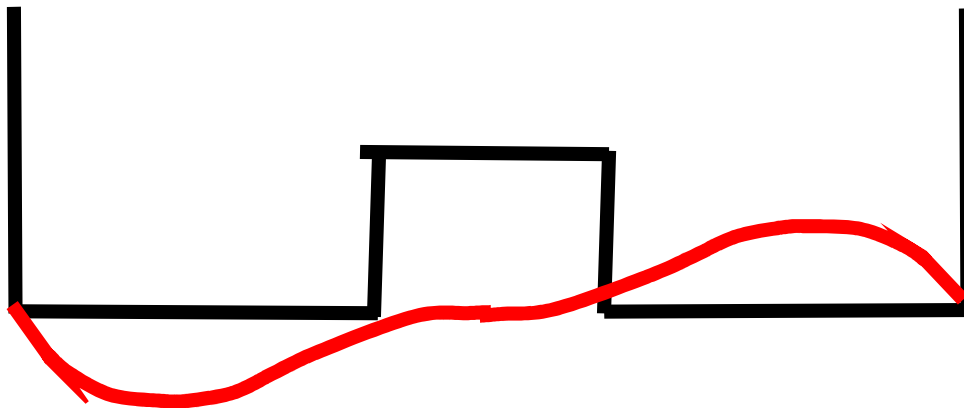
The black lines show an infinite well potential with a barrier inside. The blue lines indicate typical energy diagram.

You notice that the lowest two state in this diagram are almost degenerate; as chemists we understand that they correspond to **approximately a symmetric and antisymmetric combinations of the lowest eigenstates in each well** (see below for the lowest two eigenstates, in red):



Symmetric(lowest energy)

And: shown below, antisymmetric combination, slightly higher energy



The energy splitting can be shown to be, in the limit that the barrier is high, proportional to the tunneling matrix element, i.e.,

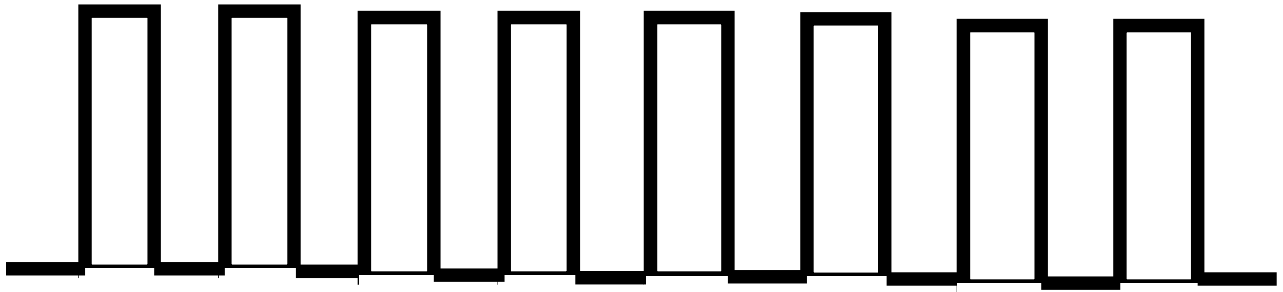
$$\exp(-\kappa A)$$

where A is the width of the barrier, $\kappa \sim \frac{1}{\hbar} \sqrt{2M(V_b - E)}$, and I introduced the barrier height, and the energy of the ground state in each well. We'll hopefully have a chance to discuss this type of **perturbation theory** (the perturbation here is that there is a finite barrier between the two wells, instead of an infinite barrier) at the end of the course.

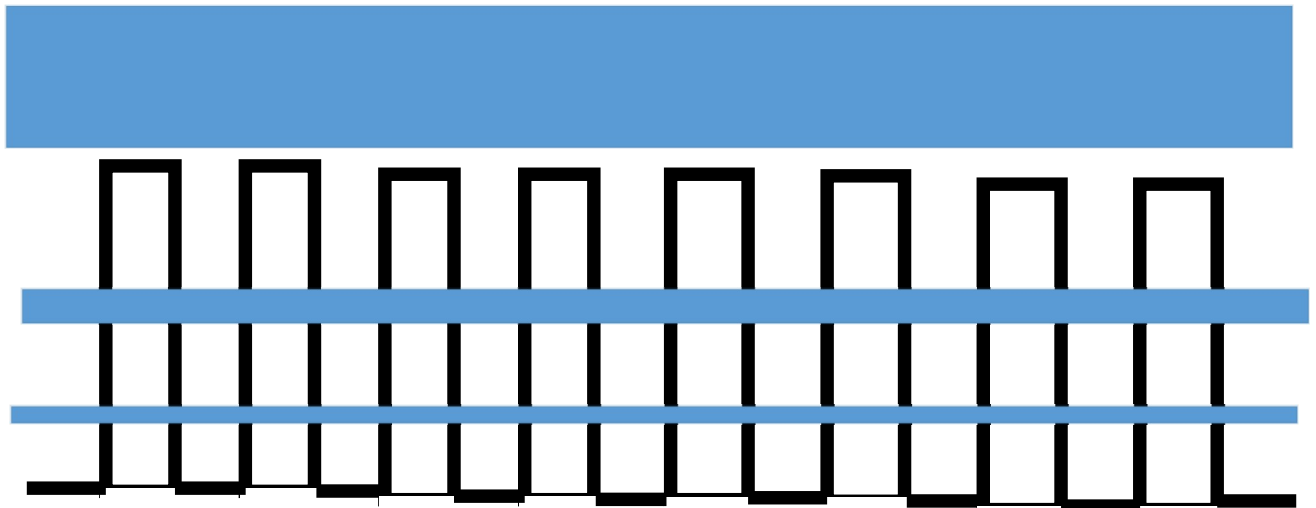
Finally:

Transmission in a lattice with periodicity:

A very important case is potential with periodic perturbations



It turns out that the splittings we saw in the case of a single barrier become now **energy bands** (allowed energy ranges are denoted in blue below):



I.e., if the coupling between the wells is not large we'll see a whole range of allowed energy which will not be too wide (its width is essentially the coupling between two adjacent wells). If you take classes on solid state chemistry or physics you'll see more of that.

Epilogue: now that we finish our too-short discussion of piecewise constant potentials, we have two directions: the Harmonic Oscillator (HO), and 3D potential (with an eye to the Hydrogen atom). We'll start with the HO, as the tools we'll get there (raising and lowering operators) will be invaluable for 3D, i.e., for the discussion of S, P, D , etc. states in 3D.

Harmonic Oscillator

Harmonic oscillator (HO) is an ideal spring where the force is proportional to the deviation from equilibrium:

$$Force = -\kappa * (x - x_0)$$

(This κ has no relation to the κ in tunneling). HO's are extremely important since they are, for both classical and quantum mechanics, the main problem that's exactly solvable in many dimensions.

I.e., a general system of coupled harmonic oscillators can be converted (if the coupling is of certain form, called linear coupling) to a set of separable harmonic oscillators, each of which can be solved independently.

Chemical dynamics, e.g. the dynamics of the motion of atoms in a molecule or in solution, is often based on describing them the vibrations as oscillators.

Here we'll therefore learn to solve a single oscillator, and then later we may if we have time briefly show how to solve in more dimensions or for more particles

Another very important feature of harmonic oscillators is that, if they are not damped (or if the damping is very weak), they will respond mostly to forces with a frequency ω_{force} which is the same as the intrinsic frequency of the oscillator,

$$\omega = \sqrt{\frac{\kappa}{M}}$$

To see that, think what happens when you push your young nephew/niece on a swing. If you apply the pushes on the spring at the "right rate", i.e., the same frequency as the "internal frequency" of the occupied-swing, then the swing will oscillate higher and higher; at other frequencies (too fast or too slow) it will barely move.

This $\omega_{forc} \sim \omega$ requirement is true in both classical mechanics and quantum mechanics. In fact, there's a lot of similarity between classical and quantum oscillators. The differences, which you learned about in 20A, are that

- There is a zero-point energy for quantum oscillators

$$E_0 = \frac{1}{2} h\nu = \frac{1}{2} \hbar\omega$$

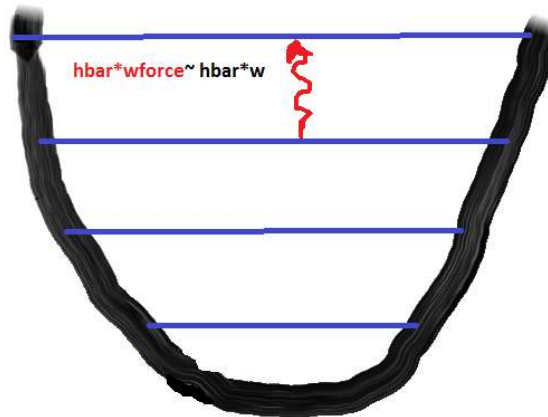
- The energies are equispaced

$$E_n = \left(n + \frac{1}{2}\right) \hbar\omega, \quad \text{i.e., } E_0 = \frac{1}{2} \hbar\omega, E_1 = \frac{3}{2} \hbar\omega, \dots$$

so

$$E_n - E_{n-1} = \hbar\omega$$

In fact, this last property makes a lot of sense; if we act on a Harmonic oscillator with light of frequency ω_{force} (we use the subscript in ω_{force} since light is really an electromagnetic wave, i.e. it applies a force) then a photon will be absorbed or emitted, i.e., the oscillator at a level "n" will respond, only if $E_n - E_{n-1} = \hbar\omega_{force}$, i.e., only if $\omega_{force} = \omega$, like we mentioned. (We'll see later that for HO, radiation usually can only cause jumps of one level up or down).



So let's solve HO's quantally.

Formally the potential energy of a spring is $\frac{1}{2} \kappa(x - x_0)^2$. Then the Hamiltonian is

$$H = \text{Kinetic} + \text{Potential} = -\frac{\hbar^2}{2M} \frac{d^2}{dx^2} + \frac{1}{2} \kappa(x - x_0)^2$$

(where I use dx and not ∂x since I only use now 1D, and we solve the time-independent Schrödinger equation, so there are no other variables such as time).

So formally we need to solve

$$H\phi_n(x) = E_n\phi_n(x)$$

i.e.,

$$-\frac{\hbar^2}{2M} \frac{d^2\phi_n(x)}{dx^2} + \frac{1}{2}\kappa x^2\phi_n(x) - E_n\phi_n(x) = 0$$

The first order of business is to “absorb” as much as possible all the constants. So, recalling that $\omega = \sqrt{\frac{\kappa}{M}}$, we get $\kappa = M\omega^2$. Inserting to H we get

$$H = -\frac{\hbar^2}{2M} \frac{d^2}{dx^2} + \frac{1}{2}M\omega^2(x - x_0)^2$$

Let's divide and multiply by $\hbar\omega$ and pulling the $\frac{1}{2}$ we get

$$H = \frac{\hbar\omega}{2} \left(-\frac{\hbar}{M\omega} \frac{d^2}{dx^2} + \frac{M\omega}{\hbar} (x - x_0)^2 \right)$$

Note that we have the factor of $\frac{M\omega}{\hbar}$ in front of $(x - x_0)^2$ and the inverse of that factor in front of the 2nd derivative. So all we need to do now is define a new dimensionless variable

$$q \equiv \sqrt{\frac{M\omega}{\hbar}} (x - x_0)$$

By the way, it is better to define now the “natural width” of the oscillator

$$\sigma \equiv \frac{1}{\sqrt{\frac{M\omega}{\hbar}}} = \sqrt{\frac{\hbar}{M\omega}}$$

So we get that q is the dimensionless length

$$q = \frac{x}{\sigma}$$

In terms of this variable, the kinetic energy term in the parentheses is simple

$$\frac{\hbar}{M\omega} \frac{d^2}{dx^2} = \frac{d^2}{dq^2}$$

So the overall Hamiltonian is

$$H = \frac{\hbar\omega}{2} \left(-\frac{d^2}{dq^2} + q^2 \right)$$

Thus, we need to solve

$$H\phi_n = \frac{\hbar\omega}{2} \left(-\frac{d^2}{dq^2} + q^2 \right) \phi_n = E_n \phi_n$$

At this stage we need to make a choice:

- We could open a differential equation book and look for the solution of this yellow highlighted equation, or worse yet, even do it ourselves, with methods of differential equations. We'll get that the solutions are something called "Hermite Polynomials". This is the usual thing to do at 113A, but is both nauseating and non-general, so you'll forget the details (like I did) immediately after the quarter is done.
- Instead, we'll use the more abstract but very elegant, powerful and general approach of Heisenberg with raising and lowering operators.

Raising and Lowering Operators for Harmonic Oscillators

The essence of the proof is simple and beautiful. We'll show two seemingly contradictory things:

- Any energy eigenvalue E (i.e., eigenvalue of the H.O. Hamiltonian) must be positive.
- Given any eigenfunction $\phi_E(x)$ with an associated eigenvalue E we'll show that we can make a new eigenfunction $\phi_{E-\hbar\omega}(x)$ with an energy that's lower by $\hbar\omega$.
This new eigenfunction will be obtained by applying a lowering operator:

$$a\phi_E \propto \phi_{E-\hbar\omega}$$

i.e., when we apply the lowering operator we get a function that is an eigenfunction of H but with an eigenvalue lower from E by $\hbar\omega$.

(Incidentally, this this lowering operator is actually very simple -- it is just proportional to $q + \frac{d}{dq}$).

Do you see the problem?

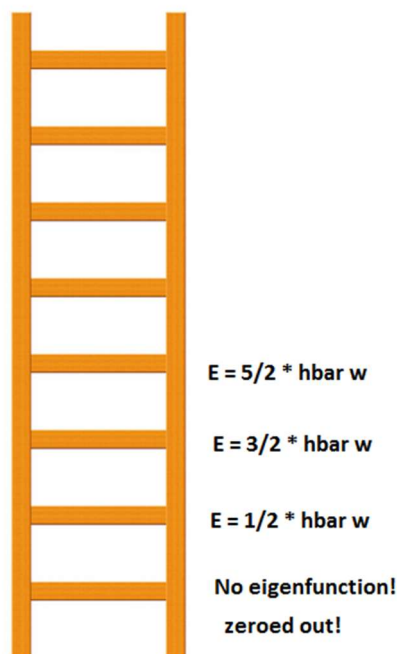
Say we had an eigenfunction with an associated eigenvalue that equals, say, $13.72 \hbar\omega$.

Then we'll apply the lowering operator once, get an eigenfunction with an energy of $13.72 \hbar\omega - \hbar\omega = 12.72 \hbar\omega$.

Then apply the lowering operator once more, get an eigenfunction with an energy of $11.72 \hbar\omega$

Etc.... apply it 11 more times and get an eigenfunction with an eigenvalue $0.72 \hbar\omega$. **And now the problem: Apply it once more, and get an eigenvalue $0.72 \hbar\omega - \hbar\omega = -0.28 \hbar\omega$...i.e., a negative eigenvalue!**

Thus the "ladder" of eigenvalues (see below) extends to negative eigenvalues, which we just said is impossible!



So what's the solution? Well it is the "proportionality" factor in $a\phi_E \propto \phi_{E-\hbar\omega}$.

I.e., when we apply the lowering operator we get an unnormalized function, which we need to normalize. **But for one value of E , i.e., $E = \frac{\hbar\omega}{2}$** it turns out that we could not normalize. I.e., when we apply the lowering operator on a function with $E = \frac{\hbar\omega}{2}$ then we get 0!

$$a\phi_{E=\frac{\hbar\omega}{2}} = 0!$$

Thus, the ladder stops (which it must!), if and only if, our energy is $E = \frac{\hbar\omega}{2}$, or of course $E = 1.5\hbar\omega$, or $E = 2.5\hbar\omega$... (see picture above).

Thus the **energy has to be of this form, i.e.,**

$$E = \frac{1}{2}\hbar\omega, \frac{3}{2}\hbar\omega, \frac{5}{2}\hbar\omega \dots$$

i.e., counting (due to historical notation) the eigenstates from $n = 0$ (not starting from $n = 1$ as we usually do), the energy must be of the form

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega, \quad n = 0, 1, 2, \dots$$

and of course the ground-state (here $n = 0$) has a "zero-point" energy of $\frac{\hbar\omega}{2}$ as you learned in 20A/20B.

That's it with the "big picture!"

Now we need to prove the two "yellow" assertions above.

- First assertion: the energy needs to be positive.

Proof:

Given a normalized $\phi_E(q)$ such that

$$H\phi_E(q) = E\phi_E(q)$$

Multiply by $\phi_E^*(q)$ and integrate

$$\int \phi_E^*(q) H\phi_E(q) dq = E \int \phi_E^*(q) \phi_E(q) dq$$

The RHS is simple, $E \int \phi_E^*(q) \phi_E(q) dq = E \int |\phi_E(q)|^2 dq = E$ (since $\phi_E(q)$ is normalized, so $\int |\phi_E(q)|^2 dq = 1$), so

$$\int \phi_E^*(q) H \phi_E(q) dq = E$$

The LHS of the gray equation is

$$\int \phi_E^*(q) H \phi_E(q) dq = \int \phi_E^*(q) \frac{\hbar\omega}{2} \left(-\frac{d^2}{dq^2} + q^2 \right) \phi_E(q) dq$$

i.e.,

$$\int \phi_E^*(q) H \phi_E(q) dq = \frac{\hbar\omega}{2} \left(-\int \phi_E^*(q) \frac{d^2}{dq^2} \phi_E(q) dq + \int \phi_E^*(q) q^2 \phi_E(q) dq \right)$$

The 2nd term on of this “green” RHS of the eq. is automatically positive

$$\int \phi_E^*(q) q^2 \phi_E(q) dq = \int |\phi_E^*(q)| q^2 dq > 0$$

The 1st term on the RHS is also positive, as we can show by integration by parts, in the same way we did earlier for the kinetic energy

$$\begin{aligned} -\int \phi_E^*(q) \frac{d^2}{dq^2} \phi_E(q) dq \\ = -[\phi_E^*(q) \frac{d}{dq} \phi_E(q)]_{q=-\infty}^{q=\infty} + \int \left(\frac{d\phi_E(q)}{dq} \right)^* \frac{d\phi_E(q)}{dq} dq \end{aligned}$$

i.e.,

$$-\int \phi_E^*(q) \frac{d^2}{dq^2} \phi_E(q) dq = 0 + \int \left| \frac{d\phi_E(q)}{dq} \right|^2 dq > 0$$

Taken together, we see that

$$\int \phi_E^*(q) H \phi_E(q) dq > 0$$

i.e., due to the “gray” equation above:

$$E > 0, \text{ Q.E.D!}$$

(II) Rewriting the Hamiltonian as (almost) a product of raising and lowering operators

Now we're ready for the ladder assertion, i.e., given ϕ_E then we'll show that $a\phi_E$ is an eigenstate with eigenvalue $E - \hbar\omega$ (and we'll show what a is, of course).

The key will be to try to write the Hamiltonian as (almost) a product instead of a sum of two terms

Actually it will be a product of two operators plus a constant, which will turn out to be the zero-point energy!

To start, look again at

$$H = \frac{\hbar\omega}{2} \left(-\frac{d^2}{dq^2} + q^2 \right)$$

Now if these were not operators but numbers, this will resemble something like

$$-\alpha^2 + \beta^2$$

where α and β are numbers. And we can write

$$-\alpha^2 + \beta^2 = (\beta - \alpha)(\beta + \alpha)$$

Let's try to write H similarly. What's

$$\left(-\frac{d}{dq} + q \right) \left(\frac{d}{dq} + q \right) = ?$$

Well, expand:

$$\left(-\frac{d}{dq} + q \right) \left(\frac{d}{dq} + q \right) = -\frac{d}{dq} \frac{d}{dq} + q^2 + q \frac{d}{dq} - \frac{d}{dq} q = -\frac{d^2}{dq^2} + q^2 + \left[q, \frac{d}{dq} \right]$$

where we used $\frac{d}{dq} \frac{d}{dq} = \frac{d^2}{dq^2}$ (derivative of derivative is 2nd derivative), and **we introduced the commutator of the two operators**

$$[A, B] \equiv AB - BA$$

which means that when acting on a function,

$$[A, B]f \equiv ABf - BAf$$

Note:

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

(proof: $[A, B] = AB - BA = -(BA - AB) = -[B, A]$).

Let's see here

$$\left[q, \frac{d}{dq} \right] = q \frac{d}{dq} - \frac{d}{dq} q$$

i.e., acting on a function

$$\left[q, \frac{d}{dq} \right] f = \left(q \frac{d}{dq} - \frac{d}{dq} q \right) f = q \frac{d}{dq} f - \frac{d}{dq} (qf)$$

i.e., since $\frac{d}{dq} (qf) = \frac{dq}{dq} f + \frac{qdf}{dq} = f + q \frac{df}{dq}$

$$\left[q, \frac{d}{dq} \right] f = q \frac{d}{dq} f - f - q \frac{df}{dq} = -qf$$

i.e., we get the **simple and extremely important result**

$$\left[q, \frac{d}{dq} \right] = -1$$

or equivalently,

$$\left[\frac{d}{dq}, q \right] = 1$$

NOTE: I am sure this is not trivial the 1st time you see it, so let's concentrate again on one strange thing. What's, **as an operator**,

$$\frac{d}{dq} q = ?$$

Well, **formally it means that when acting on a function, f , you first multiply by q and then take the derivative of the whole thing**

$$\frac{d}{dq} qf = \frac{d}{dq} (qf) = \frac{dq}{dq} f + \frac{qdf}{dq} = f + q \frac{df}{dq}$$

END OF NOTE.

Lowering and raising operators:

Going back to our expansion

$$\left(-\frac{d}{dq} + q\right)\left(\frac{d}{dq} + q\right) = -\frac{d^2}{dq^2} + q^2 + \left[q, \frac{d}{dq}\right] = -\frac{d^2}{dq^2} + q^2 - 1$$

and therefore:

$$-\frac{d^2}{dq^2} + q^2 = \left(-\frac{d}{dq} + q\right)\left(\frac{d}{dq} + q\right) + 1$$

and therefore, recalling that $H = \frac{\hbar\omega}{2}\left(-\frac{d^2}{dq^2} + q^2\right)$, we get:

$$H = \hbar\omega a^+ a + \frac{\hbar\omega}{2}$$

where we defined **the lowering operator** (a name to be justified later)

$$a = \frac{1}{\sqrt{2}}\left(\frac{d}{dq} + q\right)$$

Note that a is an operator; we usually reserve upper case letters for most operators, but for historical reason we use a lower case here.

We also wrote

$$a^+ = \frac{1}{\sqrt{2}}\left(-\frac{d}{dq} + q\right)$$

Actually, there is a reason we used the symbol a^+ . The plus-superscript stands for a **Hermitian conjugate** and we'll prove that a^+ is indeed the Hermitian conjugate of a .

But first let's define what a Hermitian conjugate (**you'll need to read the appendix first**). Formally, for any linear operator A its Hermitian conjugate, A^+ , is defined as the operator such that for any two functions, $f(x)$ and $g(x)$,

$$\int g^* A^+ f dx \equiv \left(\int f^* A g dx \right)^*$$

Of course, Hermitian operators (which we defined earlier) exactly those where $A^+ = A$, i.e., $\int g^* A f dx \equiv (\int f^* A^+ g dx)^*$. But a general operator need not be Hermitian.

Here, specifically, a is not Hermitian; let's prove that indeed the last light-blue equation is correct; indeed (and using q rather than x)

$$\begin{aligned} \int f^* a^+ g dq &\equiv \left(\int g^* a f dq \right)^* = \frac{1}{\sqrt{2}} \left(\int g^* \left(\frac{d}{dq} + q \right) f dq \right)^* \\ &= \frac{1}{\sqrt{2}} \left(\int g^* \frac{d}{dq} f dq \right)^* + \frac{1}{\sqrt{2}} \left(\int g^* q f dq \right)^* \end{aligned}$$

which, using integration by parts, gives

$$\int f^* a^+ g dq = \frac{1}{\sqrt{2}} \left(- \int \left(\frac{d}{dq} g^* \right) f dq \right)^* + \frac{1}{\sqrt{2}} \left(\int g^* q f dq \right)^*$$

i.e.,

$$\int f^* a^+ g dq = \frac{1}{\sqrt{2}} \left(\int f^* \left(- \frac{d}{dq} + q \right) g dq \right)^*$$

i.e., indeed

$$a^+ = \frac{1}{\sqrt{2}} \left(- \frac{d}{dq} + q \right)$$

as we wrote earlier.

Commutation relations

First, let's calculate the commutator of a , a^+ :

$$[a, a^+] = aa^+ - a^+a = \frac{1}{2} \left[\frac{d}{dq} + q, - \frac{d}{dq} + q \right]$$

It is easy to show that generally for any three operators

$$[A, B + C] = [A, B] + [A, C]$$

proof:

$$[A, B + C] = A(B + C) - (B + C)A = AB - BA + AC - CA = [A, B] + [A, C]$$

So applying these relations give:

$$[a, a^+] = \frac{1}{2} \left(\left[\frac{d}{dq}, -\frac{d}{dq} + q \right] + \left[q, -\frac{d}{dq} + q \right] \right)$$

but since $[A, A] = 0$ for any operator, $\left[\frac{d}{dq}, -\frac{d}{dq} + q \right] = \left[\frac{d}{dq}, q \right]$, and similarly for the 2nd term, so

$$[a, a^+] = \frac{1}{2} \left(\left[\frac{d}{dq}, q \right] + \left[q, -\frac{d}{dq} \right] \right)$$

And since , $\left[\frac{d}{dq}, q \right] = 1$, $\left[q, \frac{d}{dq} \right] = -1$ we get

$$[a, a^+] = \frac{1}{2} (1 + 1)$$

$$[a, a^+] = 1$$

i.e.,

$$aa^+ - a^+a = 1$$

Let's follow some conclusion of this commutation relation. Specifically, we'll soon need to calculate

$$aH - Ha = [a, H]$$

which is

$$[a, H] = \left[a, \hbar\omega \left(a^+a + \frac{1}{2} \right) \right] = \hbar\omega \left[a, a^+a + \frac{1}{2} \right] = \hbar\omega [a, a^+a]$$

but

$$[a, a^+a] = aa^+a - a^+aa = (aa^+ - a^+a)a = [a, a^+]a$$

i.e.,

$$[a, a^+a] = a$$

And therefore

$$[a, H] = \hbar\omega a$$

Now we can show what we looked for all along:

Lowering operators lower the energy

Let's start with a function ϕ_E fulfilling

$$H\phi_E = E\phi_E$$

Now define a function (which is not necessarily normalized)

$$f = a\phi_E$$

i.e.,

$$f(q) = \frac{1}{\sqrt{2}} \left(\frac{d}{dq} + q \right) \phi_E(q) = \frac{1}{\sqrt{2}} \frac{d\phi_E(q)}{dq} + \frac{1}{\sqrt{2}} q\phi_E(q)$$

What happens when we apply H on this function? For this we'll need a trivial property of commutators we'll use often; i.e., for any two operators

$$AB = AB - BA + BA$$

i.e.,

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

Here this means specifically, using what we learned ($[H, a] = -\hbar\omega a$) **we** get that:

$$Ha = [H, a] + aH = -\hbar\omega a + aH$$

i.e.,

$$Hf = Ha\phi_E = (-\hbar\omega a + aH)\phi_E = -\hbar\omega a\phi_E + aH\phi_E$$

i.e.,

$$Hf = -\hbar\omega a\phi_E + aE\phi_E = (E - \hbar\omega)a\phi_E$$

i.e.,

$$Hf = (E - \hbar\omega)f$$

That's what we looked for! By applying the lowering operator on an eigenfunction, we got a new eigenfunction with a lower energy! I.e., we proved that each eigenvalue gives indeed a "ladder" of lower energy eigenfunctions.

The only ingredient we need now is to prove that the ladder can stop indeed.

i.e., we need to see when will applying a give a function which is zero!

i.e., we look for a solution of

$$a\phi_0 = 0$$

Where E_0 is the yet-unknown energy of this state $\phi_0(q)$.

Amazingly, we don't even need to solve this for $\phi_0(q)$ to get E_0 (we will actually solve soon for how this wavefunction look, but interestingly we don't need the solution to just get E_0)

i.e., once we know that $a\phi_0 = 0$ we can immediately get what E_0 is:

That's because we said that ϕ_{E_0} is an eigenstate with energy E_0

$$H\phi_0 = E_0\phi_0$$

But we said that

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

So

$$\hbar\omega \left(a^+a + \frac{1}{2} \right) \phi_0 = E_0\phi_0$$

i.e.,

$$\hbar\omega a^+a\phi_0 + \frac{\hbar\omega}{2}\phi_0 = E_0\phi_0$$

But we said that E_0 is the energy such that $a\phi_0 = 0$. So the 1st term on the LHS vanishes!

i.e.,

$$0 + \frac{\hbar\omega}{2}\phi_0 = E_0\phi_0$$

i.e.,

$$E_0 = \frac{\hbar\omega}{2}$$

I.e., the energies are

$$E_0 = \frac{\hbar\omega}{2}, \quad E_1 = E_0 + \hbar\omega = \frac{3}{2}\hbar\omega, \quad E_2 = \frac{5}{2}\hbar\omega, \dots$$

i.e.,

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega \quad n = 0, 1, 2, \dots$$

Note that we did all that without even knowing what $\phi_0(q)$ is! I.e., the commutation relation gave us the energies!

Now that we have the energies we could start solving for the wavefunctions.

It turns out to be simplest to solve for ϕ_0 directly from the fact that applying a lowering operator on it gives 0.

I.e., we need to solve

$$a\phi_0 = 0$$

i.e.,

$$\left(\frac{d}{dq} + q\right)\phi_0(q) = 0$$

i.e.,

$$\frac{d\phi_0}{dq} = -q\phi_0$$

It is easy to show (even by guessing) that the solution is

$$\phi_0(q) = \text{const} * \exp\left(-\frac{q^2}{2}\right)$$

where the constant is for normalization.

Going back to our original coordinates, i.e., recalling that

$$q \equiv \frac{x - x_0}{\sigma}$$

and the natural width of the Hamiltonian is

$$\sigma = \sqrt{\frac{\hbar}{M\omega}}$$

we get that

$$\phi_0(x) = \text{const.} * \exp\left(-\frac{(x - x_0)^2}{2\sigma^2}\right)$$

i.e., an harmonic oscillator has a natural width in its ground state.

Now let's calculate the:

Higher lying states

Now let's use the raising operator, and prove that applying it on a state $\phi_n(q)$ with energy $E_n = \left(n + \frac{1}{2}\right) \hbar\omega$ gives a state $a^+ \phi_n$ with energy $E_{n+1} = \left((n + 1) + \frac{1}{2}\right) \hbar\omega$,

i.e., we'll prove that

$$\phi_{n+1}(q) = a^+ \phi_n$$

i.e., we need to prove

$$Ha^+ \phi_n = (E_n + \hbar\omega) \phi_n$$

For this, we'll prove the commutation relation of H with the raising operator, analogous to what we did for the lowering operator:

$$\begin{aligned} [H, a^+] &= \left[\hbar\omega \left(aa^+ + \frac{1}{2} \right), a^+ \right] = \hbar\omega [a^+ a, a^+] = \\ &= \hbar\omega (a^+ aa^+ - a^+ a^+ a) = \hbar\omega a^+ (aa^+ - a^+ a) = \hbar\omega a^+ [a, a^+] \end{aligned}$$

And using again

$$[a, a^+] = 1$$

We get

$$[H, a^+] = \hbar\omega a^+[a, a^+] = \hbar\omega a^+$$

This means that

$$Ha^+ = [H, a^+] + a^+H = \hbar\omega a^+ + a^+H = a^+(\hbar\omega + H)$$

So

$$Ha^+\phi_n = a^+(\hbar\omega + H)\phi_n = a^+(\hbar\omega + E_n)\phi_n$$

i.e.,

$$Ha^+\phi_n = (\hbar\omega + E_n)a^+\phi_n$$

i.e., what we wanted to show, i.e.,

$$Ha^+\phi_n = E_{n+1}a^+\phi_n$$

with

$$E_{n+1} = E_n + \hbar\omega$$

i.e., a lowers the wavefunctions through the ladder, a^+ raises them through the ladder!

Note on Normalization:

The states we get by applying repeatedly the raising operator are not normalized to unit norm.

We could actually use operator algebra, very simply, to even get the correct normalization, but we won't worry about it, and just recognize that applying the raising operator gives an unnormalized state at higher energy which we need to multiply by a constant to make into a normalized eigenstate.

First few excited states:

Let's apply the raising operator.

First, apply it once on $\phi_0(q)$

$$\phi_1(q) = \text{const.} * a^+\phi_0(q) = \text{const.} * \left(q - \frac{d}{dq}\right)\phi_0(q)$$

$$\phi_1(q) = \text{const.} * \left(q - \frac{d}{dq}\right) \exp\left(-\frac{q^2}{2}\right)$$

$$\begin{aligned}\phi_1(q) &= \text{const.} \left(q \exp\left(-\frac{q^2}{2}\right) - \left(-2 * \frac{q}{2}\right) \left(-\frac{q^2}{2}\right) \right) \\ &= \text{const.} * 2 * q \exp\left(-\frac{q^2}{2}\right)\end{aligned}$$

and absorbing the “2” in the constant finally gives

$$\phi_1(q) = \text{const.} * q \exp\left(-\frac{q^2}{2}\right)$$

Note that this function has a node – it is positive in some area of space, and negative in another. This is a general property – since $\int \phi_1 \phi_0 dq = 0$ and ϕ_0 is everywhere of one sign, the next state needs to be positive in some region of space and negative in another so the integral vanishes.

Now for the next high one, we’ll be less detailed in the math:

$$\begin{aligned}\phi_2(q) &= \text{cnst.} * \left(q - \frac{d}{dq} \right) \phi_1(q) = \text{cnst.} * \left(q - \frac{d}{dq} \right) \left(q \exp\left(-\frac{q^2}{2}\right) \right) \\ \phi_2(q) &= \text{cnst.} (2q^2 - 1) \exp\left(-\frac{q^2}{2}\right)\end{aligned}$$

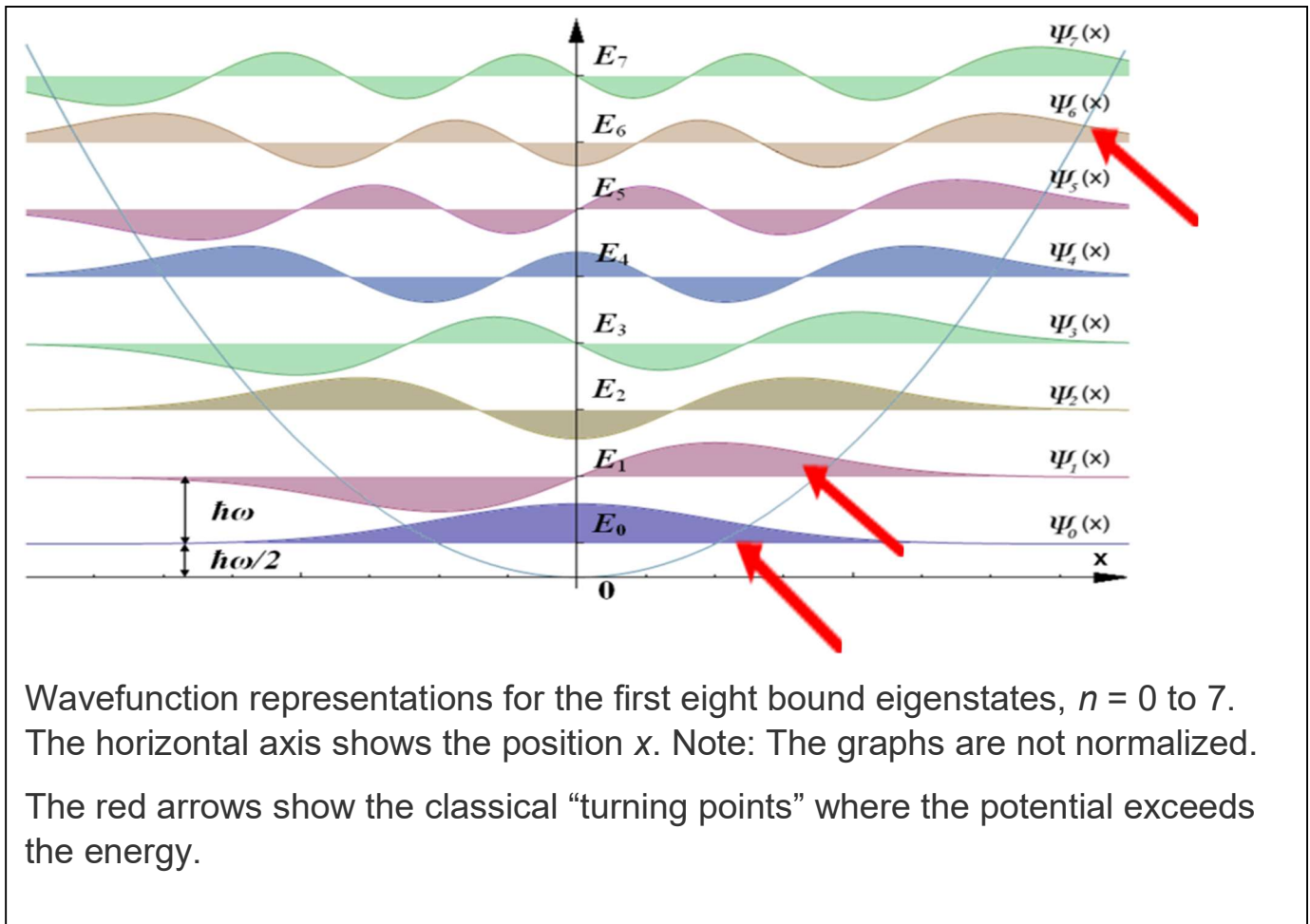
Note that this state has two nodes (at $\pm \frac{1}{\sqrt{2}}$ so it is negative for q between $-\frac{1}{\sqrt{2}}$ and $\frac{1}{\sqrt{2}}$, positive elsewhere)

And in general

$$\phi_n(q) = \text{cnst.} * \left(q - \frac{d}{dq} \right)^n \exp\left(-\frac{q^2}{2}\right)$$

Of course, we have seen these functions before – these are just the Hermite polynomials times the exponential function that we used as a simple basis (in the Appendix) before!

Graphics: The following (from Wikipedia)² shows the eigenfunctions of a perfect harmonic oscillator.



Note: the figure shows clearly that

- As we rise in energy, the wavefunctions have more nodes
- The first few eigenfunctions have significant amplitude also in values of x which are above the “classical turning points”, defined as the

² https://en.wikipedia.org/wiki/Quantum_harmonic_oscillator#/media/File:HarmOsziFunktionen.png

points where the potential $V(x)$ equals the energy E_n .
I denote classical turning points for three values by red arrows.

Note that a classical particle could not exist in areas where the potential exceeds the energy, and therefore, any non-zero value of the wavefunction outside the classical turning point (in this case, for values of x above the corresponding red arrow or below the mirror image of it at negative x) is a purely quantum effect.

- Note that as we go higher in energy most of the wavefunction is constrained to be within the “classically allowed” region, i.e., within the red arrows.

This is an indication that classical mechanics becomes relatively more valid for higher energies.

We’ll talk about that later when we discuss wavepackets for the Harmonic oscillator.

EXAMPLE: vibrations of H-Cl

Treatment based on Harmonic Oscillators:

The frequency of vibrations is in the infrared. In spectroscopic units, “wavenumbers”, it is 2990 cm^{-1}

Now the wavenumber is defined as

$$\tilde{\nu} = \frac{\nu}{c}$$

i.e., the frequency is

$$\nu = c\tilde{\nu} = 3 * 10^{10} \frac{\text{cm}}{\text{s}} * 2990 (\text{cm})^{-1} = 89.7 \text{ Thz} = 8.97 * 10^{13} \text{ Hz}$$

(don’t cry foul that I am using many digits for the wavenumber and none for the velocity of light c that’s because c is incidentally really this round number to 1 part in 10,000).

So the angular frequency is

$$\omega = 2\pi\nu = 5.63 * \frac{10^{14}}{\text{s}}$$

I like to convert everything to atomic units; in atomic units the frequency is written in terms of energy, i.e., using (Google it):

$$\hbar = 2.4188 * 10^{-17} \frac{\text{Hartree}}{\text{s}}$$

we find that

$$\hbar\omega = 2.4188 * 10^{-17} \frac{\text{Hartree}}{\text{s}} * 5.63 * \frac{10^{14}}{\text{s}} = 0.0136 \text{ Hartree}$$

i.e.,

$$\omega = 0.0136 \frac{\text{Hartree}}{\hbar}$$

Further, what's the relevant mass of the H-Cl vibration?

To a good approximation Cl does not move since it is heavy, so the mass is approximately that of H. We would be more accurate and use the reduced mass, i.e.

$$M = \frac{M_H M_{Cl}}{M_H + M_{Cl}}$$

We know that the atomic mass of H is 1.008amu, and for Cl is 35.453 AMU (well, that's wrong since it is an averaged mass over several chlorine isotopes; we really should have done the calculation of the reduced mass and vibrational frequency for each isotope separately but let's not worry about it).

So therefore

$$M = \frac{1.008 * 35.453}{1.008 + 35.453} = 0.980 \text{ amu}$$

(it is of course smaller than the mass of the lightest element in the pair, the H)

Also, it is easier if we write everything in terms of m_e , the mass of the electron; googling we find that

$$1 \text{amu} = 1822.9 m_e$$

So the reduced mass is

$$M = 0.980 * 1822.9 \simeq 1787 m_e$$

So therefore, putting it together gives:

$$\sigma = \sqrt{\frac{\hbar}{M\omega}} = \sqrt{\frac{\hbar}{1787 m_e 0.0136 \frac{\text{Hartree}}{\hbar}}} = 0.203 * \sqrt{\frac{\hbar^2}{m_e \text{Hartree}}}$$

Now in atomic units, \hbar is 1, the mass of electron is 1, and Hartree is 1, so the remaining $\sqrt{\frac{\hbar^2}{m_e \text{Hartree}}}$ term must be "1" in atomic units, i.e., $1 \text{bohr} = 1a_0$.

Therefore,

$$\sigma = 0.203 a_0$$

i.e. (since $1 \text{bohr} = 0.529 \text{Angstrom}$)

$$\sigma \simeq 0.11 \text{Angstrom}$$

Note that this is a small but non insignificant width.

The zero-point energy of this vibration is

$$E_0 = \frac{\hbar\omega}{2} = \frac{0.0136}{2} \text{Hartree} = 0.0068 \text{Hartree} = 0.185 \text{eV}$$

(we used $1 \text{Hartree} = 27.2114 \text{eV}$).

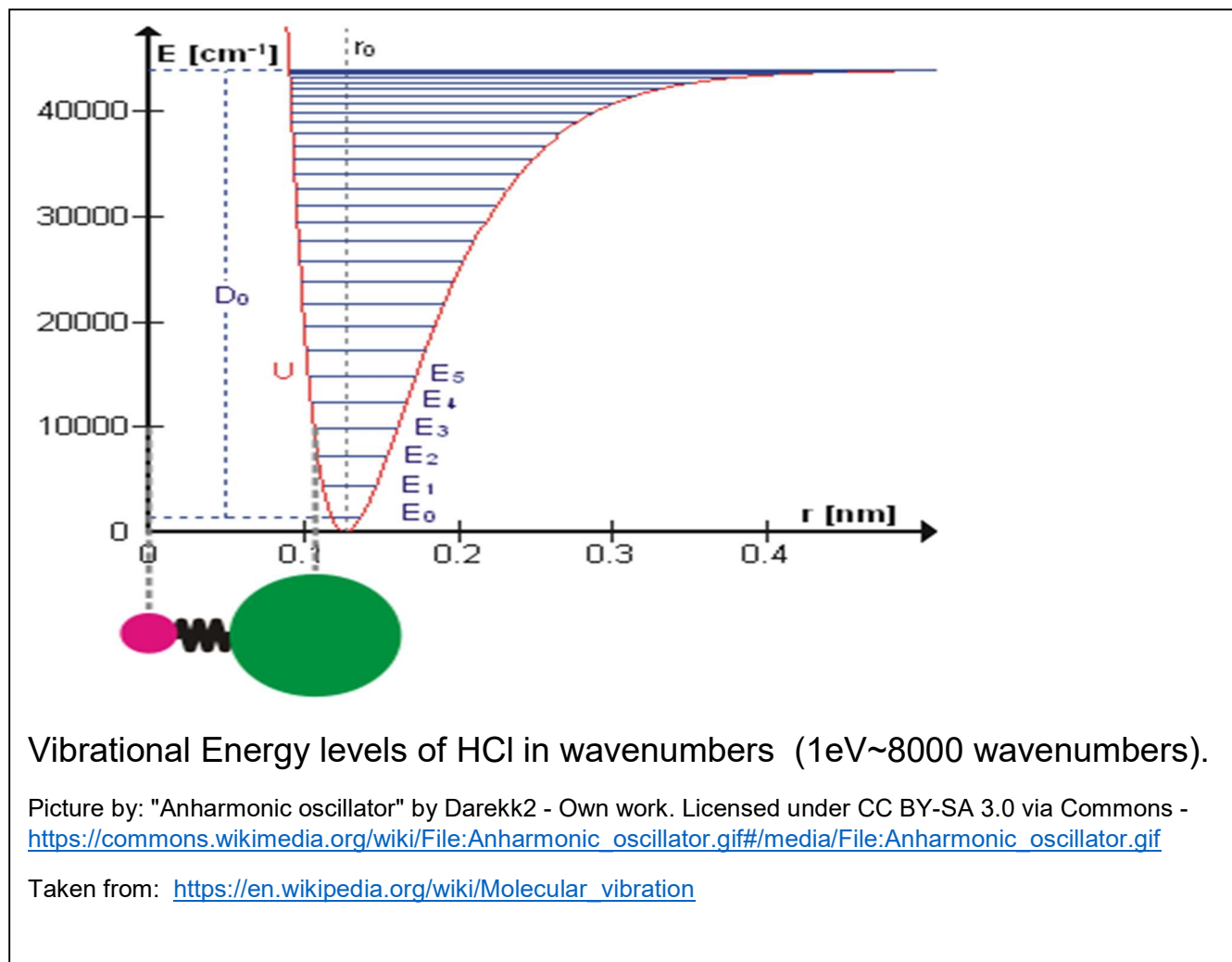
This is of course a highly non-trivial amount of energy; $k_B T_{\text{room}} = 0.026 \text{eV}$ so the zero point energy is 7 times higher than the vibrational energy.

So in accounting for chemical reactions energies one needs to account for the zero point energy.

For example if in a reaction a NaH bond is replaced by NaBr, then when we calculate the energy difference we also need to account for the different zero-point energies of the two bonds.

Note that for very heavy atoms (e.g., for NaBr vibrations) the width will be much smaller due to the heavier reduced mass.

Actual HCl molecular vibrations:



Note that the levels become denser, as the frequency is getting smaller. The first few levels can be approximated reasonably as that of an Harmonic oscillator. But then due to the fact that eventually the potential cannot rise forever when we extend the bond (i.e., at one point the bond “breaks”) the levels become denser and the eigenfunctions are not anymore those of a harmonic oscillator.

Harmonic Oscillator: Conclusions

To conclude let's note a few things:

- **In an ideal Harmonic oscillator ANY wavepacket will be periodic just like a classical particle.**

Specifically, given **any** wavepacket (which is a generic name for a wavefunction that is usually not just a single eigenstate) at $t=0$, $\Psi(r, t = 0)$, we can calculate the wavepacket after a single period, i.e., at $t = T \equiv \frac{2\pi}{\omega}$. We do it by first writing the wavepacket as a sum of the eigenfunctions,

$$\Psi(r, t = 0) = \sum_n b_n \psi_n(r),$$

where b_n are the expansion coefficients, obtained of course as usual as

$$b_n = \int \psi_n^*(r) \Psi(r, t = 0) dt.$$

Then, the wavefunction at latter times will be given, as before, by multiplying each eigenstate by a phase factor

$$\Psi(r, t) = \sum_n b_n \psi_n(r) \exp\left(-\frac{iE_n t}{\hbar}\right)$$

Prove that this is true by proving that indeed

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

So far our derivation was general, for any system.

Now specialize to the harmonic oscillator:

Take t to be a classical period, $t = T = \frac{2\pi}{\omega}$.

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

So

$$\begin{aligned}
\exp\left(-\frac{iE_n T}{\hbar}\right) &= \exp\left(-\frac{i}{\hbar}\left(n + \frac{1}{2}\right)\hbar\omega * \frac{2\pi}{\omega}\right) \\
&= \exp\left(-i\left(n + \frac{1}{2}\right) * 2\pi\right) = \exp(-i2\pi n) \exp(-i\pi)
\end{aligned}$$

And since for any n

$$\exp(-i2\pi n) = \cos(2\pi n) - i \sin(2\pi n) = 1 - i * 0 = 1$$

while

$$\exp(-i\pi) = \cos(\pi) - i \sin(\pi) = -1 - i * 0 = -1$$

We get that

$$\exp\left(-\frac{iE_n T}{\hbar}\right) = -1$$

i.e., the phase factor is the same (and is -1) for all states after a single oscillation period. Therefore

$$\begin{aligned}
\Psi(r, t = T) &= \sum_n b_n \psi_n(r) \exp\left(-\frac{iE_n T}{\hbar}\right) \\
&= \sum_n b_n \psi_n(r) * (-1) = - \sum_n b_n \psi_n(r)
\end{aligned}$$

i.e.,

$$\Psi(r, t = T) = -\Psi(r, t = 0)$$

i.e., the wavepacket is completely the same (except for an interesting overall minus factor) after one cycle, no matter what its shape (i.e., what are the b_n).

- Another strange thing is that, no matter what happens to the wavepacket and how it spreads or contracts, for H.O. the center and momentum of the wavepacket behave as if they were classical particles!

i.e., defining:

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

$$\langle p \rangle \equiv \int \Psi^*(x, t) (-i\hbar) \frac{d}{dx} \Psi(x, t) dx$$

It turns out to fulfil the classical equation of motion

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

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

Note that you proved the 1st of these “green” equations for any Hamiltonian in your week 2 H.W.; but the 2nd green equation is only valid for a H.O.

- In Harmonic oscillators there are even special wavepackets that look like at t=0 somebody shifted the ground state to be away from equilibrium. These wavepackets, called “Gaussian wavepackets”, turn out to even move (subject to the green equations above) as if their shape is “frozen”, i.e., the wavepacket center moves but its shape is unchanged.
- Quantum Harmonic Oscillators are important for radiation, for two different reasons:
 - **Electromagnetic fields are Harmonic Oscillators.** For example, a laser can be thought of being, in each mode of radiation, a wavepacket. For example, such a wavepacket, for a specific mode of radiation, may have a high amplitude to contain 1,000,000 photons, a somewhat lesser amplitude to contain 999,999 and 1,000,001 photons, and so forth, and overall it may have a non-zero amplitude to have between, say, 999,950 and 1,000,050 photons. In fact with lasers or without them it is **exceedingly difficult to prepare radiation with a very specific number of photons.** BTW, for most of our purposes as chemist we can usually ignore the quantum nature of the photons, and just approximate

the radiation as a classical function of space and time, typically of the form

$$E(\mathbf{x}, t) \sim E_0 \cos(kx - \omega t) * (\mathbf{polarization\ vector.})$$

This is true of course only if the radiation is strong; if it is weak it does not apply (i.e., see the black body radiation problem) but if it is weak we don't need to worry about it anyway usually, since then it will not affect the molecules we deal with.

- Harmonic Oscillators are also important if we want to understand the interaction of radiation with molecules. That's because, to first approximation, **the vibrations in molecules can be described as Harmonic Oscillators** (see the H-Cl example above).

We will see later that for an electromagnetic field that fulfills the following two conditions (observed in 99.9% of the experiments you'll deal with)

- Is not extremely strong and
- is not part of a specialized experiment on tiny scales where we localize the radiation to nanometer size probes (is not a "near field")

then, when such a field impinges on a molecule, then usually it **will only cause excitations and deexcitations by 1 level**, i.e., if the molecule is in the 4th vibration in one of the modes, the field will typically cause strong excitation (or de-excitation) of the molecule to the 3rd and 5th levels.

Hopefully we'll cover this in the last week or two of classes.

Hydrogen Atom – and begin Angular Momentum

Now finally we'll get the quantum mechanics equivalent of the Bohr model.

Instead of x, y, z let's use spherical coordinates,

$$r, \theta, \phi$$

where r is the length of the vector r , θ is **the angle between r and the z-axis**, and ϕ is the angle that the **projection of r to the x-y plane makes with the x-axis**, i.e.,

$$r = \sqrt{x^2 + y^2 + z^2}$$

$$\cos(\theta) = \frac{z}{r}$$

$$\operatorname{tg}(\phi) = \frac{y}{x}$$

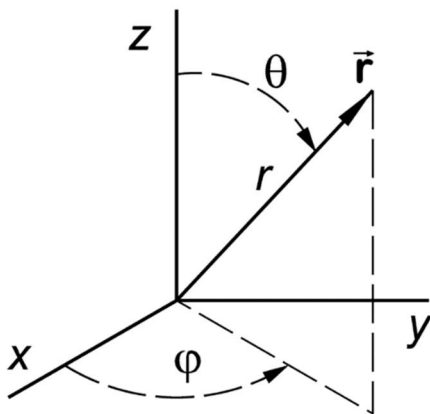
(so

$$x = r \sin(\theta) \cos(\phi)$$

$$y = r \sin(\theta) \sin(\phi)$$

$$z = r \cos(\theta)$$

See picture:³



SIDE NOTE: Spherical coordinates Volume element:

³ http://en.citizendium.org/wiki/File:Spherical_polar.png Copyright © [[User:Paul Wormer|Paul Wormer]].

In Cartesian 3D coordinates an overlap integral of two functions reads

$$\langle f|g\rangle = \int f^*(x, y, z)g(x, y, z)dxdydz$$

In spherical coordinates the volume element is different; it turns out to be

$$dvolume = r^2 dr \sin(\theta) d\theta d\phi$$

I.e., a 3D integral reads

$$\langle f|g\rangle = \int f^*(r, \theta, \phi)g(r, \theta, \phi)r^2 dr \sin(\theta) d\theta d\phi$$

We'll need to use that later.

END OF SIDE NOTE.

So Let's start classically:

Classical treatment.

Now classically the energy is (where H is the Hamiltonian, as usual)

$$E = H = K - \frac{e^2}{r}$$

where K is the kinetic energy of the electron (we assume the proton is infinitely massive, for simplicity). In Cartesian coordinates we wrote

$$K = \frac{\mathbf{p}^2}{2M} = \frac{p_x^2 + p_y^2 + p_z^2}{2M}$$

It turns out we can write the classical kinetic energy very simply in Cartesian coordinates, by separating the momentum vector to a part that is along the coordinate (a parallel part, p_r) and a perpendicular part (p_{\perp}).

I.e.,

$$\mathbf{p}^2 = p_r^2 + p_{\perp}^2$$

The perpendicular part is related to the angular momentum. Recall that the angular momentum is

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}$$

i.e., its squared norm is the value of r times the part of \mathbf{p} that's perpendicular to r , i.e.,

$$L^2 = r^2 p_{\perp}^2$$

Therefore

$$p_{\perp}^2 = \frac{L^2}{r^2}$$

So finally

$$K = \frac{p_r^2}{2M} + \frac{L^2}{2M r^2}$$

i.e.,

$$E = \frac{p_r^2}{2M} + \frac{L^2}{2M r^2} - \frac{e^2}{r}$$

This is important since classically we know that L is conserved, so in a sense we can convert the classical 3D problem of moving under a central potential $-\frac{e^2}{r}$ to a 1-D problem, where the kinetic energy is 1-dimensional $\frac{p_r^2}{2M}$, and instead of the true potential the 1-d distance “feels” an “effective potential” which is $\frac{L^2}{2M r^2} - \frac{e^2}{r}$, i.e., let's use colors:

$$E = \frac{p_r^2}{2M} + \frac{L^2}{2M r^2} - \frac{e^2}{r}$$

Very analogous properties follow for QM.

Quantum electron-in-hydrogen Hamiltonian and Angular Momentum

Formally we know

$$H = -\frac{\hbar^2}{2M} \Delta + V(r)$$

where

$$V(r) = -\frac{e^2}{r}$$

In Cartesian coordinates, of course, acting with the on a function ψ means

$$\Delta\psi = \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}$$

and it turns out that in 3D

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

Luckily we can make some sense of this mess using the **angular momentum operator**.

We'll discuss the angular momentum operator in more detail later, but for now accept that in spherical coordinates it has exactly the form that the non-r parts in the 3D spherical-coordinates Laplacian in 3D have, i.e.,

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

(we won't prove this... let's accept it... but believe me, it could be proven from the definition of the Cartesian components of L and from converting from Cartesian components to spherical components).

So therefore, we can write that the action of the kinetic energy on a general function f is

$$K\psi \equiv -\frac{\hbar^2}{2M} \Delta\psi$$

i.e.,

$$K\psi = -\frac{\hbar^2}{2M} \frac{1}{r} \frac{\partial^2}{\partial r^2} (r\psi) + \frac{1}{2Mr^2} L^2\psi$$

(Note that the factor of \hbar^2 and the minus sign were both absorbed into the definition of L^2).

Therefore, the Schrodinger equation for the three dimensional molecular orbitals, $\psi(r, \theta, \phi)$ becomes

$$K\psi + V\psi = E\psi$$

i.e.,

$$-\frac{\hbar^2}{2M} \frac{1}{r} \frac{\partial^2}{\partial r^2} (r\psi) + \frac{1}{2Mr^2} L^2 \psi + V(r)\psi = E\psi$$

Note how similar this is to the classical Hamiltonian expression we derived earlier, $\frac{p_r^2}{2M} + \frac{L^2}{2Mr^2} - \frac{e^2}{r} = E$

As usual when handling multidimensional equation, we postulate (i.e., hope) that we can find solutions which are separable, i.e., are product of a function of r , times another function of θ and ϕ , i.e., we write

$$\psi(r, \theta, \phi) = R(r)Y(\theta, \phi)$$

Later we'll add of course indices to distinguish different for R eigenfunctions and different Y eigenfunctions.

So let's insert this choice and hope that we get a solution. Of course a general wavefunction solving the Schrodinger equation does not need to be of this separable product form, but it will be the sum of separable terms (just like the Schrodinger wavefunction in x,t will be the sum of terms of the form $\exp\left(-\frac{iE_n t}{\hbar}\right) \phi_n(x)$)

So inserting the product form

$$-Y \frac{\hbar^2}{2M} \frac{1}{r} \frac{\partial^2}{\partial r^2} (rR) + \frac{R}{2Mr^2} L^2 Y + (V(r) - E)RY = 0$$

(since $Y(\theta, \phi)$ is not a function of r we pulled it to the left in the 1st term. Similarly, since L^2 is written purely in terms of derivatives of the angles, $L^2 RY = R L^2 Y$)

Now without doing too much math, you see that this equation will be possible **if**

$$L^2 Y = \lambda Y$$

i.e., if Y is an eigenstate of L^2 . In that case the blue equation becomes

$$-Y \frac{\hbar^2}{2M} \frac{1}{r} \frac{\partial^2}{\partial r^2} (rR) + \frac{R}{2Mr^2} \lambda Y + (V(r) - E)RY = 0$$

Now there are no more derivatives w.r.t. the angular coordinates, so we can divide Y and get

$$-\frac{\hbar^2}{2M} \frac{1}{r} \frac{\partial^2}{\partial r^2} (rR) + \frac{R}{2Mr^2} \lambda + (V(r) - E)R = 0$$

Interlude – Eigenvalues of L^2

We are looking for, as mentioned a solution of an eigenvalue equation for L^2

$$L^2 Y(\theta, \phi) = \lambda Y(\theta, \phi)$$

We'll talk about the formal and general solution of that equation later; suffice it is to say that the eigenvalues of L^2 have the form

$$\lambda = \hbar^2 l(l+1), \quad l = 0, 1, 2, \dots$$

i.e., (since $1 \cdot 2 = 2$, $2 \cdot 3 = 6$, $3 \cdot 4 = 12$, etc.)

$$\lambda = 0, 2\hbar^2, 6\hbar^2, 12\hbar^2, \dots$$

We note in passing that this result is similar but not equal to the Bohr model assumption

$$\text{Bohr: } L = l\hbar \text{ (wrong!)}$$

More later on that. For now, back to:

The radial equation

Multiply by r and plug in the allowed values of λ to get (I insert also the explicit form of $V(r)$)

$$-\frac{\hbar^2}{2M} \frac{\partial^2}{\partial r^2} (rR_{nl}) + \left(-\frac{e^2}{r} + \frac{\hbar^2 l(l+1)}{2Mr^2} - E \right) rR_{nl} = 0$$

You'll notice that I labeled the radial function by two indices, n, l . The latter is the index associated with the eigenvalue of the angular momentum, since for each l we'll have a different radial equation.

The n index is related to the number of nodes in the radial eigenfunctions. More precisely,

$$n = 1 + l + \text{number of nodes in radial wavefunction}$$

The reason we use this strange notation, and not label R indirectly by the number of radial nodes, is that it turns out that the energy depends only on n , not on l , as we'll explain.

Put differently, the number of nodes in the angular part is l , so we can write

$$n = 1 + \text{total number of nodes (radial + angular)}$$

For example, a 4p state will have

$$l = 1$$

And

$$n = 4$$

i.e., 2 radial nodes, 1 angular nodes, so totally 3 nodes, so $n = 1 + 3 = 4$.

You know this solution (and a lot of info on the radial and angular wavefunction) already from 20A, let's discuss a little more in depth the solution of the radial equation.

Let's accept (it will take a few lectures to prove otherwise) what you learned in 20A on the solution of the eigenvalues of the radial equation, i.e., that they are

$$E_{nl} = E_n = -\frac{e^2}{2a_0} * \frac{1}{n^2}, \quad n = 1, 2, \dots$$

i.e., depend only on the total number of nodes, not whether they are angular or radial. remind you that $a_0 = \text{bohr} = \frac{\hbar^2}{Me^2} = 0.529 \text{Angstrom}$ so

$$\frac{e^2}{2a_0} = \frac{1}{2} \text{Hartree} = 13.6 \text{eV}$$

i.e.,

$$E_n = -13.6 \text{ eV} * (1, \frac{1}{4}, \frac{1}{9}, \frac{1}{16}, \dots)$$

Now that we were told the eigenvalues, let's get some info. on the eigenfunctions.

Let's rewrite the yellow equation above by inserting the energies; and

Then

$$-\frac{1}{r} \frac{\hbar^2}{2M} \frac{\partial^2}{\partial r^2} (r R_{nl}(r)) + \left(-\frac{e^2}{r} + \frac{\hbar^2 l(l+1)}{2Mr^2} + \frac{e^2}{2a_0 n^2} \right) R_{nl}(r) = 0$$

Note that by specifying the angular momentum we converted the Schrödinger equation to be a 1-D like equation in the radial coordinate --- in a sense analogous to what we had in the classical case, where we ended up with a single equation in the radial coordinates.

We'll first

Present the formal solutions for the for a few cases (low n's)
 Explain the reasons the solutions look the way they do.

Lowest radial Solutions $R_{nl}(r)$

s ($l = 0$):

$$R_{10}(r) = \text{const.} * \exp\left(-\frac{r}{a_0}\right)$$

$$R_{20}(r) = \text{const.} * \left(2 - \frac{r}{a_0}\right) \exp\left(-\frac{r}{2a_0}\right)$$

$$R_{30}(r) = \text{const.} * \left(3 - \frac{2r}{a_0} + \frac{2r^2}{9}\right) \exp\left(-\frac{r}{3a_0}\right)$$

Let's first prove, for the simplest of these functions, $R_{10}(r)$, that it indeed solves the radial Schroedinger equation. So without the $l(l+1)$ term since $l = 0$, and using $n = 1$, we need to prove that

$$-\frac{1}{r} \frac{\hbar^2}{2M} \frac{\partial^2}{\partial r^2} \left(r \exp\left(-\frac{r}{a_0}\right) \right) + \left(-\frac{e^2}{r} + \frac{e^2}{2a_0} \right) \exp\left(-\frac{r}{a_0}\right) = ??? 0$$

To simplify we'll use atomic units, where $M = \hbar = a_0 = e^2$. So we need to prove that

$$-\frac{1}{2r} \frac{\partial^2}{\partial r^2} (r \exp(-r)) + \left(-\frac{1}{r} + \frac{1}{2} \right) \exp(-r) = ??? 0$$

Multiply by $-2r$ to get

$$\frac{\partial^2}{\partial r^2} (r \exp(-r)) + (-2 + r) \exp(-r) = ??? 0$$

But since

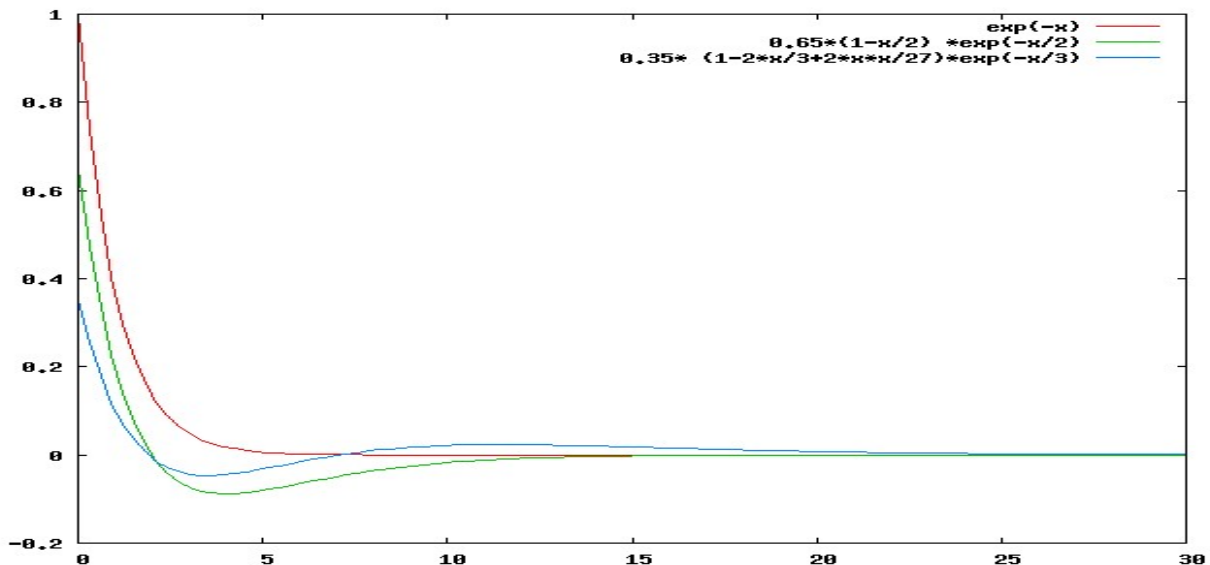
$$\begin{aligned} \frac{\partial^2}{\partial r^2} (r \exp(-r)) &= \frac{\partial}{\partial r} \frac{\partial}{\partial r} (r \exp(-r)) = \frac{\partial}{\partial r} ((1-r) \exp(-r)) \\ &= (2-r) \exp(-r) \end{aligned}$$

then indeed the “gray” equation is correct, and the form we had for $R_{10}(r)$ is correct.

Q.E.D.

Let’s plot: (red: 1s, green: 2s; blue: 3s

Note that the plots of these s-functions and the p- and d- functions soon are not of the normalized functions, as otherwise it would have been too hard to view by eye.



Let’s see some properties

Long distance: Let’s start by considering the asymptotic behavior of $R_{nl}(r)$ as $r \rightarrow \infty$. For large r , the potential and angular momentum term vanish (as they are proportional to $\frac{1}{r}$ and $\frac{1}{r^2}$, so at large r **the difference between the potential ($V(r \rightarrow \infty) = 0$) and the energy ($E = -\frac{1}{2n^2}$ in atomic units) becomes, in atomic units:**

$$E - V = -\frac{1}{2n^2} - 0 = -\frac{1}{2n^2}$$

Because the difference is **negative** (at far distance the energy is below the potential, as is always true for bound states), the wavefunction is damped, and the damping constant κ

$$\kappa = \frac{1}{\hbar} \sqrt{2M(V - E)} = (\text{in a. u., } \hbar = M = 1) = \sqrt{2(V - E)} = \sqrt{2 * \frac{1}{2n^2}}$$

i.e., at far distances

$$\kappa = \frac{1}{n}$$

i.e., at far distances the wavefunctions decay more or less as:

$$R_{n0}(r) \sim \exp\left(-\frac{r}{n}\right)$$

which is what you see in the blue/turquoise list above (there we had $\exp\left(-\frac{r}{a_0}\right)$, $\exp\left(-\frac{r}{2a_0}\right)$, $\exp\left(-\frac{r}{3a_0}\right)$, etc., and it is the same since in atomic units 1bohr = 1).

The point is that as you go to higher n , the energy gets higher so the bound-state wavefunction damps slower.

i.e., higher- n wavefunctions are more extended!

Furthermore, it is not surprising when we think on the slower and slower long-distance exponential fall off $R_{n0}(r)$ with higher n , that the average size (not just the fall off) of the wavefunction follows exactly this behavior, i.e.,

$$\langle r \rangle = n a_0$$

This is exactly like the $r = n a_0$ prediction of the Bohr model!!

Warning – this is for an isolated hydrogen atom; in a molecule and especially in liquids and solids the hydrogen wavefunction cannot extend in the same way significantly in direction of other atoms – there we need to take linear combination of atomic orbitals from different atoms to get MO, as you know and we'll learn about.

Small radius and general solutions for $l = 0$ (s states).

The “s” ($l = 0$) states extend all the way into the nucleus.

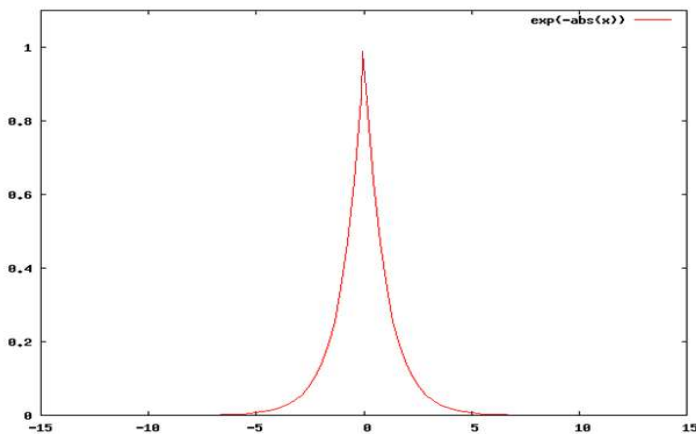
You may wonder what happens near the nucleus since the potential is $-\frac{e^2}{r}$, i.e., it becomes negative more and more and is formally infinitely negative at $r=0$.

So what balances out the infinitely negative potential energy, i.e., **what prevents the wavefunction from collapsing to the nucleus?**

The answer is: the **kinetic energy**

You may think that the solutions we have above look smooth so how come it has kinetic energy which becomes infinitely positively large at $r = 0$ (so it can cancel the infinitely negative charged potential?)

To see, let's do a cut of our radial function at $y = z = 0$, but at **all x** , i.e., at **both and positive and negative x** :



You see that the 3D wavefunction is not smooth at $x=0$, i.e., at $r=0$ in this case (since $y=z=0$, so $x=0$ is the origin).

Its non-smoothness translates to infinitely positive kinetic energy (minus of the second derivative), which cancels the infinitely negative potential energy at the origins.

As you go to higher quantum numbers, the radial eigenfunctions extend and have more and more nodes:

Look at the three $l = 0$ eigenstates.

The first has ($R_{10}(r)$) no radial nodes.

The next ($R_{20}(r)$) has one node.

The third one ($R_{30}(r)$) has two nodes.

This is a general property. It follows since the eigenstates have to be **orthogonal within each l**:

$$\int R_{jl}(r)R_{nl}(r) r^2 dr = \delta_{jn}$$

(the r^2 comes from the volume element, $dV = r^2 dr \sin(\theta) d\theta$).

For example, the 1s function $R_{10}(r)$, is all of one sign (positive, although that does not matter, it could have been all negative, it is just a convention).

So for the 2s function $R_{20}(r)$ to be orthogonal to $R_{10}(r)$ it needs to have positive and negative regions, i.e., it needs to have a node.

Note that the orthogonality only applies to the same l. If the l 's are different then the **angular** overlap (we'll see that later) will be zero, so then the radial overlap does not need to vanish.

For example, the 2p, 3p, 4p, etc. radial functions will not be orthogonal to any of the s functions

$$\int R_{j0}(r)R_{n1}(r)r^2 dr \neq 0$$

Higher “p, d, etc.” ($l > 0$) radial eigenstates:

The next “p” ($l = 1$), “d” ($l = 2$) eigenstates, etc. $R_{n1}(r)$, $R_{n2}(r)$, etc. are different in one major category – they do not extend to $r = 0$.

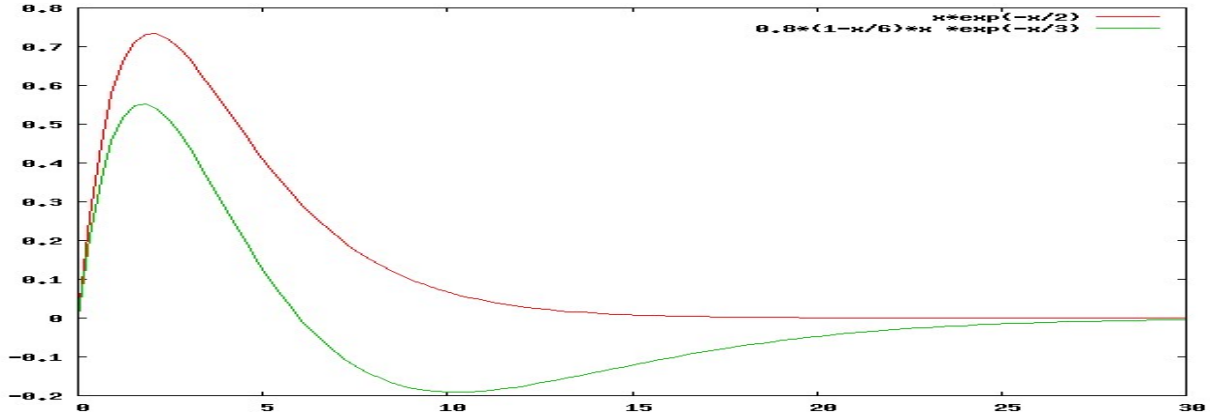
Let's write a few of them down, plot, and then understand why they don't extend to the nucleus.

First, the two lowest radial **p states**:

$$R_{n=2,l=1}(r) = R_{21}(r) = \text{const} * r * \exp\left(-\frac{r}{2a_0}\right)$$

$$R_{n=3,l=1}(r) = R_{31}(r) = \text{const} * \left(1 - \frac{r}{6a_0}\right) * r * \exp\left(-\frac{r}{3a_0}\right)$$

And their plot

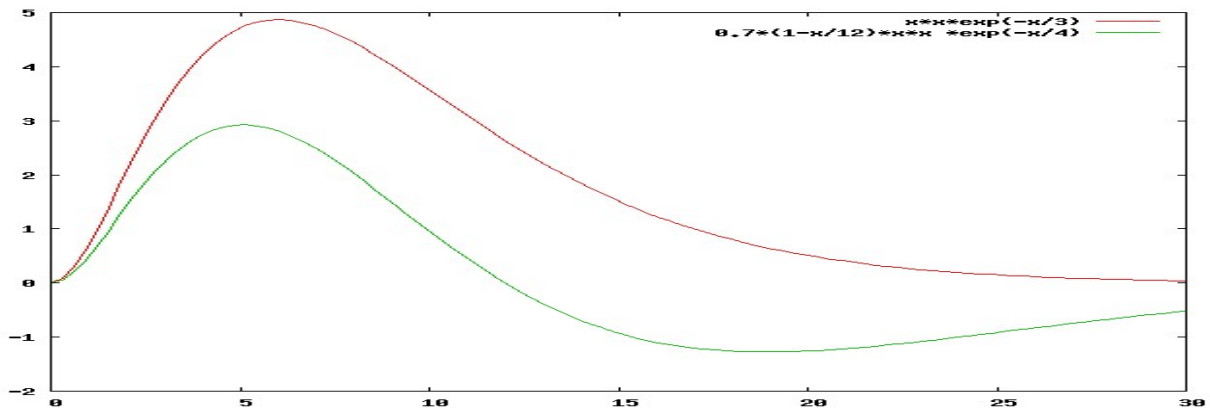


d-states:

Then, the two lowest radial d states

$$R_{n=3,l=2}(r) = R_{32}(r) = \text{const} * r^2 * \exp\left(-\frac{r}{3a_0}\right)$$

$$R_{n=4,l=2}(r) = R_{42}(r) = \text{const} * \left(1 - \frac{r}{12a_0}\right) * r^2 * \exp\left(-\frac{r}{4a_0}\right)$$



We see a pattern from the equations and from comparing all three plots:

With each higher l , we get an r^l factor.

That factor forces them to extend further and further from the nucleus (i.e., for $l = 1$ $r^l = r$ is linear in r ; $l = 2$, $r^l = r^2$ is smaller near the nucleus, etc.)

We'll use that r^l factor later to simplify the angular functions.

At very far-away distances they also damp exponentially, for the same reason that the s functions damp exponentially – since the energy is negative and the potential far away is zero. Of course, the higher the n the slower they damp (since the higher the energy is, i.e., the closer to zero).

Angular Momentum and Angular eigenfunctions for the Hydrogen atom (and in general).

Now it is time to deal with the angular part, i.e., with angular Momentum.

To recall, we said that the kinetic energy of the electron in a central potential has an angular part (times $\frac{1}{2Mr^2}$) which we highlight below

$$-\frac{\hbar^2}{2M}\Delta\psi = \frac{\hbar^2}{2M}\frac{1}{r}\frac{\partial^2}{\partial r^2}(r\psi) + \frac{1}{2Mr^2}(-\hbar^2)\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2}\right)$$

And we also said that it can be proved that this green part of the kinetic energy can be written as L^2

$$-\frac{\hbar^2}{2M}\Delta\psi = -\frac{\hbar^2}{2M}\frac{1}{r}\frac{\partial^2}{\partial r^2}(r\psi) + \frac{1}{2Mr^2}L^2\psi$$

i.e., given a function $Y(\theta, \phi)$,

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

We further said that it can be proved that the eigenstates of L^2 fulfil

$$L^2Y = \hbar^2 l(l+1)Y \quad l = 0, 1, 2, \dots$$

Now we'll expand on this.

First, recall that in classical mechanics

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}$$

This definition turns out to be true also quantum mechanically!

i.e.,

$$\mathbf{L} = \mathbf{r} \times (-i\hbar\nabla) = -i\hbar \mathbf{r} \times \nabla$$

i.e.,

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

and similarly for L_x, L_y (the trick to remember the indices when changing from the z to the y to the x components is to circle around the list of components of L, r, p that appear, ie., $zxy \rightarrow xyz \rightarrow yzx$):

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

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

and of course we define the squared angular momentum operator:

$$\mathbf{L}^2 \equiv L_x^2 + L_y^2 + L_z^2$$

Also, \mathbf{L} designates henceforth the angular-momentum operator, not a general operator).

The usual thing at this stage is to just solve for the green highlighted equation above. The solution is a little nauseating, and worse, this way of solving hides the very general properties of the eigenstates of the momentum operator which comes about **directly from the relations between different components (the commutation relation)**.

I will take a middle road:

I'll half-prove half-tell you what the regular solutions are,

And then later

I'll mention a few facts about the general operator-relations that the angular momentum operator fulfills so you can see the generality of the properties of the eigenvalues.

So let's start.

As mentioned, with some painful replacement of Cartesian with spherical coordinates could prove that the

Two (yellow and green) definitions above of \mathbf{L}^2 are equal – the Cartesian and spherical coordinates definitions are equal. Further, we could prove that in spherical coordinates L_z has a particularly simple form:

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

Eigenfunctions and eigenvalues of L^2

The derivatives w.r.t. ϕ and θ in the definition of L^2 are separated (even though the 2nd derivative w.r.t. ϕ is multiplied by $\frac{1}{\sin^2 \theta}$) so this hints that we should again try to separate the coordinate.

Further, the only term associated with ϕ is a 2nd derivative w.r.t. it, $\frac{\partial^2}{\partial \phi^2}$ and we have seen that the eigenfunctions of such terms are of the form $\exp(im\phi)$

i.e., we write the solution as

$$Y(\theta, \phi) = g(\theta) * \exp(im\phi)$$

where we need to determine what $g(\theta)$ is.

Now comes the 1st key point: **m must be integer**. This is so that if we rotate by 360 degrees the wavefunction comes back to its same value! (i.e., $\exp(im(\phi + 2\pi)) = \exp(im\phi)$).

(Actually, it is a weak argument, which for spin will not be valid – we are really never observing the wavefunction so we have to be careful on assigning to it as “single value”... we’ll explain the true argument later).

Obviously, $\exp(-im\phi)$ is also an eigenstate of L_z , i.e.,

$$L_z \exp(im\phi) = -i\hbar \frac{\partial}{\partial \phi} \exp(im\phi)$$

i.e.,

$$L_z \exp(im\phi) = m\hbar \exp(im\phi)$$

so the eigenvalues of L_z are of the form $m\hbar$ with m integer, i.e., $0, \pm\hbar, \pm 2\hbar$, etc.

Inserting these equations to the green equation above we get that the eigenvalue equation

$$L^2(g(\theta) \exp(im\phi)) = \lambda \exp(im\phi)$$

becomes

$$\frac{-\hbar^2}{\sin \theta} \left(\sin \theta \frac{\partial g(\theta) \exp(im\phi)}{\partial \theta} \right) - \frac{\hbar^2}{\sin^2 \theta} \frac{\partial^2 g(\theta) \exp(im\phi)}{\partial \phi^2} = \lambda g(\theta) \exp(im\phi)$$

i.e., since $\frac{\partial^2 g(\theta) \exp(im\phi)}{\partial \phi^2} = -m^2 g(\theta) \exp(im\phi)$, we get that we're looking for solutions of

$$\frac{-\hbar^2}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial g(\theta)}{\partial \theta} \right) - \frac{\hbar^2}{\sin^2 \theta} m^2 g(\theta) = \lambda g(\theta)$$

I'll spare you the details of solving the equations and just state the results (many of which you essentially know already):

The eigenvalues λ are indeed of the form

$$\lambda = \hbar^2 l(l+1), \quad l = 0, 1, 2, \dots$$

The eigenstates depend on the eigenvalue of L^2 and on m , the eigenvalue of L_z ; so we label them as $g_{lm}(\theta)$

There's a restriction on l

$$|m| \leq l, \quad \text{i.e., } m = -l, -l+1, \dots, l$$

That restriction clearly could be understood based on the fact that

$$L^2 - L_z^2 = L_x^2 + L_y^2$$

We'll show soon (quite easily) that the expectation value and eigenvalues of L_x^2, L_y^2 are positive, so that means that the eigenvalue of the LHS must be all positive, and that will imply that $l \geq |m|$

And most interestingly, while the eigenstates depend on m (which determines the direction) the eigenvalues don't depend on it!

So

for $l = 1$ there's a three-fold degeneracy between:

$$m = -1, m = 0, m = 1$$

A fivefold degeneracy for $l = 2$, etc.

The lowest eigenstates are:

"s" state : $l = 0$: a constant:

$$Y_{00}(\theta, \phi) \equiv Y_{m=0, l=0}(\theta, \phi) = \frac{1}{\sqrt{4\pi}}$$

The reason for this specific constant is normalization, i.e.,

$$\int |Y_{00}(\theta, \phi)|^2 \sin(\theta) d\theta d\phi = \frac{1}{4\pi} \int \sin(\theta) d\theta d\phi$$

Integrate over ϕ ($\int d\phi = 2\pi$) to get

$$\int |Y_{00}(\theta, \phi)|^2 \sin(\theta) d\theta d\phi = \frac{1}{4\pi} 2\pi \int \sin(\theta) d\theta$$

and since

$$\int_{\theta=0}^{\theta=\pi} \sin(\theta) d\theta = -[\cos(\theta)]_{\theta=0}^{\theta=\pi} = -(-1 - 1) = 2$$

we get

$$\int |Y_{00}(\theta, \phi)|^2 \sin(\theta) d\theta d\phi = \frac{1}{4\pi} 2\pi * 2 = 1 \quad Q.E.D.$$

Similarly:

$l = 1$: three states

$$Y_{-1,1}(\theta, \phi) \equiv Y_{m=-1,l=1}(\theta, \phi) = \sqrt{\frac{3}{8\pi}} \sin(\theta) \exp(-i\phi)$$

$$Y_{0,1}(\theta, \phi) \equiv Y_{m=0,l=1}(\theta, \phi) = \sqrt{\frac{3}{4\pi}} \cos(\theta)$$

$$Y_{1,1}(\theta, \phi) \equiv Y_{m=1,l=1}(\theta, \phi) = \sqrt{\frac{3}{8\pi}} \sin(\theta) \exp(i\phi)$$

These expression seems difficult, but it is trivial to remember them (except for the normalization constants) once we realize that

$$x = r \sin(\theta) \cos(\phi)$$

$$y = r \sin(\theta) \sin(\phi)$$

$$z = r \cos(\theta)$$

and therefore

$$\sin(\theta) \exp(-i\phi) = \frac{x - iy}{r}$$

and therefore

$$Y_{-1,1}(\theta, \phi) = \sqrt{\frac{3}{8\pi}} \frac{x - iy}{r}$$

$$Y_{0,1}(\theta, \phi) = \sqrt{\frac{3}{4\pi}} \frac{z}{r}$$

$$Y_{-1,1}(\theta, \phi) = \sqrt{\frac{3}{8\pi}} \frac{x + iy}{r}$$

This strange combinations (x-iy, x+iy) are there because we looked for the eigenstates of L^2 (in this case with $l = 1$, i.e., eigenvalue $\hbar^2 l(l + 1) = \hbar^2 * 1 * 2 = 2\hbar^2$) that are also eigenvalues of L_z .

Rotations around the axis:

It turns out that such eigenstates are in a certain sense rotating around the z-axis;

Y_{11} rotates counterclockwise,

$Y_{-1,1}$ rotates clockwise,

and Y_{01} does not rotate.

More generally, we can think that:

Y_{ml} rotates around the z axis:

counterclockwise if $m > 0$,

clockwise if $m < 0$

and with a speed that's higher as $|m|$ grows.

Proof (sort of):

ϕ is the angle of rotations around the z-axis. And the ϕ –dependent part of $Y_{ml}(\theta, \phi)$ is, as we learned, $\exp(im\phi)$.

So if we include the time-dependent phase factor the time-dependent wavefunction is $\exp\left(im\phi - \frac{iE_n t}{\hbar}\right)$.

This is reminiscent of $\exp(i(kx - \omega t))$.

Like there, here too the for phase in the exponent (in blue) to remain constant means $m\phi - E_n t/\hbar=0$, i.e., ϕ grows in time if $m>0$, i.e., a counterclockwise rotations

(You may be confused since ϕ is a coordinate so how can it grow; all I mean for now is that the phase at one time at a value of phi will become, a little later in time, the phase at a slightly higher value of phi (if $m>0$)).

Real combinations:

In chemistry, when solving the time-independent Schrodinger equation (to get molecular orbitals) we don't want to deal, unless we have to, with complex quantities. Therefore, we'll take **linear combinations and define the real p states** (of course, no relation to p the momentum operator)

$$Y_{x1}(\theta, \phi) \equiv \frac{Y_{-1,1}(\theta, \phi) + Y_{1,1}(\theta, \phi)}{\sqrt{2}} = \sqrt{\frac{3}{4\pi}} \sin(\theta) \cos(\phi) = \sqrt{\frac{3}{4\pi}} \frac{x}{r}$$

$$Y_{z1}(\theta, \phi) \equiv Y_{0,1}(\theta, \phi) = \sqrt{\frac{3}{4\pi}} \frac{z}{r}$$

$$Y_{y1}(\theta, \phi) \equiv \frac{-Y_{-1,1}(\theta, \phi) + Y_{1,1}(\theta, \phi)}{\sqrt{2} i} = \sqrt{\frac{3}{4\pi}} \sin(\theta) \sin(\phi) = \sqrt{\frac{3}{4\pi}} \frac{y}{r}$$

These are perfectly fine combinations, it is easy to prove that they are orthonormal; and since they are combination of eigenvalues of L^2 with the same eigenvalue ($l = 1$, i.e., eigenvalue $2\hbar^2$) they will also be eigenstates of L^2 with the same eigenvalue.

But they are not (except for $Y_{z1}(\theta, \phi)$) eigenstates of L_z ! They don't have a definite value of the z-angular momentum!

To conclude let's look quickly at this table, which includes also write the $l = 2$ (i.e., "d") states and $l = 3$ ("f") states.

Note: d states involve product of two components from x,y,z,
f states involve product of three terms from x,y,z

TABLE 3.7
Some Spherical Harmonics

Symbol	Polar	Cartesian	Normalizat Constan
Y_{00}	1	1	$\frac{1}{2}(1/\pi)^{1/2}$
Y_{10}	$\cos \theta$	z/r	$\frac{1}{2}(3/\pi)^{1/2}$
$Y_{1\pm 1}$	$\mp(\sin \theta)e^{\pm i\phi}$	$\mp(x \pm iy)/r$	$\frac{1}{2}(3/2\pi)^{1/2}$
Y_{20}	$(3 \cos^2 \theta - 1)$	$(3z^2 - r^2)/r^2$	$\frac{1}{4}(5/\pi)^{1/2}$
$Y_{2\pm 1}$	$\mp(\sin \theta)(\cos \theta)e^{\pm i\phi}$	$\mp z(x \pm iy)/r^2$	$\frac{1}{2}(15/2\pi)^{1/2}$
$Y_{2\pm 2}$	$(\sin^2 \theta)e^{\pm 2i\phi}$	$(x \pm iy)^2/r^2$	$\frac{1}{4}(15/2\pi)^{1/2}$
Y_{30}	$(5 \cos^3 \theta - 3 \cos \theta)$	$z(5z^2 - 3r^2)/r^3$	$\frac{1}{4}(7/\pi)^{1/2}$
$Y_{3\pm 1}$	$\mp \sin \theta(5 \cos^2 \theta - 1)e^{\pm i\phi}$	$\mp(x \pm iy)(5z^2 - r^2)/r^3$	$\frac{1}{8}(21/\pi)^{1/2}$
$Y_{3\pm 2}$	$(\sin^2 \theta)(\cos \theta)e^{\pm 2i\phi}$	$z(x \pm iy)^2/r^3$	$\frac{1}{4}(105/2\pi)^{1/2}$
$Y_{3\pm 3}$	$\mp(\sin^3 \theta)e^{\pm 3i\phi}$	$\mp(x \pm iy)^3/r^3$	$\frac{1}{8}(35/\pi)^{1/2}$

Douglas, Bodie E. and Hollingsworth, Charles A. *Symmetry in Bonding and Spectra - Introduction* (Orlando, Florida: Academic Press, Inc., 1985) p. 88.

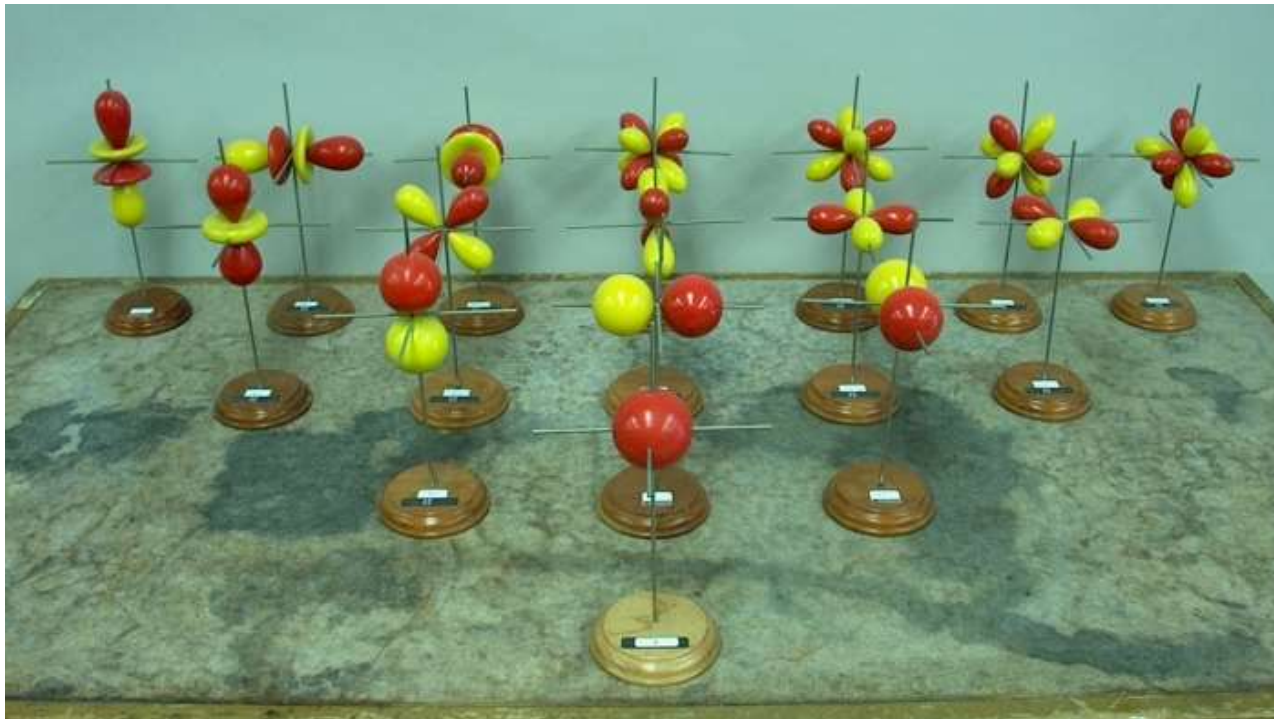
While here are real combinations of these states:

TABLE 4.1
Angular Factors of Conventional Atomic Orbitals

Symbol	Polar	Cartesian	Normalizing factor
s	1	1	$\frac{1}{2}\left(\frac{1}{\pi}\right)^{1/2}$
p_x	$\sin \theta \cos \phi$	x/r	$\frac{1}{2}\left(\frac{3}{\pi}\right)^{1/2}$
p_y	$\sin \theta \sin \phi$	y/r	$\frac{1}{2}\left(\frac{3}{\pi}\right)^{1/2}$
p_z	$\cos \theta$	z/r	$\frac{1}{2}\left(\frac{3}{\pi}\right)^{1/2}$
d_{z^2}	$(3 \cos^2 \theta - 1)$	$(3z^2 - r^2)/r^2$ $(2z^2 - x^2 - y^2)/r^2$	$\frac{1}{4}\left(\frac{5}{\pi}\right)^{1/2}$
d_{xz}	$\sin \theta \cos \theta \cos \phi$	xz/r^2	$\frac{1}{2}\left(\frac{15}{\pi}\right)^{1/2}$
d_{yz}	$\sin \theta \cos \theta \sin \phi$	yz/r^2	$\frac{1}{2}\left(\frac{15}{\pi}\right)^{1/2}$
$d_{x^2-y^2}$	$\sin^2 \theta \cos 2\phi$	$(x^2 - y^2)/r^2$	$\frac{1}{4}\left(\frac{15}{\pi}\right)^{1/2}$
d_{xy}	$\sin^2 \theta \sin 2\phi$	xy/r^2	$\frac{1}{4}\left(\frac{15}{\pi}\right)^{1/2}$
f_{z^3}	$(5 \cos^3 \theta - 3 \cos \theta)$	$z(5z^2 - 3r^2)/r^3$ $[2z^3 - 3z(x^2 + y^2)]/r^3$	$\frac{1}{4}\left(\frac{7}{\pi}\right)^{1/2}$
f_{xz^2}	$(5 \cos^2 \theta - 1) \sin \theta \cos \phi$	$x(5z^2 - r^2)/r^3$ $[4xz^2 - x(x^2 + y^2)]/r^3$	$\frac{1}{8}\left(\frac{42}{\pi}\right)^{1/2}$
f_{yz^2}	$(5 \cos^2 \theta - 1) \sin \theta \sin \phi$	$y(5z^2 - r^2)/r^3$ $[4yz^2 - y(x^2 + y^2)]/r^3$	$\frac{1}{8}\left(\frac{42}{\pi}\right)^{1/2}$
f_{xyz}	$\sin^2 \theta \cos \theta \sin 2\phi$	xyz/r^3	$\frac{1}{4}\left(\frac{105}{\pi}\right)^{1/2}$
$f_{z(x^2-y^2)}$	$\sin^2 \theta \cos \theta \cos 2\phi$	$z(x^2 - y^2)/r^3$	$\frac{1}{4}\left(\frac{105}{\pi}\right)^{1/2}$
$f_{x(x^2-3y^2)}$	$\sin^3 \theta \cos 3\phi$	$x(x^2 - 3y^2)/r^3$	$\frac{1}{8}\left(\frac{70}{\pi}\right)^{1/2}$
$f_{y(3x^2-y^2)}$	$\sin^3 \theta \sin 3\phi$	$y(3x^2 - y^2)/r^3$	$\frac{1}{8}\left(\frac{70}{\pi}\right)^{1/2}$

Douglas, Bodie E. and Hollingsworth, Charles A. *Symmetry in Bonding and Spectra - An Introduction* (Orlando, Florida: Academic Press, Inc., 1985) p. 123.

And here are some models for these states, where “s” is in front, then the Cartesian “p’s”, etc.⁴



Now to conclude we note that these angular functions we introduced should be multiplied by the radial functions.

For the “s” states there is no angular factor to multiply by (except for a constant $1/\sqrt{4\pi}$.)

For the p states the multiplication cancels the “r” from the radial; function and the “1/r” in the Cartesian form of the angular function, for example

$$\psi_{n=1,l=1,m=0}(r, \theta, \phi) = \text{const.} * r \exp\left(-\frac{r}{2a_0}\right) * \frac{z}{r}$$

$$\psi_{n=1,l=1,m=0}(r, \theta, \phi) = \text{const.} * \exp\left(-\frac{r}{2a_0}\right) * z$$

The same for all other combinations.

⁴ Taken from <http://web.physics.ucsb.edu/~lecturedemonstrations/Composer/Pages/88.18.html>, where I also took the tables from.

Operator Approach to Angular Momentum

So far you learned that the solution to the blue equation a few pages ago,

$$\frac{-\hbar^2}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial g(\theta)}{\partial \theta} \right) - \frac{\hbar^2}{\sin^2 \theta} m^2 g(\theta) = \lambda g(\theta)$$

yields that the eigenvalues are $\lambda = -\hbar^2 l(l + 1)$.

We could have shown it with quite a bit of differential equations solving-skills. But viewing it this way does a disservice to angular momentum, as it is hard to generalize (to many particles, spin etc.).

I'll therefore quickly enumerate the main steps over the more advanced and elegant operator approach to angular momentum. The key will be that the properties we will find will depend only on the commutation relations:

Let's enumerate these properties (some of which we touched on earlier); we won't prove many of them.

The angular momentum operator is Hermitian $L_x^\dagger = L_x$, etc

Each element of the angular momentum does not commute with the other elements, $[L_x, L_y] = -i\hbar L_z$, etc.

All individual elements do commute however with L^2 ,

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

Conservation of the angular momentum is because the potential does not depend on the angle; formally it translates to

$$[H, L_x] = [H, L_y] = [H, L_z] = [H, L^2] = 0$$

Properties (ii) and (iii) imply that we cannot find a complete common set of eigenfunctions of L_z and L_x or L_y , but we can find a set of eigenfunction of L^2 and one of the components, traditionally taken as L_z . These eigenfunctions will fulfill therefore

$$L^2 Y = \lambda Y$$

$$L_z Y = m\hbar Y$$

Because L_x has real eigenvalues, L_x^2 is positive definite, as is L_y^2 . This implies that the eigenvalues of $L^2 - L_z^2 = L_x^2 + L_y^2$ are positive. This implies that the eigenvalue of the total angular momentum is higher or equal to the square of the eigenvalue of L_z :

$$\lambda \geq m^2 \hbar^2$$

The next stage will be to do something similar to the ladder for Harmonic oscillators. I.e., we define a raising operator

$$L_+ \equiv L_x + iL_y$$

and its Hermitian conjugate, the lowering operator

$$L_- \equiv L_x - iL_y$$

The raising operator does not change the energy of the state it acts on (unlike in the harmonic oscillator case). **But**

it raises m by 1 (keeping λ fixed). I.e., it raises the angular momentum in the z direction, keeping the total angular momentum fixed.

And it multiplies the resulting state (with m+1) by an important factor, $\sqrt{\lambda - \hbar^2 m(m+1)}$, i.e.

Similarly, the lowering operator

Lowers m by 1 (keeping λ fixed).

Multiplies the state with (m-1) by $\sqrt{\lambda - \hbar^2 m(m-1)}$

So now comes an argument analogous to the ladder in the harmonic oscillator: given any m, we can make a state, with the same λ , but with m+1, m+2, m+3, ...

Eventually m will be so $m^2 \hbar^2$ will be bigger than λ !

The only solution, as in the Harmonic oscillator case, is that when acting with L_+ on the highest state possible for a given λ , (a state we'll label as m_{top}) we'll get a zero!

But we know what the factor is (I wrote it above), so we can write it:

$$\sqrt{\lambda - \hbar^2 m_{top}(m_{top} + 1)} = 0$$

i.e.,

$$\lambda = \hbar^2 l(l + 1)$$

And

$$m_{top} = l$$

Furthermore, if we apply the lowering operator on a state with a given m , we'll go to $m-1, m-2, m-3, \dots$ and since m could be negative eventually we will reach a state with a large negative m such that $\hbar^2 m^2 > \lambda$ unless applying L_- on the lowest possible state, m_{bot} will give 0.

i.e., in addition to the equation above we'll need

$$\sqrt{\lambda - \hbar^2 m_{bot}(m_{bot} - 1)} = 0$$

The only way for this to be true is

$$m_{bot} = -m_{top}$$

But since when we move from one m to the other we step down by an integer, it must be that

$$m_{top} - m_{bot} = Integer$$

i.e.,

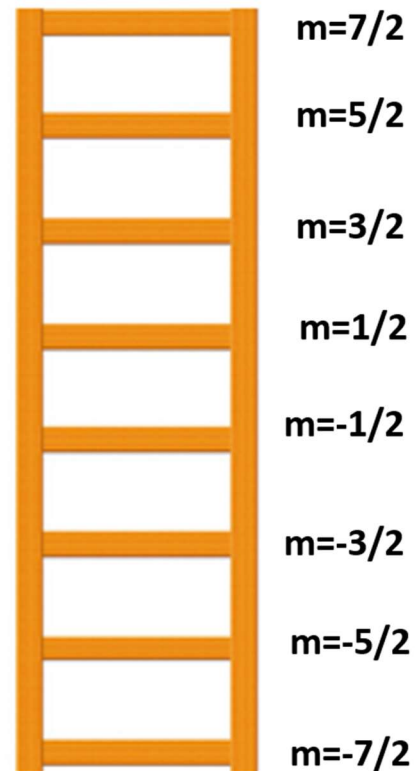
$$2m_{top} = integer$$

i.e.,

$$2l = integer.$$

Ladder for $l = \frac{7}{2}$ (so $\lambda = l(l + 1)\hbar^2$)

Note that if we were to try to shift the m series to be non-integer and non-half-integer Then it will never terminate – i.e., to terminate It needs to terminate both at the bottom and at the top , which requires the top and bottom m 's to have the same absolute value (7/2 here).



MADE-UP EXAMPLES.

Let's make up numbers, so you can see. I'll use atomic units, $\hbar = 1$, for simplicity.

Say we had $\lambda = 32.6$ (a made up value, which you'll soon see is impossible).

And say there was a state with $m = -1.3$ (again, a made up value which is impossible, you'll soon see).

Then we can make using a lowering operators several times, states with $m = -2.3, -3.3, -4.3, -5.3, -6.3, \dots$

And by applying the raising operator, we'll have states with

$$m = -0.3, \quad 0.7, \quad 1.7, \quad 2.7, \quad 3.7, \quad 4.7, \quad 5.7, \quad 6.7, \dots$$

Well, that creates a problem! Since m^2 for the state with $m=5.7$ is already high ($5.7^2 = 32.49$), i.e., higher than λ .

So we may think that all we have to do is change λ , to be for example, $5.7^2 = 32.49$.

But that won't work, since when we use the negative m , we get that the most negative m we have on the list already has

$$m_{bot}^2 = (-6.3)^2 > 32.49, \quad i.e.,$$
$$m_{bot}^2 > \lambda$$

So the only way to have a solution will be for m to be either

$$m = 0, \pm 1, \pm 2$$

Or

$$m = \pm \frac{1}{2}, \pm \frac{3}{2}, \dots$$

As shown in the figure, and therefore

$$\lambda = \hbar^2 l(l+1)$$

l integer or half - integer

$$|m| \leq l$$

Some comments on angular momentum

Before going on to spin (which is also a form of angular momentum) note a few important points:

Never confuse ϕ in the spherical coordinates with ϕ_n used for eigenfunctions... not enough letters in the alphabet!

Angular momentum is related to rotation. We see it in

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

and a derivative is related to difference, i.e., change.

This implies (with some math) that when a wavefunction is rotated, or when the axis are rotated, the components of L are involved.

Since the components of L don't commute among themselves, they cannot share the same list of eigenfunctions.

We can prove this in general, before seeing implications

Two operators that share the same eigenfunction must commute.

Proof:

If A, B share the same eigenfunctions, $\phi_n(x)$, with eigenvalues a_n, b_n then a general function $f(x)$ can be written (assuming the set of eigenfunctions is complete) as

$$f = \sum_n c_n \phi_n$$

Therefore:

$$\begin{aligned} ABf &= A(Bf) = A\left(B \sum_n c_n \phi_n\right) = A\left(\sum_n c_n B\phi_n\right) = A\left(\sum_n c_n b_n \phi_n\right) \\ &= \left(\sum_n c_n b_n A\phi_n\right) \end{aligned}$$

i.e., putting back the coordinates so we remember what are constants and what are functions (or vectors)

$$ABf(x) = \sum_n c_n b_n a_n \phi_n(x)$$

Clearly we'll get the same result if we were to apply it in the reverse order, BAf , i.e.,

$$ABf = B Af$$

i.e.,

$$[A, B]f = 0$$

Since f is a general function, it proves that in general

$$[A, B] = 0$$

END OF PROOF.

The implications are here that, for example, L_x and L_z don't share the same list of eigenfunctions.

This is easy to see, e.g., look at the “pz” eigenstate of L_z , i.e.,

$$Y_{01}(\theta, \phi) = \sqrt{\frac{3}{4\pi}} \frac{z}{r}$$

(i.e., $L_z Y_{01} = 0 * L_{0,1}$, i.e., a state with $m=0$, i.e., an eigenstate of L_z with eigenvalue 0.)

Clearly, the eigenstate of L_x with eigenvalue 0 must be “px”, i.e., x/r

But recall that

$$Y_{\pm 1,1}(\theta, \phi) = \sqrt{\frac{3}{8\pi}} \frac{x \pm iy}{r}$$

so

$$\sqrt{\frac{3}{4\pi}} \frac{x}{r} = \frac{Y_{-1,1}(\theta, \phi)}{\sqrt{2}} + \frac{Y_{1,1}(\theta, \phi)}{\sqrt{2}}$$

i.e., “px”, a state with an eigenvalue 0 for L_x , is a specific combination of:
 a state with an $+1 * \hbar$ eigenvalue L_z , and
 a state with eigenvalue $-1 * \hbar$ of L_z .

- I.e., the eigenfunction of one operator L_x is not the eigenfunction of another operator, L_x

Meaning of rotation – interaction with magnetic field

Finally, I want to come back to the meaning of $m < 0$ or $m > 0$.

Let's say we have a magnetic field. Classically, this magnetic field will interact with rotating currents;

Therefore, a magnetic field (\mathbf{B}) would raise or lower the energy of particle that rotates clockwise or anticlockwise around the axis of the magnetic field.

So if we have a magnetic field then this field will interact differently with wavefunction that's rotating clockwise, not rotating, or rotating anticlockwise.

Formally, the interaction will be of the form

$$\text{const.} * \mathbf{B} \cdot \mathbf{L}$$

So if we define the z-axis to be along the field ($\mathbf{B} = B_z \hat{\mathbf{z}}$), then the field will be:

$$\text{const.} * B_z L_z$$

So you see that when acting on $Y_{ml}(\theta, \phi)$ states then the energy will be

$$\text{const.} * B_z m$$

i.e., the **energies will split** (be different for different m states). This ("Zeeman") split is small, but it could be measured. We'll talk about a similar split when we talk about spin and about perturbation theory.

Adding up angular momenta

This is a very important subject.

BTW, we'll use atomic units ($\hbar = 1$) in this section

Angular momentum addition crops up in three major places

Adding up the angular momenta of two particles, labeled I and II;
 $\mathbf{L}^{tot} = \mathbf{L}^{(I)} + \mathbf{L}^{(II)}$; for examples, what is the angular momenta of a molecular where the only contribution to the angular momenta are from two electrons, one with $l^I = 4$ and the other with $l^{II} = 1$.

First of all, the sum of two angular momenta for each particle is also an angular momentum operator, i.e., it satisfies for example

$$\text{claim: } [L_x^{tot}, L_y^{tot}] = i\hbar L_z^{tot}$$

This is because the angular momenta for one particle and for the other commute, e.g., $[L_x^I, L_y^{II}] = 0$ (they involve coordinate and derivatives of two different particles, so the derivative w.r.t. the position of one does not affect the coordinate of the other). Therefore

$$\begin{aligned} [L_x^{tot}, L_y^{tot}] &= [L_x^I + L_x^{II}, L_y^I + L_y^{II}] = \\ &= [L_x^I, L_y^I] + [L_x^I, L_y^{II}] + [L_x^{II}, L_y^I] + [L_x^{II}, L_y^{II}] = \\ &= i\hbar L_z^I + 0 + 0 + i\hbar L_z^{II} \\ &= i\hbar L_z^{tot} \quad Q.E.D. \end{aligned}$$

Because the angular momentum fulfills the commutation relations, it also is guaranteed to have integer values of l^{tot} , i.e., the eigenvalues of L_{tot}^2 will be $\hbar^2 l_{tot}(l_{tot} + 1)$, etc. Further

$$m_{tot} = m^I + m^{II}$$

and

$$-l_{tot} \leq m_{tot} \leq l_{tot}$$

What values will l_{tot} take?

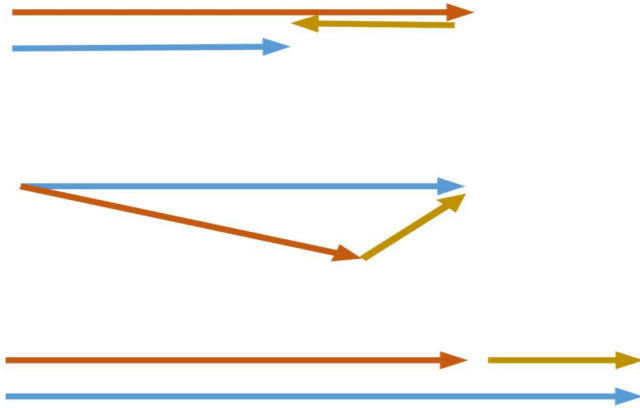
It turns out the allowed values, here and in the other case, are the same as in classical addition of vectors, i.e., for two vectors which can point each in any direction, the length of their sum can range between the sum of their value to the difference in their value, i.e., classically if we add a “red” and “brown” vector to get a “blue” vector

$$\mathbf{L} = \mathbf{L}^I + \mathbf{L}^{II}$$

Then the length of the blue vector is within the range:

$$\left| |\mathbf{L}^I| - |\mathbf{L}^{II}| \right| \leq |\mathbf{L}| \leq |\mathbf{L}^I| + |\mathbf{L}^{II}|$$

(see picture)



Quantally the same rule (known as the triangle rule) applies

$$|l^I - l^{II}| \leq l_{tot} \leq l^I + l^{II}$$

So in our case ($l^I = 4$, $l^{II} = 1$) $3 \leq l_{tot} \leq 5$,

i.e., its allowed values are 3,4,5.

Next, turn to adding up the orbital angular momenta ($L = r \times p$) that we learned about, together with angular momenta due to **spin** (S), which as we'll see in the next chapter, has only one possible

eigenvalue, $s = \frac{1}{2}$ (*i.e.*, the eigenvalue of S^2 is

$$\hbar^2 * \frac{1}{2} * \left(\frac{1}{2} + 1\right) = \frac{3}{4} \hbar^2$$

The total angular momentum is labeled $J = L + S$. We'll see that the eigenvalue of the spin from one electron is $s = \frac{1}{2}$ (no relation of this s to the "1s", "2s", etc). So say $l = 4$, what's the allowed values of **J**?

The triangle rule is valid here too, telling us that

$$3.5 \leq j \leq 4.5$$

i.e., $j = 3.5$ or $j = 4.5$ (the eigenvalues have to be shifted by 1, which we won't prove, so $j = 4$ is not allowed).

Adding up two spins (spin for one electron and for another): same story.

$$s^I = \frac{1}{2}, \quad s^{II} = \frac{1}{2},$$

And therefore the total spin is:

$$s = \left(\frac{1}{2} - \frac{1}{2}\right) \text{ or } \left(\frac{1}{2} + \frac{1}{2}\right) \quad \text{i.e.,} \quad 0 \text{ or } 1.$$

i.e., either:

$$s = 0 \rightarrow s_z = 0$$

$$s = 1 \rightarrow s_z = -1, 0, 1$$

If the two spins add up to a total spin of 0, we'll call this **singlet** (since only one value of s_z is present)

And when the two spins add up to a total spin of 1, it's a **triplet** (three values of s_z possible).

We'll construct the explicit eigenfunctions of the total spin for this case in the next chapter.

SPIN

When a beam of electrons is passed through a magnet some will be repelled and some will be attracted.

This points out that there's an underlying angular momentum "spin" to the electrons, since we said earlier that the interaction of a magnetic field with a system is of the form

$$\text{const. } \mathbf{B} \cdot \mathbf{L}$$

Since the beam splits to two sub-beams, we need to have an angular momentum with only two eigenvalues, i.e.,

$$s = \frac{1}{2}$$
$$s_z = -\frac{1}{2}, +\frac{1}{2}$$

This is exactly the $\frac{1}{2}$ integer angular momenta we talked about.

But it turns out we cannot find internal coordinates for the spin of the electron (i.e., an electron is not a rigid three-D body where it is meaningful to talk about it spinning around).

Instead, we'll have to be content with a matrix description.

That's ok since we learned before that any function can be described in terms of a basis set times coefficients.

I.e., when we describe a function $f(x)$ we can talk about $f(x)$ or the expansion coefficients f_n in terms of an orthogonal basis, $f(x) = \sum_n f_n \phi_n(x)$.

So here we will say that an electron spin wavefunction (or more precisely the spin part of the wavefunction) can be written as a linear combination of "spin-up" and "spin-down" states.

What do we mean by up and down? That's relative to a specific axis, i.e., as we'll see later an "up" for one axis is not "up" for another axis, just like a $2p_z$ function is not an $m=0$ function if we measure relative to another axis.

So anyway, choose an axis (usually we use the z-axis), and then we'll write in abstract notation two states, a spin up state denoted usually as α ,

$$|\alpha\rangle \equiv \left|s_z = \frac{1}{2}\right\rangle$$

And an analogous spin-down state

$$|\beta\rangle \equiv \left|s_z = -\frac{1}{2}\right\rangle$$

Any spin wavefunction could be written as a linear combination

$$c_+|\alpha\rangle + c_-|\beta\rangle$$

i.e., a spin wavefunction can be represented as a length-two vector

$$\begin{pmatrix} c_+ \\ c_- \end{pmatrix}$$

That's as far as the wavefunction. But what's the spin operator, \mathbf{S} ?

Well each of the spatial components of \mathbf{S} (i.e., S_x, S_y, S_z) needs, when it acts on a spin wavefunction to produce another wavefunction, i.e., if we use the language of column vectors and matrices, it needs to take a length-2 column vector and transform it to another column vector.

i.e., S_z , the z-part of the spin- angular momentum, needs to **be a 2*2 matrix** (or more formally it can be represented by a 2*2 matrix).

It is easy to see what it is. If acting on a spin-up, S_z should give back the same spin-up eigenfunction times the value of spin up, i.e., times $\hbar * \frac{1}{2}$

i.e., the matrix S_z acting on a spin up should yield

$$S_z \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

And similarly for spin-down

$$S_z \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\frac{\hbar}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

This implies immediately that

$$S_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

When I was at your stage I was very confused by this statement.

What does it mean that S_z is a matrix?

All it means actually is something very simple: in the basis which uses spin up and spin down along the z-axis, the action of the “formal” S_z operator, i.e., the operator of the spin along the z-axis, can be represented by the yellow matrix above.

Next, to S_x, S_y . We need to find what their value should be. Well, we know that they need to fulfill the commutation relation for angular momenta, which are

$$[S_x, S_y] = i\hbar S_z$$

And similarly (rotate the indices as usual, $x \rightarrow y \rightarrow z \rightarrow x$)

$$[S_y, S_z] = i\hbar S_x$$

$$[S_y, S_x] = i\hbar S_y$$

Pauli worked out the math and in turns out that the matrices which fulfill these relations (i.e., the matrices representing S_x, S_y) are

$$S_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$S_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

Note

These spin matrices are Hermitian, as they should be for measurable operators.

Let’s check, for example, whether $[S_x, S_y] = i\hbar S_z$ is fulfilled; we worked it out for operators but the same relation works for matrices, so:

$$[S_x, S_y] = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} - \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$= \frac{\hbar}{2} \frac{\hbar}{2} \left(\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} - \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \right) = i\hbar \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = i\hbar S_z \quad Q.E.D.$$

The other light-blue relations can be shown similarly.

Let's check on this simplest of cases (smallest angular momentum possible) what we wrote earlier on raising and lowering operators works. For example, we said that the raising operator is

$$S_+ \equiv S_x + iS_y$$

Which in this case will be

$$S_+ \equiv \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + i \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \hbar \\ 0 & 0 \end{pmatrix}$$

And indeed when we act with the raising operator on a spin down state we get a spin up state

$$S_+ \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 & \hbar \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \hbar \\ 0 \end{pmatrix}$$

But when we act on spin up state the result vanishes, as it should (we cannot raise a single-electron spin state above spin up)

$$S_- \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & \hbar \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad Q.E.D$$

Interlude: Angular-momentum related Interactions in atoms.

In addition to the dominant electron-nucleus electrical interaction, there are several weaker interaction that need to be considered.

We're generally considering three such static interactions at present (later we'll consider also interactions with an oscillating electric field).

They are due to the interaction of magnetic fields with currents. Currents relate to angular momentum (which is an operator related to angular motion).

(i) So the first is orbital-angular-momentum interaction with a magnetic field:

$$H_{O.B.} = -const.* \mathbf{L} \cdot \mathbf{B}$$

(where "O.B." means the interaction of the Orbital angular momentum with the magnetic field). And I include the minus, which ensures that if the m-

state (in an axis parallel to B) is positive, i.e., L is in some sense along B (and not anti-parallel to it) then the energy will decrease, i.e., will be more negative so preferable.

(ii) Then there's a similar interaction with the spin

$$H_{S,B} = -const.* \mathbf{S} \cdot \mathbf{B}$$

As you can imagine the constants in the $H_{O,B}$, $H_{S,B}$ interactions are quite similar.

Similarly, note the H_{SB} has the minus so a spin parallel to B (and not antiparallel to it) will be preferable (i.e., have lower energy).

(iii) And finally, the orbital angular momentum itself creates, as it is a current, a magnetic field, and this magnetic field interacts with the spin (and of course this could be viewed the opposite, i.e., the spin current creates a magnetic field which interacts with the orbital angular momentum; or simply we know from magnetic theory that two currents will interact.

At any rate, this leads to the spin orbit interaction.

Turns out this interaction also depends strongly on the type of atom, due to a dependence on the distance from the nucleus, etc.

$$H_{SO} = -c_{atom} \mathbf{S} \cdot \mathbf{L}$$

You notice the spin orbit interaction is an internal interaction, i.e., exists regardless of whether there's an external field. It is negligible for light atoms, but for heavy atoms, due to relativistic effects, it is very large and affects the spectrum of transition metals. We'll deal with parts of it below and parts at the end.

END OF INTERLUDE

Spin Measurement

Spin measurement is a very simple way to see quantum measurement at action.

(I) First, a single measurement:

Pass in a strong magnetic field \mathbf{B} a beam of hydrogen atoms

They are all at their ground state, so $l = 0$ so we don't have to worry about magnetic-field interaction with \mathbf{L} (i.e., the $\mathbf{L} \cdot \mathbf{B}$ term is zero).

But there is a strong $\mathbf{S} \cdot \mathbf{B}$ interaction.

If the magnetic field is along the z-axis, then electrons with spin up will go towards the field, and electrons with spin down will go away

I.e., a beam with a magnetic field is a measurement device which measures the spin component along the magnetic field.

Say initially we had an unpolarized beam and the field is along the z-axis.

Then after passing through the magnetic field $\frac{1}{2}$ the atoms in average will have $|\alpha\rangle$ state, i.e., be spin up (along z). And they will be deflected one way (say to the right, if the z-axis is horizontal and positive to the right).

And another $\frac{1}{2}$ will have $|\beta\rangle$ and be deflected the other way.

So if an electron initially had a spin wavefunction of say,

$$0.8|\alpha\rangle - 0.6 * i * |\beta\rangle$$

(I am just making some numbers up)

Then there's a $0.64 = 0.8^2$ chance that it will end up to the the right, with a spin wavefunction of $|\alpha\rangle$

and 0.36 chance that it will end up to the left with a spin off $|\beta\rangle$

But it will not stay in the middle! It must make a choice! – our device measures the spin!

Now a consecutive measurement

Now we take the electrons that went to the right. All of them have spin up, $|\alpha\rangle$, which we write also as

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Now pass them through another region where the magnetic field is along the x axis, (i.e., B_x is positive, the rest are zero).

This field will split this right beam to two sub-beams, one that has spin-up along the x-axis, one that has spin-down along the x-axis!!!

The electrons, which are in spin state $|\alpha\rangle$ (associated with spin-up along the z-axis) are not in an eigenstate of S_x .

The eigenstates of S_x are easily shown to be (in the $|\alpha\rangle, |\beta\rangle$ basis):

$$\begin{pmatrix} 1 \\ \frac{1}{\sqrt{2}} \\ 1 \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

And

$$\begin{pmatrix} 1 \\ \frac{1}{\sqrt{2}} \\ 1 \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$

i.e., we can write them a

$$\frac{|\alpha\rangle + |\beta\rangle}{\sqrt{2}}: \text{ spin - up - along - } x$$

$$\frac{|\alpha\rangle - |\beta\rangle}{\sqrt{2}}: \text{ spin - down - along - } x$$

The electrons which enter the 2nd magnetic-field region therefore with spin $|\alpha\rangle$, must make a choice – some of them will end up with $\frac{|\alpha\rangle + |\beta\rangle}{\sqrt{2}}$, some with $\frac{|\alpha\rangle - |\beta\rangle}{\sqrt{2}}$.

How many? From symmetry we can guess $\frac{1}{2}$ and $\frac{1}{2}$ (i.e., $\frac{1}{4}$ and $\frac{1}{4}$ of the original beam).

Indeed, to find out how many we need to take the electron wavefunction (as the entrance to the second magnetic field), and calculate the absolute-value of its overlap with the eigenstate, i.e.,

Prob. for the spin to point eventually along the positive x-axis= square of overlap of the associated eigenstate of S_x

$$\begin{pmatrix} 1 \\ \frac{1}{\sqrt{2}} \\ 1 \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

with the initial post-first-beam spin wavefunction

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

i.e., the overlap will be

$$\left| 1 * \frac{1}{\sqrt{2}} + 0 * \frac{1}{\sqrt{2}} \right|^2 = \frac{1}{2}$$

So indeed $\frac{1}{2}$ of the post-first-field beam will emerge with spin parallel to the positive x-axis.

LONG Exercise: let's repeat this exercise if the magnetic field in the 2nd stage points along an axis we'll label as **d**, in the x-z plane at 60 degrees to the positive x and 30 degrees to positive z axis.

(i.e., the first magnetic field points along the z-axis; the 2nd along the **d** axis).

For the exercise we'll need to determine the eigenstates of the associated spin-operator.

Remember that we said that the spin is a vector, and found out what S_x and S_z are (and S_y , irrelevant here).

So the spin along the **d** axis must be a (normalized) combination of these two vector components.

We know that a unit vector along \mathbf{d} has components

$$\hat{\mathbf{d}} = \cos(30^\circ) \hat{\mathbf{x}} + \cos(60^\circ) \hat{\mathbf{z}} = \frac{\sqrt{3}}{2} \hat{\mathbf{x}} + \frac{1}{2} \hat{\mathbf{z}}$$

So

$$S_{\mathbf{d}} = \cos(60^\circ) S_x + \cos(30^\circ) S_z$$

i.e.,

$$S_{\mathbf{d}} = \frac{1}{2} S_x + \frac{\sqrt{3}}{2} S_z = \frac{1}{2} \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \frac{\sqrt{3}}{2} \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

i.e.,

$$S_{\mathbf{d}} = \hbar \begin{pmatrix} \frac{\sqrt{3}}{4} & \frac{1}{4} \\ \frac{1}{4} & -\frac{\sqrt{3}}{4} \end{pmatrix}.$$

To simplify, let's use atomic units, where $\hbar = 1$, so

$$S_{\mathbf{d}} = \begin{pmatrix} \frac{\sqrt{3}}{4} & \frac{1}{4} \\ \frac{1}{4} & -\frac{\sqrt{3}}{4} \end{pmatrix}$$

What are the eigenvalues and eigenfunctions of this operator? Actually, we know the eigenvalues. They **must be** $\frac{1}{2}$ **and** $-\frac{1}{2}$! This is because we are looking at the spin operator along a given direction, and the spin along one direction cannot be fundamentally different than the spin along another, i.e., we should get the same measurable values, i.e., eigenvalues.

Let's check that – I remind you of what you learned in linear algebra, i.e., **for γ to be an eigenvalue of the matrix $S_{\mathbf{d}}$, the determinant of $S_{\mathbf{d}} - \gamma I$ must vanish!**

$$0 = \det(S_d - \gamma I) = \det \begin{pmatrix} \frac{\sqrt{3}}{4} - \gamma & \frac{1}{4} \\ \frac{1}{4} & -\frac{\sqrt{3}}{4} - \gamma \end{pmatrix}$$

$$0 = \left(\frac{\sqrt{3}}{4} - \gamma\right) \left(-\frac{\sqrt{3}}{4} - \gamma\right) - \frac{1}{4} \frac{1}{4}$$

$$0 = \gamma^2 - \frac{3}{16} - \left(\frac{1}{4}\right)^2$$

$$0 = \gamma^2 - \frac{4}{16}$$

i.e.,

$$\gamma = \pm \frac{1}{2}$$

as we predicted!

Now we can determine the eigenstates of S_d , as promised.

Take $\gamma = \frac{1}{2}$, i.e., spin up when measured along the \mathbf{d} direction. That's the eigenvector we're interested in, since our measurement will be to find the proportion of particles that will emerge with spin up.

Write the eigenvector as the first column

$$v \equiv \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

So we know that

$$S_d v = \gamma v = \frac{1}{2} v$$

i.e.,

$$\begin{pmatrix} \frac{\sqrt{3}}{4} & \frac{1}{4} \\ \frac{1}{4} & -\frac{\sqrt{3}}{4} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

Let's look at the first row of the resulting equation

$$\frac{\sqrt{3}}{4}v_1 + \frac{1}{4}v_2 = \frac{1}{2}v_1$$

i.e.,

$$\left(\frac{\sqrt{3}}{4} - \frac{1}{2}\right)v_1 + \frac{1}{4}v_2 = 0$$

i.e.,

$$v_2 = 2 - \sqrt{3}$$

So one choice could be, for example, choosing $v_1 = 1$ and then

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 - \sqrt{3} \end{pmatrix} = \begin{pmatrix} 1 \\ 0.268 \end{pmatrix}$$

But that choice is not normalized so we need to divide by $\sqrt{1 + 0.268^2} = 1.035$, i.e.,

$$v = \frac{1}{1.035} \begin{pmatrix} 1 \\ 0.268 \end{pmatrix} = \begin{pmatrix} 0.966 \\ 0.255 \end{pmatrix}$$

i.e., we can think of a spin-up along \mathbf{d} as being

$$0.966|\alpha\rangle + 0.255|\beta\rangle$$

So therefore when particles come after the first detector with spin up along z , i.e., with spin state $|\alpha\rangle$, the probability to emerge with spin-up along \mathbf{d} will be the overlap-squared, i.e.,

$$\text{overlap} - \text{squared of: } \mathbf{0.966}|\alpha\rangle + \mathbf{0.255}|\beta\rangle \quad \textit{with } |\alpha\rangle$$

i.e.,

$$P = (0.966)^2 = 0.933$$

So 93% of the particles that enter the second detector (and all have spin up along \hat{z}) will emerge from the second detector with spin up along \mathbf{d} .

i.e., since $\frac{1}{2}$ only emerged from the 1st detector with spin up, we get that the total probability to:

1) leave the 1st detector with spin up along \hat{z} and then

2) emerge from the 2nd detector with spin up along **d**

Will be

$$\frac{1}{2} * 0.933 = 46.6\%$$

Exercise:

If now, after 46.6% of the particles emerge from the 2nd detector with spin up along **d**, and we add another measurement, this time again along z, how many particles will emerge with spin down along z? (i.e., spin up along z, then up along **d**, then down along z?)

Answer: I leave it to you show that the answer is

$$0.258^2 = 0.067$$

of the particles that enter the 3rd stage, i.e.,

$$0.466 * 0.067 = 0.03125$$

i.e., 3% of all particles

Next back to chemistry: triplets and singlets!

Spin triplets and singlets: Pauli principle and Hund's rules.

The most important aspect of spins to chemistry is, of course that because of them two electrons can be in the same spatial state (otherwise H and B would have been a closed-shell entities, as they would have filled the 1S and the 2s-2p shells – but in reality it is He and Ne, of course)

The essential principle guiding the placement of electrons in shells is the Pauli principle, which states

The wavefunction of the full system is completely antisymmetric w.r.t. exchanging two electrons.

Let's see what this tells us on He.

In He, to a good approximation, the two electrons (which we'll label as "1" and "2" _are in the same "1s" spatial state. (**again, this 1s has nothing to do with spin, it just refers historically to l=0**).

So if we were to exchange electron '1" and "2" then just the spatial part would not change , i.e., it would be symmetric under exchange. So where's the "minus"? where's the antisymmetry?

It is in the spin part! So let's talk about:

Spin combinations:

Give electron 1 and electron 2, how do we write the spin state of each?

Notation can be cumbersome, but I'll use following notation: order the spin states and put in a bra-ket. And at times label with the electron.

I.e.,

$$|\beta_1 \alpha_2\rangle$$

or just

$$|\beta\alpha\rangle$$

will mean that the 1st electron is down (in state $|\beta\rangle$) and the 2nd one is up.

(Similarly for 3 electrons, $|\beta\alpha\beta\rangle$ will mean that the 1st and 3rd electrons have spin down, and the 2nd has spin up).

Now, let's look (for 2 electrons) for spin combinations that are completely antisymmetric:

Antisymmetric ("singlet"):

$$\text{Singlet: } \frac{|\alpha\beta\rangle - |\beta\alpha\rangle}{\sqrt{2}}$$

Proof: when we exchange 1 and 2, then $|\beta\alpha\rangle$ becomes $|\alpha\beta\rangle$ (i.e., instead of electron 1 being in spin-down, it is in spin-up now, etc.), so

$$\text{Exchange} \left(\frac{|\alpha\beta\rangle - |\beta\alpha\rangle}{\sqrt{2}} \right) = \frac{|\beta\alpha\rangle - |\alpha\beta\rangle}{\sqrt{2}} = - \left(\frac{|\alpha\beta\rangle - |\beta\alpha\rangle}{\sqrt{2}} \right)$$

Side note: The $\sqrt{2}$ is for **normalization**:

$$\begin{aligned} \text{Norm squared of } \left(\frac{|\alpha\beta\rangle - |\beta\alpha\rangle}{\sqrt{2}} \right) &= \frac{\langle\alpha\beta| - \langle\beta\alpha|}{\sqrt{2}} \cdot \frac{|\alpha\beta\rangle - |\beta\alpha\rangle}{\sqrt{2}} \\ &= \frac{1}{2} (\langle\alpha\beta|\alpha\beta\rangle - \langle\beta\alpha|\alpha\beta\rangle - \langle\alpha\beta|\beta\alpha\rangle + \langle\beta\alpha|\beta\alpha\rangle) \\ &= \frac{1}{2} (1 - 0 - 0 + 1) = 1 \quad Q.E.D. \end{aligned}$$

Note: $\langle\beta\alpha|\alpha\beta\rangle=0$ since it means – overlap of (spin-down for electron 1 and spin-up for electron 2) with (spin-up for electron 1 and spin-down for electron 1),

Which is 0, as spin-down for electron 1 has zero overlap with spin-up for electron 1, etc.

END OF SIDE NOTE

Symmetric: Similarly, we'll have three (triplet) states that are completely symmetric states

$$\text{Triplet states: } |\alpha\alpha\rangle, \frac{|\alpha\beta\rangle + |\beta\alpha\rangle}{\sqrt{2}}, |\beta\beta\rangle$$

It is trivial to see that if we exchange particles 1 and 2 these states don't change, i.e., they are symmetric w.r.t particle exchange. Let's see on the middle one:

$$\text{Exchange} \left(\frac{|\alpha\beta\rangle + |\beta\alpha\rangle}{\sqrt{2}} \right) = \frac{|\beta\alpha\rangle + |\alpha\beta\rangle}{\sqrt{2}} = \left(\frac{|\alpha\beta\rangle + |\beta\alpha\rangle}{\sqrt{2}} \right) \quad Q.E.D.$$

Ground and Excited Helium wavefunction:

Now that we have the anti-symmetric spin combinations, we see that we can write the w.f. **of the full 2-electron system** as (again, not exactly but to a good approximation)

$$\Psi_{\text{He,ground}} = \phi_{1s}(\mathbf{r}_1)\phi_{1s}(\mathbf{r}_2) \frac{|\alpha\beta\rangle - |\beta\alpha\rangle}{\sqrt{2}}$$

(where the 1s function is essentially what we discussed for H, although it may look a little different, a little compressed/stretched – we'll discuss that later).

What about the excited states?

Let's excite one electron from 1s→2s. (We won't worry now about excitations from 1s→2p, although those are similar)

I.e., before symmetrization we'll want something like:

$$\Psi_{\text{He}}^* (\text{not properly anti-symmetric}) = \phi_{1s}(\mathbf{r}_1)\phi_{2s}(\mathbf{r}_2) |\alpha\beta\rangle$$

Or equally possibly

$$\Psi_{\text{He}}^* (\text{not properly anti-symmetric}) = \phi_{1s}(\mathbf{r}_1)\phi_{2s}(\mathbf{r}_2) |\alpha\alpha\rangle$$

etc.

So let's write possible combinations which the Pauli principle allows.

Let's consider combinations that are symmetric in the spin and antisymmetric in the spatial parts, or vice versa (in principle we don't have to have both the spin and spatial part symmetric/antisymmetric, only the full w.f., but we won't discuss that now).

The combinations are:

$$\text{Singlet excited state: } \frac{\phi_{1s}(\mathbf{r}_1)\phi_{2s}(\mathbf{r}_2) + \phi_{2s}(\mathbf{r}_1)\phi_{1s}(\mathbf{r}_2)}{\sqrt{2}} \frac{|\alpha\beta\rangle - |\beta\alpha\rangle}{\sqrt{2}}$$

And the triplet excited states:

Triplet excited states:

$$i) \frac{\phi_{1s}(\mathbf{r}_1)\phi_{2s}(\mathbf{r}_2) - \phi_{2s}(\mathbf{r}_1)\phi_{1s}(\mathbf{r}_2)}{\sqrt{2}} |\alpha\alpha\rangle$$

$$ii) \frac{\phi_{1s}(\mathbf{r}_1)\phi_{2s}(\mathbf{r}_2) - \phi_{2s}(\mathbf{r}_1)\phi_{1s}(\mathbf{r}_2)}{\sqrt{2}} \frac{|\alpha\beta\rangle + |\beta\alpha\rangle}{\sqrt{2}}$$

$$iii) \frac{\phi_{1s}(\mathbf{r}_1)\phi_{2s}(\mathbf{r}_2) - \phi_{2s}(\mathbf{r}_1)\phi_{1s}(\mathbf{r}_2)}{\sqrt{2}} |\beta\beta\rangle$$

We see that:

All fulfill the Pauli principle; if we exchange electrons 1 and 2 then the spatial part will either change sign (in the triplet) and the spin part will not, or vice versa (in the singlet).

Turns out that the triplet states are degenerate (have the same energy). Makes sense, since they have the same spatial part, and are only different by the spin functions.

Turns out that spin part of them all have the same spin ("1") and different spin along the z-axis (-1, 0, and 1).

Which one will have a lower (more negative) energy? Definitely the triplet. Hund's rule tells us that, and now we can understand why.

Physically, electrons will want to be away from each other if possible, due to electron-electron repulsion.

In the triplet state, the electrons are naturally away from each other. **i.e.**, in the triplet state the amplitude for the electrons to be at the same point is zero:

$$\Psi = 0 \quad \text{if} \quad \mathbf{r}_1 = \mathbf{r}_2$$

That's because the spatial part is antisymmetric, so

$$\mathbf{r}_1 = \mathbf{r}_2 \quad \rightarrow \quad \frac{\phi_{1s}(\mathbf{r}_1)\phi_{2s}(\mathbf{r}_1) - \phi_{2s}(\mathbf{r}_1)\phi_{1s}(\mathbf{r}_1)}{\sqrt{2}} = 0$$

This symmetry explains Hund's rules too. Basically, if electrons can be in two orbitals with similar energies (e.g., 2s and 2px, or 2px and 2pz, etc.)

they'll prefer the spatially antisymmetric combination, which means their spins will be in a triplet combination (what we think of colloquially as electrons having parallel spins, but that's not accurate – for example, $\frac{|\alpha\beta\rangle+|\beta\alpha\rangle}{\sqrt{2}}$ is a symmetric triplet eigenstate with opposite spins).

Spin of the singlet and triplet states.

Let's prove what we asserted, that the singlet state has a total spin of 0 (i.e., the two $s=1/2$ electron spins combine to give a total spin of 0) using similar technique we could prove that the triplet has spin 1.

Apply S_x, S_y, S_z on the singlet state, and start with S_z .

$$S_z = S_{I,z} + S_{II,z}$$

Where the "I" and "II" refer to electron one and two.

So the singlet is $\frac{|\alpha\beta\rangle-|\beta\alpha\rangle}{\sqrt{2}}$ and therefore

$$S_z \left(\frac{|\alpha\beta\rangle - |\beta\alpha\rangle}{\sqrt{2}} \right) = (S_{I,z} + S_{II,z}) \left(\frac{|\alpha\beta\rangle - |\beta\alpha\rangle}{\sqrt{2}} \right)$$

Also.:

$$S_{I,z}|\alpha\beta\rangle = \frac{1}{2}|\alpha\beta\rangle \quad \text{since the spin of electron I is up in this w. f.}$$

$$S_{II,z}|\alpha\beta\rangle = -\frac{1}{2}|\alpha\beta\rangle \quad \text{since the spin of electron II is down in this w. f.}$$

Therefore

$$S_z|\alpha\beta\rangle = (S_{I,z} + S_{II,z})|\alpha\beta\rangle = \frac{1}{2}|\alpha\beta\rangle - \frac{1}{2}|\alpha\beta\rangle = 0$$

And we can prove similarly that $S_z|\beta\alpha\rangle=0$, so $S_z \left(\frac{|\alpha\beta\rangle-|\beta\alpha\rangle}{\sqrt{2}} \right) = 0$.

What about S_x and S_y – similarly, a little more involved and interesting.

Recall $S_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ so for one particle $S_x|\alpha\rangle = \frac{1}{2}|\beta\rangle$, $S_x|\beta\rangle = \frac{1}{2}|\alpha\rangle$ so if we act with the spin of the 1st particle on a wavefunction of two particles we get

$$S_{I,x}|\alpha\beta\rangle = \frac{1}{2}|\beta\beta\rangle$$

And similarly when acting the with the spin the 2nd particle

$$S_{II,x}|\alpha\beta\rangle = \frac{1}{2}|\alpha\alpha\rangle$$

And thus

$$S_x|\alpha\beta\rangle = S_{I,x}|\alpha\beta\rangle + S_{II,x}|\alpha\beta\rangle = \frac{1}{2}|\beta\beta\rangle + \frac{1}{2}|\alpha\alpha\rangle$$

Similarly

$$S_x|\beta\alpha\rangle = \frac{1}{2}|\alpha\alpha\rangle + \frac{1}{2}|\beta\beta\rangle$$

Therefore

$$S_x * \text{singlet_wf} = S_x \frac{(|\alpha\beta\rangle - |\beta\alpha\rangle)}{\sqrt{2}} = \frac{1}{2\sqrt{2}} (|\beta\beta\rangle + |\alpha\alpha\rangle - |\alpha\alpha\rangle - |\beta\beta\rangle) = 0$$

Similarly,

$$S_y * \text{singlet_wf} = 0$$

Therefore

$$S^2 \frac{(|\alpha\beta\rangle - |\beta\alpha\rangle)}{\sqrt{2}} = (S_x^2 + S_y^2 + S_z^2) \frac{(|\alpha\beta\rangle - |\beta\alpha\rangle)}{\sqrt{2}} = 0^2 + 0^2 + 0^2 = 0$$

i.e.,

$$S^2 \frac{(|\alpha\beta\rangle - |\beta\alpha\rangle)}{\sqrt{2}} = 0 * \frac{(|\alpha\beta\rangle - |\beta\alpha\rangle)}{\sqrt{2}}$$

So the singlet has zero-spin indeed!

And similarly when we act on any of the three triplet functions we could show that S^2 on them gives $1*(1+2) = 2$ times them, so $s_{tot} = 1$ (s_{tot} is the equivalent of l , i.e., the eigenvalue of S^2 is $s_{tot}(s_{tot} + 1)$.)

Spin Orbit

The last topic relating to spin is spin-orbit. Basically, the spin-orbit interaction is a relativistic effect so it is pronounced in heavier atoms where the core electrons are close to nucleus and have very high velocities which reach a significant fraction of the velocity of light (so they are relativistic)

The spin-orbit interaction is, as mentioned, of the general form

$$H_{S.O.} = \text{const.} \cdot f(r) \mathbf{L} \cdot \mathbf{S}$$

where $f(r)$ is some function of r and radial derivatives which we won't worry about (that is near the origin, i.e., is stronger for heavier nuclei).

The interesting part is $\mathbf{L} \cdot \mathbf{S}$. **The reason is that it mixes spin up and spin down.** (i.e., it has terms which are not just proportional to S^2 nor just to S_z , as in the interaction with magnetic field; instead it mixes spin up with spin down.

Side note: total angular momentum

Let's define the total angular momentum

$$\mathbf{J} = \mathbf{L} + \mathbf{S}$$

Now the total angular momentum eigenstates will be, as we mentioned earlier, combination of eigenstates of \mathbf{L} and \mathbf{S} .

Let's talk about transition metals with $3d$ states (i.e., $l = 2$)

Couple these $l = 2$ states with $s = 0.5$ from spin; as we learned without proof, from the triangle rule we can get states with

$$j = 2 - 0.5, 2 + 0.5$$

i.e., states with

$$j = 1.5, 2.5$$

Since $J_z = L_z + S_z$ it follows that the eigenvalue of J_z , which we label as m_j will equal $m_l + m_s$ i.e., the z-component of the angular momentum of the orbital motion and of the spin.

For example, the state with $j = 1.5$, $m_j = j = 1.5$ will be a combination of two components, one with $l_z = 2$ and $m_s = -0.5$, and the other with $l_z = 1$ and $m_s = 0.5$ the coefficients can be calculated and the final result is:

$$\Psi_{j=1.5, m_j=1.5, s=\frac{1}{2}, l=2}(\theta, \phi, spin) \simeq 0.89 * Y_{m=2, l=2}(\theta, \phi)|\beta\rangle - 0.45 * Y_{m=1, l=2}(\theta, \phi)|\alpha\rangle$$

Note that the first state on the RHS has $m_l = 2, m_s = -\frac{1}{2}$,

so $m_j = 2 - \frac{1}{2} = 1.5$,

and the second state has the same m_j since $m_l = 1, m_s = +\frac{1}{2}$.

So both the states in green in the RHS contribute to the total state on the left (i.e., they contribute to the eigenfunction of the combined angular momentum operator with $m_j = 1.5$).

It is not surprising to learn that the other combination of these two states will give $j = 2.5$, i.e.,

$$\Psi_{j=2.5, m_j=1.5, s=\frac{1}{2}, l=2}(\theta, \phi, spin) \simeq 0.45 * Y_{m=2, l=2}(\theta, \phi)|\beta\rangle + 0.89 * Y_{m=1, l=2}(\theta, \phi)|\alpha\rangle$$

Now back to the spin orbit.

What we'll show now is that these total angular momentum combination (in green above) which I wrote are eigenstates of the spin-orbit interaction.

I.e., the spin orbit interaction mixes states with different L

Now notice that for numbers, $ab = (a + b)^2 - a^2 - b^2$. This is also true here, i.e.

$$\mathbf{L} \cdot \mathbf{S} = \frac{1}{2}(\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2)$$

(The reason is that \mathbf{L} and \mathbf{S} commute – unlike \mathbf{L} with \mathbf{L}).

That simplifies the spin orbit interaction. When acting on the top green state, with $j = 1.5$, it follows that acting with $L \cdot S$ is the same as acting with

$$\frac{1}{2}(j(j+1) - l(l+1) - s(s+1)) = \frac{1}{2}\left(1.5 * 2.5 - 2 * 3 - \frac{1}{2} * \frac{3}{2}\right) = -1.5$$

While when acting on the bottom green state we get that acting with $L \cdot S$ is the same as acting with

$$\frac{1}{2}(j(j+1) - l(l+1) - s(s+1)) = \frac{1}{2}\left(2.5 * 3.5 - 2 * 3 - \frac{1}{2} * \frac{3}{2}\right) = 2$$

Therefore, the spin orbit interaction has a different sign and magnitude depending on which state it acts!

This naturally leads to splitting between states with the same l but different j . We'll get back to splitting when we talk about perturbation theory.

He Atom

At the H atom⁵ we are at the end of our ability to solve analytically the quantum mechanics of an atom.

The problem is that for any species with 2 or more electrons, we have to include the $e^- - e^-$ repulsive interaction in V .

This precludes an analytical solution to the time-independent Schrödinger Equation.

Specifically, the Hamiltonian for the 2 electrons in He is (I use the hat on the Hamiltonian operator so you don't confuse H for Hamiltonian with H for Hydrogen).

$$\hat{H} = \left(-\frac{\hbar^2}{2M}\Delta_1 - \frac{\hbar^2}{2M}\Delta_2 - \frac{2e^2}{r_1} - \frac{2e^2}{r_2} \right) + \frac{e^2}{r_{12}}$$

where the **2** is due to the interaction of each electron with the $Z = 2$ charge of the He nucleus

Note that $-\frac{\hbar^2}{2m_e}\Delta_1$ is the kinetic energy of electron #1, and similarly for electron #2.

We could rewrite the Hamiltonian as

$$\hat{H} = \hat{H}_1 + \hat{H}_2 + \frac{e^2}{r_{12}}$$

where

$$\hat{H}_1 = -\frac{\hbar^2}{2M}\Delta_1 - \frac{2e^2}{r_1}$$

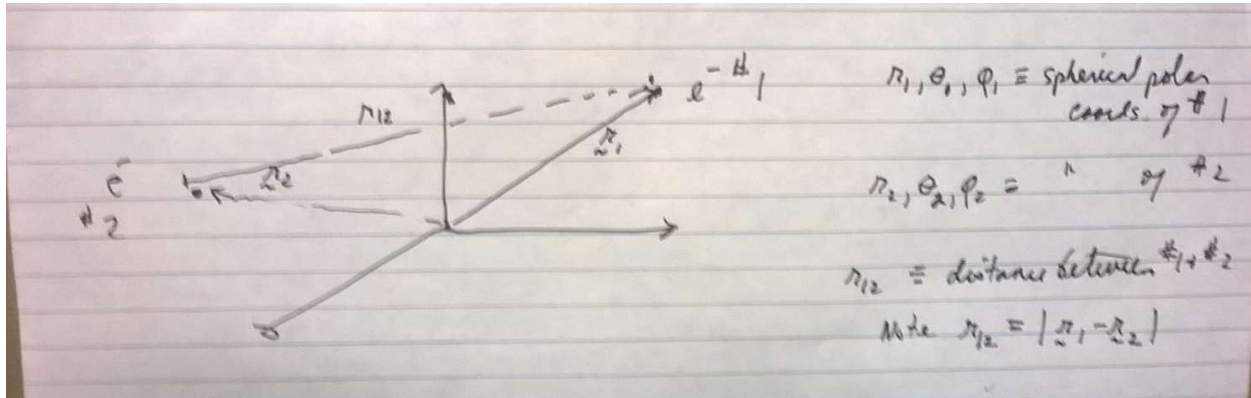
$$\hat{H}_2 = -\frac{\hbar^2}{2M}\Delta_2 - \frac{2e^2}{r_2}$$

⁵ Note: this part and the next few parts are copied almost verbatim from notes of Prof. Peter Felker.

Note that for each electron "j" (where $j = 1, 2$ is an index over the electrons)

$$-\frac{\hbar^2}{2M}\Delta_j = -\frac{\hbar^2}{2M}\frac{1}{r_j}\frac{\partial^2}{\partial r_j^2} + \frac{L_j^2}{2Mr_j^2}$$

See picture below:



Note that if the e-e repulsion was not present, the He atom would be simple to solve since the Hamiltonian would then be the sum of two one-electron-atom Hamiltonian's.

We could use separation of variables and obtain the He atom solution for H_0 (where the e's don't interact) in terms of an H-like state:

$$(\hat{H}_1 + \hat{H}_2)\Phi = \left[\left(-\frac{\hbar^2}{2M}\Delta_1 - \frac{2e^2}{r_1} \right) + \left(-\frac{\hbar^2}{2M}\Delta_2 - \frac{2e^2}{r_2} \right) \right] \Phi = E\Phi$$

Again postulate a separable form for Φ (i.e., hope you can find a solution in the form:)

$$\Phi = \phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2)$$

Then

$$(\hat{H}_1 + \hat{H}_2)\Phi = [(K_1 + V_1) + (K_2 + V_2)]\phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2) = E\phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2)$$

Divide by $\phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2)$, to get, when you move one term to the right

$$\frac{\left(-\frac{\hbar^2}{2M}\Delta_1 - \frac{2e^2}{r_1} \right) \phi_1(\mathbf{r}_1)}{\phi_1(\mathbf{r}_1)} = E - \frac{\left(-\frac{\hbar^2}{2M}\Delta_2 - \frac{2e^2}{r_2} \right) \phi_2(\mathbf{r}_2)}{\phi_2(\mathbf{r}_2)}$$

The LHS is only a function of \mathbf{r}_1 , while the RHS is only a function of \mathbf{r}_2 , so both sides must be constant, i.e.,

$$\frac{\left(-\frac{\hbar^2}{2M}\Delta_1 - \frac{2e^2}{r_1}\right)\phi_1(\mathbf{r}_1)}{\phi_1(\mathbf{r}_1)} = E^{(1)}$$

$$\frac{\left(-\frac{\hbar^2}{2M}\Delta_2 - \frac{2e^2}{r_2}\right)\phi_2(\mathbf{r}_2)}{\phi_2(\mathbf{r}_2)} = E^{(2)} \equiv E - E^{(1)}$$

so

$$E^{(1)} + E^{(2)} = E$$

i.e., both ϕ_1 and ϕ_2 are solutions of the Schrodinger equation

$$\left(-\frac{\hbar^2}{2M}\Delta_1 - \frac{2e^2}{r_1}\right)\phi_1(\mathbf{r}_1) = E^{(1)}\phi_1(\mathbf{r}_1)$$

and

$$\left(-\frac{\hbar^2}{2M}\Delta_2 - \frac{2e^2}{r_2}\right)\phi_2(\mathbf{r}_2) = E^{(2)}\phi_2(\mathbf{r}_2)$$

Note that ϕ_1 and ϕ_2 solve a Hydrogen-like Schrodinger equation (for a higher charge nucleus, with $Z=2$, of course)

Unfortunately the e-e repulsion, which involves both \mathbf{r}_1 and \mathbf{r}_2

(via $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$), messes up the wavefunction, which is not really a product state anymore.

Other considerations: Electron spin and Indistinguishability of electrons.

As we've seen in the previous chapter, spin is important even without considering magnetic interaction because of the indistinguishability of electrons and the consequences of the Pauli principle. Specifically:

We'll now explain the issue of symmetry, i.e., the fact that under an exchange of BOTH coordinates and spin the wavefunction of two identical particles must either be antisymmetric (e.g., for electrons) or symmetric

(e.g., in the case of photons). We used it already in the previous chapter, but now we'll explain more.

Label by σ_1 the spin of electron 1, which can be α (spin up) or β (spin down).

A complete description of the coordinate of electron 1 is labeled as

$$q_1 \equiv \mathbf{r}_1, \sigma_1$$

i.e., both the spatial coordinate and the spin coordinate.

Similarly : $q_2 \equiv \mathbf{r}_2, \sigma_2$

We know (calling the full w.f. Ψ) that

$$|\Psi(q_1, q_2)|^2 d\mathbf{r}_1 d\mathbf{r}_2$$

is the probability the find:

electron 1 in a volume of size $d\mathbf{r}_1$ around \mathbf{r}_1 , with spin σ_1 , and
electron 2 in a volume of size $d\mathbf{r}_2$ around \mathbf{r}_2 , with spin σ_2 .

Now because of the indistinguishability of the electrons, we cannot tell electron #1 from electron #2. So if we look at atom and call one electron #1 and the other #2, and then look away, and finally look back again, we have no way of knowing which electron is the one we called #1 and which is #2.

As such, any observable must not be affected by the labeling of the electrons.

In short, $|\Psi(q_1, q_2)|^2$ **must equal** $|\Psi(q_2, q_1)|^2$

i.e.,

$$\Psi(q_1, q_2) = \pm \Psi(q_2, q_1)$$

So Ψ must be symmetric (+) or antisymmetric (-) w/rt the exchange of the coords. of any two e's.

It turns out that for electrons (an example of a class of particles called "fermions") Ψ must be antisymmetric w.r.t. full-coordinate (spin+spatial coordinates) interchange. This is the celebrated Pauli Principle which we mentioned before.

So, we get that for the electrons

$$\Psi(\mathbf{r}_1, \sigma_1, \mathbf{r}_2, \sigma_2) = -\Psi(\mathbf{r}_2, \sigma_2, \mathbf{r}_1, \sigma_1)$$

Or simply (a notation we'll use later):

$$\Psi(1,2) = -\Psi(2,1)$$

Now that we understand the Pauli principle, we recall from the previous chapter that the combination:

$$\Psi(1,2) = \phi_{1s}(r_1)\phi_{1s}(r_2) \frac{(\alpha(1)\beta(2) - \alpha(2)\beta(1))}{\sqrt{2}} = -\Psi(2,1)$$

Side-note: we'll use slightly different notation, no bra-kets for the spin now – but same meaning; a combination: spin up for electron#1 and spin-down for electron #2, with the opposite, spin down for #1 with spin up for #2.

So the spatial part of ψ is symmetric w.r.t. interchange of electron coordinates, and the spin part of ψ is antisymmetric and therefore renders ψ itself antisymmetric w.r.t. coord. Interchange.

The main point is: the dependence of the full w.f. in spatial and spin coordinates must be taken into account in assessing the acceptability of an electronic wavefunction.

Note finally: the form of the w.f. above is not the exact w.f. (due to the electron-electron interaction), even though it obeys the correct symmetry, i.e., $\Psi(1,2) = -\Psi(2,1)$ (which the full w.f. also fulfills). **The form we have above is an approximation.** So we need to learn about approximations and how to use and improve them – next chapters!

Approximation methods:

Two main classes – the variational method; and perturbation theory.

Variational Method

We proved it in class a few weeks ago, but let's repeat. The claim is that if we pick a general wavefunction Φ then the **normalized expectation value of the Hamiltonian will be greater than or equal to the ground-state (lowest energy)**.

BTW, we sometimes denote the ground state as E_1 but we'll follow the convention we used for the Harmonic oscillator and denote it as E_0 , i.e., enumerate the eigenstates and eigenvalues of H as ϕ_n, E_n with $n = 0, 1, 2, \dots$

So the claim is that if we define the average energy associated with a w.f. as

$$\langle \hat{H} \rangle \equiv (\text{definition}) \frac{\langle \Phi | \hat{H} | \Phi \rangle}{\langle \Phi | \Phi \rangle}$$

(and I remind you that $\langle \Phi | \hat{H} | \Phi \rangle$ means $\int \Phi^* \hat{H} \Phi dx$ in 1D, and similarly in more dimensions).

Then

$$\langle \hat{H} \rangle \geq E_0 \quad (\text{where } E_0 \text{ is the exact ground state energy}).$$

Proof: expand Φ in terms of the exact eigenstates of \hat{H} :

$$\Phi = \sum_{n=0}^{\infty} a_n \phi_n$$

Then

$$\langle \hat{H} \rangle = \frac{\langle \Phi | \hat{H} | \Phi \rangle}{\langle \Phi | \Phi \rangle} = \frac{\int (\sum_n a_n^* \phi_n^*) H (\sum_j a_j \phi_j) dx}{\int (\sum_n a_n^* \phi_n^*) (\sum_j a_j \phi_j) dx}$$

i.e. (highlighting important terms)

$$\langle \hat{H} \rangle = \frac{\langle \Phi | \hat{H} | \Phi \rangle}{\langle \Phi | \Phi \rangle} = \frac{\sum_{jn} a_n^* a_j \int \phi_n^* \mathbf{H} \phi_j dx}{\sum_{jn} a_n^* a_j \int \phi_n^* \phi_j dx} = \frac{\sum_{jn} a_n^* a_j \mathbf{E}_j \int \phi_n^* \phi_j dx}{\sum_{jn} a_n^* a_j \int \phi_n^* \phi_j dx}$$

But

$$\int \phi_n^* \phi_j dx = \delta_{jn} = 1 \text{ if } j = n, 0 \text{ otherwise}$$

So the delta function means we can replace the sum over n by picking $n = j$, so

$$\langle \hat{H} \rangle = \frac{\sum_j a_j^* a_j \mathbf{E}_j}{\sum_j a_j^* a_j} = \frac{\sum_j |a_j|^2 \mathbf{E}_j}{\sum_j |a_j|^2}$$

But

$$E_j \geq E_0$$

So

$$\langle \hat{H} \rangle = \frac{\sum_j |a_j|^2 \mathbf{E}_j}{\sum_j |a_j|^2} \geq \frac{\sum_j |a_j|^2 \mathbf{E}_0}{\sum_j |a_j|^2} = \mathbf{E}_0$$

i.e.,

$$\langle \hat{H} \rangle \geq E_0$$

There will be equality of course only if the wavefunction has no contribution (i.e., $a_j = 0$) for those states j with energy higher than the ground-state energy.

The power of the variational method (as formulated here) is in determining the ground state of a system.

Specifically, we are guaranteed that any choice of Φ will give rise to $\langle \hat{H} \rangle \geq E_0$. Hence, we make provision to vary Φ in such a way to minimize $\langle \hat{H} \rangle$.

The lower we can make $\langle \hat{H} \rangle$ by such variation, the better the approximation we have for the ground-state energy and the ground state eigenfunction.

One way to accomplish this variation is to build into the trial wavefunction, Φ , one or more adjustable parameters(s) The parameters can then be systematically adjusted to minimize $\langle \hat{H} \rangle$ by usual Calculus minimization.

Example: particle in an infinite-depth box of size L (a problem we solved earlier)

For simplicity we'll assume $L = 1$ in the following.

We know of course that the true wavefunction is

$$\text{const.} * \sin\left(\frac{\pi x}{L}\right) = \text{const.} * \sin(\pi x)$$

But say we did not know that. Then say we take a function which is symmetric around $\frac{L}{2} = \frac{1}{2}$ (the potential is symmetric around the midpoint, so it makes sense and can be proved that the ground-state w.f. is symmetric around the midpoint).

The function also needs to be 0 at $x = 0$, and $x = L = 1$.

One such function will be

$$x(1 - x)$$

Say we wanted to do even better, so we can take that function in a combination with its square

$$\Phi(x) = x(1 - x) + \beta (x(1 - x))^2$$

Let's find what the best β , i.e., the one that minimizes the expectation value of the Hamiltonian.

$$\frac{\langle \Phi | \hat{H} | \Phi \rangle}{\langle \Phi | \Phi \rangle} = ?$$

Well, first the denominator

$$\begin{aligned} \langle \Phi | \Phi \rangle &= \int_0^1 \left(x(1 - x) + \beta (x(1 - x))^2 \right)^2 dx \\ &= \int_0^1 x^2(1 - x)^2 dx + 2\beta \int_0^1 x^3(1 - x)^3 dx + \beta^2 \int_0^1 x^4(1 - x)^4 dx \end{aligned}$$

Here, all integrals are from 0 to 1 (i.e., to L) since the w.f. cannot exist outside that region.

We could do the integrals analytically, but I'll shortcut with Wolframalpha, to get

$$\int_0^1 x^2(1-x)^2 dx = \frac{1}{30}$$

$$2 \int_0^1 x^3(1-x)^3 dx = \frac{1}{70}$$

$$\int_0^1 x^4(1-x)^4 dx = \frac{1}{630}$$

i.e.,

$$\langle \Phi | \Phi \rangle = \frac{1}{30} + \frac{\beta}{70} + \frac{\beta^2}{630} = \frac{1}{630} (\beta^2 + 9\beta + 21)$$

Similarly,

$$\langle \Phi | \hat{H} | \Phi \rangle = -\frac{\hbar^2}{2M} \int_0^1 \Phi(x) \frac{d^2 \Phi(x)}{dx^2} dx = \frac{\hbar^2}{2M} \int_0^1 \frac{d\Phi(x)}{dx} \frac{d\Phi(x)}{dx} dx$$

But

$$\frac{d\Phi(x)}{dx} = \frac{d}{dx} \left(x(1-x) + \beta (x(1-x))^2 \right) = (1-2x)(1+2\beta(x-x^2))$$

So (with $\hbar = M = 1$)

$$2 \langle \Phi | \hat{H} | \Phi \rangle = \int_0^1 (1-2x)^2 (1+4\beta(x-x^2) + 4\beta^2(x-x^2)^2) dx =$$

$$\int_0^1 (1-2x)^2 dx + 4\beta \int_0^1 (1-2x)^2 (x-x^2) dx + 4\beta^2 \int_0^1 (1-2x)^2 (x-x^2)^2 dx$$

But

$$\int_0^1 (1-2x)^2 dx = \frac{1}{3}$$

$$4 \int_0^1 (1-2x)^2 x(1-x) dx = \frac{2}{15}$$

$$4 \int_0^1 (x(1-x)(1-2x))^2 dx = \frac{2}{105}$$

So

$$2\langle\Phi|\hat{H}|\Phi\rangle = \frac{1}{3} + \frac{2\beta}{15} + \frac{2\beta^2}{105} = \frac{1}{105}(2\beta^2 + 14\beta + 35)$$

i.e., the expectation value is

$$\langle\hat{H}\rangle \equiv \frac{\langle\Phi|\hat{H}|\Phi\rangle}{\langle\Phi|\Phi\rangle} = \frac{1}{2} * \frac{630}{105} * \frac{2\beta^2 + 14\beta + 35}{\beta^2 + 9\beta + 21} = 3 \frac{2\beta^2 + 14\beta + 35}{\beta^2 + 9\beta + 21}$$

So we're looking for the value of the parameter that minimizes the expression in red.

We would solve by differentiation

$$\frac{d}{d\beta} \left(\frac{2\beta^2 + 14\beta + 35}{\beta^2 + 9\beta + 21} \right) = 0$$

But if $\frac{d(f/g)}{d\beta} = 0$ then $0 = \frac{g \frac{df}{d\beta} - f \frac{dg}{d\beta}}{g^2}$ i.e., $g \frac{df}{d\beta} = f \frac{dg}{d\beta}$. Thus in our case

$$(\beta^2 + 9\beta + 21) \frac{d}{d\beta} (2\beta^2 + 14\beta + 35) = (2\beta^2 + 14\beta + 35) \frac{d}{d\beta} (\beta^2 + 9\beta + 21)$$

i.e.,

$$(\beta^2 + 9\beta + 21)(4\beta + 14) = (2\beta^2 + 14\beta + 35)(2\beta + 9)$$

Expand to get

$$21 * 14 + (21 * 4 + 14 * 9)\beta + (4 * 9 + 14)\beta^2 + 2\beta^3 =$$

$$35 * 9 + (35 * 2 + 14 * 9)\beta + (2 * 9 + 2 * 14)\beta^2 + 2\beta^3$$

Luckily the red terms, cubical in β , fall away, as do the blue terms, so after subtracting the RHS from the LHS we're left with a 2nd order equation

$$a\beta^2 + b\beta + c = 0$$

where

$$c = 35 * 9 - 21 * 14 = 21$$

$$b = 70 - 84 = -14$$

$$a = -2 * 9 + 14 = -4$$

And using that the solutions are $\beta_{1,2} = -\frac{b}{2a} \pm \frac{1}{2a}\sqrt{b^2 - 4ac}$ we get

$$\beta_1 = -\frac{28}{8} + \frac{\sqrt{14^2 + 4 * 4 * 21}}{8} = -\frac{7}{4} + \frac{\sqrt{133}}{4} = 1.13314$$

$$\beta_2 = -\frac{7}{4} - \frac{\sqrt{133}}{4} = -4.63314$$

Let's evaluate to see which solution gives lower energy; I remind you that

$$\langle \hat{H} \rangle = 3 \frac{2\beta^2 + 14\beta + 35}{\beta^2 + 9\beta + 21}$$

Therefore, substituting we get

for $\beta = \beta_1 = 1.13314$, $\langle \hat{H} \rangle = 4.93487$

for β_2 : $\langle \hat{H} \rangle = 51.6$. Obviously β_2 is a maximum, not a minimum.

How are we doing?

Well, if we did not have β and tried just a function of the form

$x(1-x)$, i.e., $\beta = 0$, we would get $\langle \hat{H} \rangle = 3 \frac{2\beta^2 + 14\beta + 35}{\beta^2 + 9\beta + 21} = 3 * \frac{35}{21} = 5.0$

The true solution is of course

$$E_{ground} = \hbar^2 \frac{\pi^2}{2ML^2} = \frac{\pi^2}{2} = 4.93480$$

Note the impressive improvement! By playing with one parameter we got the initial error, (i.e., the error for $\beta = 0$, which was $5 - E(\text{ground}) = 0.064$, i.e., 1.2% of the initial value), all the down to 0.00007, i.e., tiny!

The Variational Method and Application to the He atom.⁶

As we wrote, the variational principle amounts to:

- Writing the wavefunction as dependent on parameters (the more the better accuracy wise, but the more effort this makes).
- Optimizing the parameter(s) to give the lowest $\langle \hat{H} \rangle$.

Let's apply the variational method to the He atom, which, recall, has (where we use again a.u. , so $e^2 = 1, M = 1, \hbar = 1$)

$$\hat{H} = \left(-\frac{1}{2}\Delta_1 - \frac{1}{2}\Delta_2 - \frac{2}{r_1} - \frac{2}{r_2} \right) + \frac{1}{r_{12}}$$

We explained before that for the ground-state the spin-part of the wavefunction will be antisymmetric, so the spatial part is symmetric. Since we ignore spin orbit effects the spin part drops out of the calculation and we can just postulate a spatial wavefunction.

Let's choose

$$\Phi = \frac{\xi^3}{\pi} \exp(-\xi r_1) \exp(-\xi r_2)$$

This function is a product of two hydrogen-like 1s wavefunctions for a nuclear charge equal to ξe .

I.e., recall that we said (or should have said) that the 1s ground-state for a single-electron system with a charge Z is

$$\phi_{1s} = \text{const.} * \exp\left(-\frac{Zr}{a_0}\right)$$

Here we set $a_0 = 1$ and instead of Z we use ξ .

As written the wavefunction Φ is always normalized (i.e., $\langle \Phi | \Phi \rangle = 1$. **prove that!!**) and it obeys the boundary conditions, i.e., it is finite everywhere and goes to 0 as either of two electrons goes to infinity.

⁶ Copied from Peter Felker's notes.

We'll treat ξ as an adjustable parameter to minimize $\langle \hat{H} \rangle$.

To proceed we'll do a trick. Rewrite the Hamiltonian as a zero-order part associated with a charge ξ

$$\hat{H} = \hat{H}_0 + \hat{H}'$$

where

$$\hat{H}_0 = \left[-\frac{1}{2} \Delta_1 - \frac{\xi}{r_1} \right] + \left[-\frac{1}{2} \Delta_2 - \frac{\xi}{r_2} \right]$$

and the "extra"

$$\hat{H}' = \frac{\xi - 2}{r_1} + \frac{\xi - 2}{r_2} + \frac{1}{r_{12}}$$

Since the w.f. is normalized Φ then

$$\langle \hat{H} \rangle = \langle \Phi | \hat{H} | \Phi \rangle = \langle \Phi | \hat{H}_0 | \Phi \rangle + \langle \Phi | \hat{H}' | \Phi \rangle$$

The first term on the RHS, $\langle \Phi | \hat{H}_0 | \Phi \rangle$, is easily calculated once we realize that the first term in \hat{H}_0 is identical to the Hamiltonian of a one-electron atom in which the electron interacts with a nucleus of charge ξe .

The second term in brackets is also of this form.

Since Φ is a product of two 1s orbitals for such hydrogen like atoms, Φ is an eigenfunction of \hat{H}_0 with eigenvalue equal to

$$W_0 = 2 * \left(-\frac{\xi^2 e^2}{2a_0} \right) = -\xi^2 * \text{Hartree}$$

Here:

- The two is due to having two terms (one for each electron)
- The ξ^2 is because the energy is proportional to the square of the charge of the nucleus; to see that quickly realize that if we double the charge of the nucleus, then the electron will be twice as close to the nucleus in average, and the charge to which the electron is attracted will be twice as large, resulting in an increase of the energy by $2^2 = 4$.

Next we need to find $\langle \Phi | \hat{H}' | \Phi \rangle$. Plugging in for Φ and for \hat{H}' gives

$$\langle \Phi | \hat{H}' | \Phi \rangle = \left[\frac{1}{\pi^2} \xi^6 \right] \iint r_1^2 dr_1 \sin(\theta_1) d\theta_1 r_2^2 dr_2 \sin(\theta_2) d\theta_2 * \exp(-2\xi r_1) \exp(-2\xi r_2) * \left(\frac{\xi - 2}{r_1} + \frac{\xi - 2}{r_2} + \frac{1}{r_{12}} \right)$$

Where the “blue” term is the volume element for the two integrals; the gray term come from $|\Phi|^2$, and the last term (the parenthesis) is coming from \hat{H}' .

The integrals can be simplified; each of the red terms really involves only one electron and can be shown to be $\xi(\xi - 2) \frac{e^2}{a_0}$, which we multiply by 2 since there are two such terms.

The last integral (over $\frac{1}{r_{12}}$) is painful but can be shown to be $\xi \frac{5e^2}{8a_0}$.

So the combined result for $\langle \Phi | \hat{H}' | \Phi \rangle$ is

$$\langle \Phi | \hat{H}' | \Phi \rangle = 2 * \xi(\xi - 2) \frac{e^2}{a_0} + \xi \frac{5e^2}{8a_0} = \left(2\xi^2 - 4\xi + \frac{5}{8}\xi \right) \frac{e^2}{a_0} = 2\xi^2 - 4\xi + \frac{5}{8}\xi$$

where the last equality is since we use a.u.

And the total energy is, as written earlier,

$$\langle \hat{H} \rangle = \langle \Phi | \hat{H} | \Phi \rangle = \langle \Phi | \hat{H}_0 | \Phi \rangle + \langle \Phi | \hat{H}' | \Phi \rangle = -\xi^2 + 2\xi^2 - 4\xi + \frac{5}{8}\xi$$

i.e.,

$$\langle \hat{H} \rangle = \xi^2 - 4\xi + \frac{5}{8}\xi$$

i.e.,

$$\langle \hat{H} \rangle = \xi^2 - \frac{27}{8}\xi$$

So to minimize the energy by the variational principle we differentiate,

$$0 = \frac{d}{d\xi} \langle \hat{H} \rangle = \frac{d}{d\xi} \left(\xi^2 - \frac{27}{8} \xi \right) \quad \text{for } \xi_{\text{optimal}}$$

i.e.,

$$2\xi_{\text{optimal}} - \frac{27}{8} = 0$$

i.e.,

$$\xi_{\text{optimal}} = \frac{27}{16} \approx 1.69$$

Energies:

Let's look at the energies.

If we were to employ in Φ (the turquoise equation) a value of $\xi = 1$, i.e., the product of 1s-functions of a hydrogen atom, then

$$\langle \hat{H}(\xi = 1) \rangle = \xi^2 - \frac{27}{8} \xi + \frac{5}{8} \xi = 1^2 - \frac{27}{8} * 1 = -2\frac{3}{8} \text{ Hartree}$$

But **1 Hartree = 27.2114eV** so

$$\langle \hat{H}(\xi = 1) \rangle = -64.627\text{eV}$$

Conversely, If we were to employ in Φ a value of $\xi = 2$, i.e., the product of 1s-functions associated each with a He^+ atom (i.e., one electron near a nucleus with a charge of 2 protons), then

$$\langle \hat{H}(\xi = 2) \rangle = 2^2 - \frac{27}{8} * 2 = -2\frac{3}{4} \text{ Hartree} = -74.83\text{eV}$$

This is a lower (and therefore better) energy. Now let's check the value of the energy at the optimal value

$$\left\langle \hat{H} \left(\xi = \frac{27}{16} \right) \right\rangle = \left(\frac{27}{16} \right)^2 - \frac{27}{8} \frac{27}{16} = -2.848 \text{ Hartree} = \mathbf{-77.489\text{eV}}$$

Comparing this last energy with the true ground-state energy of He, which is $E_{ground}(exact) = -79.01 \text{ eV}$, one sees that we have done pretty well with a relatively inflexible trial wavefunction (only one adjustable parameter).

Note that the best value for ξ is between 1 and 2. Since the way in which ξ appears in our trial wavefunction leads one to interpret it as an *effective nuclear charge*, we conclude that the effective nuclear charge that an electron “feels” in helium is between that of a proton and that of proton. We interpret this as arising from the shielding of each electron from the nucleus by the other electron. Complete shielding would give rise to $\xi = 1$, no shielding to $\xi = 2$. The intermediate value of $\xi_{optimal} = 1.69$ indicates a partial shielding between these two extremes.

It is also instructive to compare our results with that obtained from the He-atom Schrödinger equation when electron repulsion is completely neglected. Then the ground-state energy is the sum of the ground-state energies of two one-electron atoms with nuclear charge equal to $2e$, i.e.

$$E(\text{no } e - e \text{ repulsion}) = 2(\text{i. e., two electrons}) * \left(-\frac{2^2}{2} \text{ Hartree} \right) = -4 \text{ Hartree} \\ = -108.846 \text{ eV}$$

(Note that the 2^2 in the parentheses is since the energy of an electron near a proton with a charge Z is proportional to Z^2 , as the Bohr model showed and as we mentioned above).

Correlation effect:

It is interesting to ask what is the origin of the difference between the true result (-79.01eV) and the value we found (-77.489eV).

In principle, if we used a more flexible function than a damped exponential, i.e., used a function of the form

$$\Phi = \phi(r_1)\phi(r_2)$$

but did not constrain the wavefunction to be the damped exponential, then we could have got an energy lower than what we found (i.e., lower than -77.489 but of course higher than the true ground-state energy).

Such a solution (i.e., ϕ) is called the Hartree-Fock, HF, (or in this case just the Hartree) molecular orbital, and the associated energy associated with

Φ , i.e., the expectation value of the energy with the “best” product form, is called the Hartree-Fock energy.

Well, as a matter of pure coincidence it turns out that the ϕ we had, i.e., $\phi(r) \propto \exp(-\xi r)$ is in this case almost the exact best form, i.e., almost equals to the Hartree Fock orbital. So the Hartree Fock energy is only very very slightly lower than -77.489 eV.

So what about the rest of the energy? I.e., why is the exact energy of Helium lower by an amount, labeled $E_{correlation}$ (or just E_c), which equals here

$$E_c(\text{He}) = E_{\text{exact}}(\text{He}) - E_{\text{HF}}(\text{He}) = -79.01 - (-77.489) = -1.52 \text{ eV} ?$$

The reason for the correlation energy is simple: In the exact wavefunction $\Psi_{\text{exact}}(\mathbf{r}_1, \mathbf{r}_2)$, the regions where r_{12} is small have very low probability. Basically, each electron avoids the other.

Think of a classical analogue – when electron 1 is to the north of the nucleus, electron 2 will like to be south of the nucleus, when 1 is to the east of the nucleus, 2 is to the west, etc.

This type of correlation is not captured in the Hartree Fock solution, but is important for chemistry. Much of the research at present on simulating molecules in chemistry amounts to developing numerically feasible correlation approaches.

Perturbation theory: read Prof. Felker Notes.

Perturbation theory applied to the He atom:

Let's apply perturbation theory to He, especially the ground state.

Zero-Order Hamiltonian:

Recall that the Hamiltonian is (with $Z = 2$):

$$\hat{H} = -\frac{\hbar^2}{2M}\Delta_1 - \frac{Ze^2}{r_1} - \frac{\hbar^2}{2M}\Delta_2 - \frac{Ze^2}{r_2} + \frac{e^2}{r_{12}}$$

i.e., in a.u.

$$\hat{H} = -\frac{\Delta_1}{2} - \frac{Z}{r_1} - \frac{\Delta_2}{2} - \frac{Z}{r_2} + \frac{1}{r_{12}}$$

Let's choose now the non-interacting part to be our zero order Hamiltonian, and the electron-electron interaction is the perturbation, i.e., write

$$\hat{H} = \hat{H}_0 + \hat{H}'$$

where

$$\hat{H}_0 = -\frac{\Delta_1}{2} - \frac{Z}{r_1} - \frac{\Delta_2}{2} - \frac{Z}{r_2}$$

$$\hat{H}' = \frac{1}{r_{12}}$$

Now \hat{H}_0 is the sum of two hydrogen-like Hamiltonians. We know the eigenfunctions and eigenvalues of \hat{H}_0 . The eigenfunctions are products of hydrogen-like eigenfunctions (e.g., $\psi_{1s}(\mathbf{r}_1)\psi_{1s}(\mathbf{r}_2)$) and the eigenenergies are sums of hydrogen-like energies. Thus, we have the zeroth-order eigenfunctions (f) and eigenvalues (denoted as ϵ) for the problem:

$$f_{n_1, l_1, m_1, n_2, l_2, m_2}(\mathbf{r}_1, \mathbf{r}_2) = \psi_{n_1, l_1, m_1}(\mathbf{r}_1)\psi_{n_2, l_2, m_2}(\mathbf{r}_2)$$

So $H_0 f = \epsilon f$, i.e.,

$$H_0 f_{n_1, l_1, m_1, n_2, l_2, m_2} = \epsilon_{n_1, n_2} f_{n_1, l_1, m_1, n_2, l_2, m_2}$$

where the zero-order energies (i.e., energies associated with the zero-order Hamiltonian) are the sum of the individual energies

$$\epsilon_{n_1, n_2} = -\frac{Z^2 e^2}{2a_0} \left(\frac{1}{n_1^2} + \frac{1}{n_2^2} \right), \quad n_1, n_2 = 1, 2, 3, \dots$$

Let's determine the:

1st order perturbation to the ground-state energy

The zeroth-order ground-state eigenfunction is $\phi_{1s}(r_1) \cdot \phi_{1s}(r_2)$, i.e., formally

$$\begin{aligned} u_{000,000}(\mathbf{r}_1, \mathbf{r}_2) &= \phi_{000}(r_1) \cdot \phi_{000}(r_2) \\ &= \left(\frac{1}{\pi^{0.5}} \left(\frac{Z}{a_0} \right)^{\frac{3}{2}} e^{-\frac{Zr_1}{a_0}} \right) \cdot \left(\frac{1}{\pi^{0.5}} \left(\frac{Z}{a_0} \right)^{\frac{3}{2}} e^{-\frac{Zr_2}{a_0}} \right) \end{aligned}$$

And the associated energy is

$$\epsilon_{0,0} = -\frac{Z^2 e^2}{2a_0} \left(\frac{1}{1} + \frac{1}{1} \right) = -\frac{Z^2 e^2}{a_0} = -Z^2 \text{ Hartree}$$

The correction to 1st order to the ground-state energy is denoted as $W_1^{1s^2}$

$$W_1^{1s^2} = \langle u_{000,000} | \hat{H}' | u_{000,000} \rangle = \int u_{000,000}^*(r_1, r_2) \frac{1}{r_{12}} u_{000,000}^*(r_1, r_2) d^3\mathbf{r}_1 d^3\mathbf{r}_2$$

i.e.,

$$\begin{aligned} W_1^{1s^2} &= \left(\frac{1}{\pi} \left(\frac{Z}{a_0} \right)^3 \right)^2 \int e^{-\frac{Zr_1}{a_0}} e^{-\frac{Zr_2}{a_0}} \frac{1}{r_{12}} e^{-\frac{Zr_1}{a_0}} e^{-\frac{Zr_2}{a_0}} r_1^2 dr_1 r_2^2 dr_2 \sin\theta_1 d\theta_1 d\phi_1 \sin\theta_2 d\theta_2 d\phi_2 \end{aligned}$$

We actually have seen this integral previously in conjunction with the variational treatment of the ground-state; I remind you that there we wrote:

$$\langle \Phi | \hat{H}' | \Phi \rangle =$$

$$\begin{aligned} &\left[\frac{1}{\pi^2} \xi^6 \right] \iint r_1^2 dr_1 \sin(\theta_1) d\theta_1 r_2^2 dr_2 \sin(\theta_2) d\theta_2 * \exp(-2\xi r_1) \exp(-2\xi r_2) \\ &* \left(\frac{\xi - 2}{r_1} + \frac{\xi - 2}{r_2} + \frac{\mathbf{1}}{r_{12}} \right) \end{aligned}$$

where I highlight in red the term we need now, and in the present perturbative calculation we don't vary the effective charge, i.e., here $\xi = Z = 2$ is the charge of the Helium, so the two " $\xi - 2$ " in the integral drop-out.

So we'll use the result we quoted there, which was

$$W_1^{1s^2} = \frac{5}{8} Z \frac{e^2}{a_0} = \frac{5}{8} Z \text{ Hartree}$$

So adding to the zero-order energy, $\epsilon_{00} = -Z^2$ Hartree as mentioned earlier, we get the total ground-state energy to 1st order:

$$W^{1s^2} \simeq \epsilon_{00} + W_1^{1s^2} = -Z^2 + \frac{5}{8} Z$$

So for He, $Z=2$, and we get

$$\begin{aligned} W^{1s^2} &= \left(-2^2 + 5 * \frac{2}{8} \right) \text{ Hartree} = -4 \text{ Hartree} + \frac{5}{4} \text{ Hartree} \\ &= (-108.83 + 34.01) \text{ eV} = -74.82 \text{ eV} \end{aligned}$$

Note:

- 1) In this case (and usually) perturbation theory to 1st order is not nearly as good as a comparable variational calculation. There is no flexibility built into the perturbation-theory wavefunction, except for the initial choice of \hat{H}_0 and therefore the associated zero-order eigenfunctions.
- 2) Going into higher orders in perturbation theory requires systematically more in the way of calculations. Usually we'll need to go to 2nd order, beyond that the results may actually get worse in many cases.
- 3) We only treated here the non-degenerate case, meaning that the zero-order eigenvalue is not degenerate (or more precisely, that the strength of the perturbation is somehow smaller than the difference between the energy of the state we're looking at, in this case the ground-state, and that of nearby eigenstates.)

The degenerate perturbation theory case is more important but more complicated, probably we won't treat it.

- 4) A time-dependent version of perturbation theory is readily formulated and is very powerful in studying time-dependent problem, most of all

in the treatment of the absorption of radiation. We may have enough time to cover that topic (next), or you may need to wait to 113B.